Network analysis of a tourism destination

by
Rodolfo Baggio

Thesis submitted for the degree of
Doctor of Philosophy

School of Tourism
THE UNIVERSITY OF QUEENSLAND
Queensland, Australia
July 2008

That in which we mainly need to insist is on persuading our principals that this is a complete art, although just born, based on principles and means that are new, but noble and commendable, and needs to be embraced, cultivated and promoted, so that with exercise and time it will be possible to benefit from the fruits of which it has in itself the seeds and the roots.

Galileo, Letter to Orso d'Elci
Florence, 25 December 1617

Abstract

The nature of development and evolution of a tourism destination has profound implications not only for public authorities, as an aid in control and planning, but also for all the destination stakeholders. A tourism destination is analysed here as a dynamic evolving complex system with the series of techniques and methods drawn from the area of network analysis.

Recent research has shown that the topology of a network has predictable properties that may greatly affect the overall dynamic behaviour and explain and influence a number of processes from the diffusion of ideas to robustness to external attacks and the optimisation of the relationships among the network components. These analysis techniques can be considered a diagnostic method for collecting and analysing data about the patterns of relationships among people in groups or among organisations. They provide a view into the network of relationships that may give tourism organisation managers a strong leverage to improve the flow of information and to target opportunities where this flow may have the most impact on regulatory or business activities.

The main objective of this thesis is to apply methods and techniques of so called 'network science' in order to study the evolution of the destination system and to simulate dynamic processes such as information and knowledge diffusion, and efficiency optimisation.

The first part of the thesis concerns the theoretical background for this work. The most recent results of the investigations on network theories are reviewed. Both static and dynamic characteristics of a network and of the main processes which may occur (robustness with respect to external or internal shocks and diffusion of knowledge and information) are discussed. The models and the metrics presented have been chosen from the larger literature on the subject based upon their recognised importance for the assessment of the structural characteristics of a network, or because they are deemed particularly suited for the objectives of this study.

The tourism destination is then discussed as a peculiar form of industrial district. Industrial economics are presented together with the main issues on the structure and the evolution of a cluster of firms. In this framework the characteristics of a tourism destination are

examined and the current ideas on its possible evolutionary growth are taken into account. These discussions set the composite and multidisciplinary theoretical framework for this thesis.

A specific methodology for the study conducted is then presented. The methodology chapter contains the account of the reasons for the choice of the test case: the island of Elba, Italy. Furthermore it reports on the issues related to the collection of the data needed and their assessment from a completeness and reliability point of view. Two networks are built: the network of relationships between the tourism stakeholders and one of the hyperlinks among their websites. Specific algorithms and measurement techniques for the main static and dynamic characteristics of the destination network are detailed.

The results of the study are then presented. They concern the measurement of the topological features of both networks (the real and the virtual) and the comparison between them. A simulation study then allows the robustness of the system and its responses to the process of diffusion of knowledge and information to be examined. The simulation also allows the influence of possible modifications in the network topology on these processes to be determined. A reconstruction of the past structure of the destination is finally used to compare network characteristics at different times. This comparison is employed to present a model of destination evolution based on its dynamic network topology.

The last part of the thesis contains the discussion of the results. Their interpretation is presented first from a 'network science' viewpoint. The implications of this interpretation for the management of the destination are discussed by highlighting the contribution of this work for the scholarly study of the subject and by providing insights into the possible usage of the methodology by practitioners (destination managers or single stakeholders).

Keywords: tourism destinations, economic and social integration, socio-economic systems, complex systems, network analysis.

Candidate's statement of originality

I hereby declare that this thesis, submitted in fulfilment of the requirements for the degree of Doctorate of Philosophy and entitled 'Network analysis of a tourism destination', represents my own original work and has not been previously submitted, either in whole or in part, to this or any other institution for any degree, diploma or other qualification.

To the best of my knowledge and belief, this work contains no material published or written by any other person except as acknowledged in the text.

Rodolfo Baggio
Brisbane, July 2008

Prof. C. Cooper
Principal advisor

Acknowledgments

It is commonly believed that a PhD thesis is the work of a single person. The author is even asked to sign an official statement in this regard. This is only partially true. During the hard work that leads to the final thesis, even if they are unaware, quite a large number of people are involved. Some give profound insights on a specific topic, some discuss ideas or interpretations, some offer counselling and advice in the inevitable sad moments that may occur, or provide some essential logistic support.

And then, there are all those who have contributed in shaping my mind and attitude towards study and research, even from long ago, since the very beginning of my education. Probably, they do not know or do not think that what they did was important, but they play a crucial role in the success of the whole enterprise.

I shall not, like many do, write here a long list of names. I am too worried about forgetting someone. So, to all those wonderful people who helped me, thank you! I cannot find other ways to acknowledge all, or even any of you properly.

I make only a very few exceptions. As well as the whole family, who tolerated my efforts with incredible patience, I want to express my gratitude to Magda Antonioli, the volcanic director of the Master in Economics and Tourism at Bocconi University. Her friendship and her support have been, and still are, beyond any possible rationalisation.

Two more exceptions are for Chris Cooper and Noel Scott.

Expressing acknowledgments to supervisors appears to be a courteous customary habit. The reader may think the same in the present case. Chris and Noel know this it is not just a formal expression of appreciation. And I am not able to find the right words to communicate my feelings. So, again: thanks.

Foreword

The main motivation of the work described in the following pages may be traced back to my personal history.

After having attended a classical high school, I graduated in physics with an experimental thesis on the detection of high energy cosmic rays in the atmosphere. For a couple of years I then worked at the Institute of Physics of the Milan University, researching in the fields of astrophysics and radio astronomy. After that, I spent almost twenty years working in the Information Technology industry dealing with both technical and organisational aspects of information science.

A few years ago I started collaboration with Bocconi University, teaching on technological topics. In this position I happened to meet Magda Antonioli Corigliano and her staff: the group responsible for the Master in Economics and Tourism. There, besides teaching, I conducted a number of research projects on the relationships between tourism and information and communication technologies (ICT).

My approach to the study of the various subjects I have dealt with has always been strongly influenced by my background and by the idea that a mixture of different disciplines and different experiences can greatly help in understanding a phenomenon, whatever the field of study is.

In my early university years I had the luck and the honour to meet Ludovico Geymonat, probably the greatest living Italian philosopher of science. We had only a few occasions to discuss ideas but I still remember his words. In his monumental *Storia del pensiero filosofico e scientifico* (History of Philosophical and Scientific Thought, 1972) he writes (vol. 6, p. 1042):

> … one of the most serious crises of modern culture is, without doubt, connected to the affirmation of the specialist spirit in almost every scientific discipline: and not only – it is worth noting – in those of the sciences specifically called as such (mathematics, physics, biology etc.) but also in research fields such as history, philology, economy and so forth.

and, a little further on in the same book, speaking of methodology as a first step to overcome excessive specialism (pp. 1045-1046):

> […] recently, the same scientists have started to realize the negative effects of pure specialism. One of the factors that has mostly contributed to this awareness has been, without doubt, a renewed interest in the methodological problems. […] One more important effect of the methodological reflection has been the stimulus impressed on comparing the structures of the different sciences discussing the legitimacy of reducing a theory (secondary) to another one (thought primary), or also only of using the results attained in one of them outside the field in which they had been demonstrated; to specify scrupulously analogies and differences – sometimes well hidden – of the principles produced in support of single theories; to evaluate the meaning of generalisations etc.

Since then, I have always thought that the efforts of the scientific community towards better and more effective methods of analysis cannot be those of isolated and overspecialised groups of researchers, but that the collaboration of different, and even conflicting, knowledge realms is the only viable way to achieve important results.

Some years ago, I stumbled, serendipitously, upon a set of papers in which 'familiar' methods were applied to the study of complex networks and upon a strange new discipline named *econophysics* (Mantegna & Stanley, 2000) and I decided to go into more depth in this subject. Working in a Master in Tourism program, it was natural to think of applying these methods to this field of study.

One more fortuitous, and lucky, encounter with Chris Cooper, then Head of the School of Tourism at The University of Queensland, gave me the opportunity of embarking on the adventure of a PhD. Our common interest in the structure of a tourism destination and the relationships among its stakeholders led, almost naturally, to this work.

This thesis may look a little perplexing to many readers.

Typically, many PhD researchers (and dissertations) in social sciences, economics and tourism start with some clear questions and a definite theoretical framework and propose reconsiderations, improvements or empirical trials to endorse it (or to invalidate it). This thesis is different from the 'usual' and may, at first sight, confuse the reader with a strong background in these social science disciplines. But this work did not start with a firmly set framework. It started, rather, with the idea of applying a number of methods and techniques from different fields (mathematics, graph theory, physics, economics etc.) to one object of study (a tourism destination) and trying to find a way to use these to improve our ability to describe a destination and to have better means to follow its evolution.

So, as in many other disciplines, and mainly in physics, my main objective was to build a theory, or better (more modestly) a valid methodological path and to verify whether the results obtained are meaningful.

The basis on which I moved was to use the analogy of a complex network for a tourism destination and to apply to the latter the considerations and the models that are used with the former. This may look like a disputable way of proceeding, but shares a long tradition in the physical sciences. Faraday, Coulomb, Helmoltz, Maxwell, Boltzmann, just to name a few examples, moved along this path. An initial analogy served them as an aid in the development of their ideas.

The reader will judge the final result. I can only hope that my small contribution is of some help in drawing better and more useful pictures of the phenomena studied.

As a final remark, as a scientist, I obviously know that this is only a first step in a long and problematic route and that everything contained in these pages can be upset by different or newer empirical evidence. Therefore, I present my results as a working hypothesis.

Neque quisquam, quod ad hypotheses attinet, quicquam certi ab astronomia expectet, cum ipsa nihil tale praestare queat, ne si in alium usum conficta pro veris arripiat, stultior ab hac disciplina discedat quam accesserit. Vale.

Nicolaus Copernicus, De Revolutionibus Orbium Caelestium, 1543

Publications by the candidate relevant to this thesis

Network Analysis and Tourism: From Theory to Practice.

Scott, N., Cooper, C., & Baggio, R. (2008), Clevedon, UK: Channel View.

Abstract

This book aims to provide a comprehensive review of the contribution of network analysis to the understanding of tourism destinations and organizations. Theoretical and methodological aspects are discussed along with a series of applications. While this is a relatively new approach in the tourism literature, in other social and natural sciences network analysis has a long tradition and has provided important insights for the knowledge of the structure and the dynamics of many complex systems. The study of network structures, both from a quantitative and qualitative point of view, can deliver a number of useful outcomes also for the analysis of tourism destinations and organizations.

Knowledge Management and Transfer in Tourism: An Italian Case.

Baggio, R., & Cooper, C. (2008), *Proceedings of the IASK Advances in Tourism Research 2008 (ATR2008), Aveiro, Portugal, 26-28 May, 44-53.*

Abstract

In the twenty first century tourism destinations have an imperative to innovate and remain competitive in an increasingly global competitive environment. A pre-requisite for innovation is the understanding of how destinations source, share and use knowledge. This chapter examines the nature of networks and how their analysis can shed light upon how destinations can share and benefit from knowledge as they strive to innovate and be competitive. The chapter outlines the notion of destinations as networks of connected organisations, both public and private sector. These organisations can be considered as stakeholders of the destination, and in network theory represent the nodes within the system. The chapter outlines the use of epidemic diffusion models and shows how they can act as a analogy for knowledge communication and transfer within a destination network. These mathematical models can be combined with other network analysis approaches to shed light on how destination networks operate, and how they can be optimised with policy intervention. The chapter closes with a practical tourism example taken from the Italian destination of Elba. Using simulations the case demonstrates how the Elba network can be optimised. Overall this chapter demonstrates the considerable utility of network analysis for tourism.

Symptoms of complexity in a tourism system.

Baggio, R. (2008). *Tourism Analysis*, *13*(1), 1-20.

Abstract

Tourism destinations behave as dynamic evolving complex systems, encompassing numerous factors and activities which are interdependent and whose relationships might be highly nonlinear. Traditional research in this field has looked after a linear approach: variables and relationships are monitored in order to forecast future outcomes with simplified models and to derive implications for management organisations. The limitations of this approach have become apparent in many cases, and several authors claim for a new and different attitude.

While complex systems ideas are amongst the most promising interdisciplinary research themes emerged in the last few decades, very little has been done so far in the field of tourism. This paper presents a brief overview of the complexity framework as a means to understand

structures, characteristics, relationships, and explores the implications and contributions of the complexity literature on tourism systems. The objective is to allow the reader to gain a deeper appreciation of this point of view.

Destination Networks - Theory and practice in four Australian cases.

Scott, N., Cooper, C., & Baggio, R. (2008). *Annals of Tourism Research*, *35*(1), 169-188.

Abstract

Tourism involves a network of organizations interacting to produce a service. In this paper, network analysis is used to examine the structural properties of inter-organizational networks within tourism destinations. Network analysis is particularly useful as it adopts a whole of destination approach and imposes predefined groupings on the organization of tourism in a region. Information flows between key organizations provide the basis for analyzing a destination's organizational structure and linkages, allowing identification of strategic weaknesses in the cohesiveness of the destination that can be addressed by policy and management approaches. The paper outlines four Australian destination case studies that demonstrate the utility of network analysis by illustrating structural features such as product clusters, structural divides and central organizations.

What network analysis can reveal of tourism destinations [Poster].

Baggio, R. (2007). *Presented at the Complex Networks: from Biology to Information Technology, Pula (CA, Italy) 2-6 July*.

Abstract

Tourism is probably the largest economic sector of the World's economy, and a tourism destination is considered to be a fundamental unit of analysis for the understanding of this industry. This research examines two such systems, Fiji islands and the island of Elba (Italy), investigating their structural characteristics. Network theoretic metrics are used to gauge the static and dynamic attributes of the networks formed by the websites belonging to the different tourism operators. The general topology is found only partly similar to the one peculiar to many complex socio-economic systems. The differences appear due to the rather poor connectivity and clusterisation of the networks. The structural characteristics are then interpreted in terms of the evolutionary growth of a tourism destination.

Tourism destinations: a network analysis of the Web space.

Baggio, R., & Antonioli Corigliano, M. (2007). *Proceedings of the Advances in Tourism Marketing Conference (ATMC), Valencia, Spain, 10-12 September*.

Abstract

Complex networks have been extensively studied in the last few years, and an interdisciplinary community of researchers have proposed and investigated a whole series of models with the objective of explaining the relationship between the structural characteristics of a network, its functions and its dynamic evolution. This paper aims at examining how network thinking can help in understanding the interactions between tourism stakeholders within a destination.

Quantitative network analysis methods are used to compare the network characteristics of two tourism destinations: the Fiji Islands and Elba, Italy. The main structural characteristics are measured, both from a static and a dynamic point of view. The network characteristics are

compare with some distinctive features of the destinations and with possible evolutionary models.
The results show that network topological measurements can be used to highlight the main features of the destination network and of the underlying social and economic system and can provide a way to assess their evolutionary history.
The paper proposes an original quantitative approach to the study of tourism destinations and to the relationships between their stakeholders.

Use of network analysis in tourism research.

Scott, N., Cooper, C., & Baggio, R. (2007). *Proceedings of the Advances in Tourism Marketing Conference (ATMC), Valencia, Spain, 10-12 September*.

Abstract

This paper discusses the status of the network concept in tourism, examining the historical development of network thinking in the wider literature and the origins of network thinking in sociology, anthropology as well as mathematics. It then examines the usefulness of the network concept for the study of tourism and reviews a number of different applications in tourism research reported in three tourism journals over a six year period. Based on this analysis and previous typologies of approaches to the study of networks, the paper develops a fourfold topology of types of network research in tourism.
The paper reviews recent network research in tourism using a convenience sample of refereed journal articles from Current Issues in Tourism, Tourism Management and Annals of Tourism Research between 2000 and 2006 and categorises these on a fourfold typology.
The paper highlights a number of research areas where the authors consider that further application of the network concept would be of benefit. These areas include application of complexity and chaos theory and the study of the tourists' networks of friends and acquaintances that influence tourist behaviour.
The review of the tourism literature is limited to three journals and a six year period.
The paper develops an original typology of network studies in tourism and discusses complexity theory and the application of physical network concepts in social systems.

The websites of a tourism destination: a network analysis.

Baggio, R., Antonioli Corigliano, M., & Tallinucci, V. (2007). In M. Sigala, L. Mich & J. Murphy (Eds.), *Information and Communication Technologies in Tourism 2007 - Proceedings of the International Conference in Ljubljana, Slovenia* (pp. 279-288). Wien: Springer.

Abstract

How to reveal throughout a quantitative survey of the websites the cooperation development of a tourism destination? The analysis of the links among the tourist websites of a destination allows it. The website network of a tourism destination is examined and the main statistical characteristics of the underlying graph are calculated. The general topology of the network is similar to the one characterizing such systems. However, differences are found, mainly due to the relatively poor connectivity among the vertices of the network. These results are discussed and interpreted. Moreover, the usage of quantitative network parameters is shown to be able to measure the degree of collaboration and cooperation among the destination stakeholders and their tendency to act in such a way.

What network analysis of the WWW can tell us about the organisation of tourism destinations.

Baggio, R., Scott, N., & Wang, Z. (2007). *Proceedings of the CAUTHE 2007, Sydney, Australia, 11-14 February.*

Abstract

This research examines the comparative characteristics of two tourism destinations based on quantitative analysis of the network of hyperlinks among their tourism operators' websites. Methods and techniques of the 'science of networks' are used to characterise and compare the static and dynamic characteristics of a tourism destinations webspace. Network metrics are proposed as quantitative assessments of collaboration and cooperation among destination stakeholders. Using two cases, Fiji islands and the island of Elba (Italy), we show that the networks exhibit a topology similar to that characterising many complex social systems which have been studied and reported in the literature. The differences found of these tourism networks from other types appear due to the relatively poor connectivity among the elements of the networks. The structural characteristics are then interpreted in terms of the evolutionary growth of tourism destinations.

Destination management plans: use of language as representation of power.

Baggio, R., & Marzano, G. (2007). *Proceedings of the CAUTHE 2007, Sydney, Australia, 11-14 February.*

Abstract

The analysis of discursive practices and language is critical to understand power relationships and power struggles. In particular, control over orders of discourse by institutional and societal power-holders is one factor in the maintenance of their power. This paper analyses how Tourism Queensland used political language and symbols to legitimate and generate support for its decisions. The destination management plans, developed for the 13 Queensland regions, were analysed using an automated content analysis software combined with quantitative network theoretic measurements. These methods prove to be effective in identifying the main concepts contained within the documents and their interrelationships. Results show that the core concepts of the destination management plans relate to the social impact that tourism has on communities as well as to the importance of the marketing and promotion of tourism. On the other hand, some concepts on which academics have focused, such as sustainability and collaboration, do not appear core elements in how Tourism Queensland prescribed tourism to be managed through its official discourse.

The Web Graph of a Tourism System.

Baggio, R. (2007). *Physica A, 379*(2), 727-734.

Abstract

The website network of a tourism destination is examined. Network theoretic metrics are used to gauge the static and dynamic characteristics of the webspace. The topology of the network is found partly similar to the one exhibited by similar systems. However, some differences are found, mainly due to the relatively poor connectivity and clusterisation of the network. These results are interpreted by considering the formation mechanisms and the connotation of the linkages between websites. Clustering and assortativity coefficients are proposed as quantitative estimations of the degree of collaboration and cooperation among destination stakeholders.

Complex systems, information technologies and tourism: a network point of view.

Baggio, R. (2006). *Information Technology and Tourism*, *8*(1), 15-29.

Abstract

There is a growing interest in complexity science as a framework for understanding social and economic systems. This paper aims at presenting this approach giving a brief overview of the complexity framework and illustrating some of the methods in order to allow the reader to gain a deeper appreciation of this perspective. The role of information management and information technology in tourism, emphasized on numerous occasions, is examined in this context. It is argued that this framework can offer tools and techniques able not only to better understand the general state from a theoretical point of view, but can also provide practical guidance in specific situations. As an example, the structure of the community of websites belonging to Italian travel agencies is analyzed.

Table of contents

Abstract .... **v**

Candidate's statement of originality .... **vii**

Acknowledgments .... **ix**

Foreword .... **xi**

Publications by the candidate relevant to this thesis .... **xv**

1 Introduction **1**

1.1 Preliminary framework to the research questions 3
1.2 Objectives and research questions 6
1.3 The approach of this thesis 6
1.4 Delimitations of the study 10
1.5 Outline of the thesis 11

2 Theoretical background: models of complex networks **13**

2.1 Complexity and complex systems 15
2.2 Tools for the analysis of complex systems 21
2.2.1 Nonlinear dynamics 21
2.2.2 Statistical physics 22
2.2.3 Network theory 24
2.3 Elementary graph theory 24
2.3.1 A brief historical sketch of graph and network theories 24
2.3.2 Graphs: main definitions 29
2.3.3 Graphs of social networks 36
2.3.4 Graph spectra 37
2.4 Random networks 42
2.4.1 The Erdős-Rényi model 42
2.5 Small-world networks 46
2.5.1 Features of small-world networks 49
2.6 Scale-free networks 55
2.6.1 Features of scale-free networks 58
2.6.2 Graph spectra 62
2.7 Conclusions 65

3 Theoretical background: network dynamics **71**

3.1 Basic graph evolution 72
3.1.1 Dynamic scale-free networks 75
3.2 Dynamic processes on complex networks 81
3.2.1 Robustness and fragility 81
3.2.2 Diffusion in complex networks 84
3.2.3 Self-organised criticality 88
3.2.4 Network optimisation 93
3.2.5 Communities and community discovery 98
3.3 The structure of the WWW 101
3.4 Conclusions 104

4 Theoretical background: industrial economics and tourism destinations 107

4.1 Industrial clusters, districts and networks 109
4.1.1 Structure and evolution of an industrial cluster 113
4.1.2 Cluster economics and benefits 121
4.1.3 Clusters and districts 125
4.2 Tourism destinations 126
4.2.1 The tourism destination as an industrial district 129
4.2.2 The management of a tourism destination 134
4.2.3 Structure and evolution of a tourism destination 137
4.2.4 The contribution of network analysis 148
4.3 Conclusions: a theoretical framework for the thesis 151

5 Methods 153

5.1 The research path 155
5.2 The approach to this work: an epistemological digression 158
5.2.1 The epistemology of a physicist 158
5.2.2 The use of analogy 161
5.2.3 Physics and socio-economic systems 163
5.3 The case study 165
5.3.1 Elba Island 170
5.4 Data collection 174
5.4.1 Sampling issues 176
5.4.2 The empirical data 180
5.5 Network analysis 185
5.5.1 Static characteristics 186
5.5.2 Dynamic processes 188
5.5.3 System evolution 193
5.5.4 The technological network 195
5.6 Research boundaries 195
5.7 Ethical issues 197
5.8 Conclusions: the methodological approach 198

6 Results 201

6.1 The Elba network, a static characterisation 202
6.2 Data completeness evaluation 209
6.3 The technological network 210
6.3.1 The technological and socio-economic networks: a comparison 213
6.4 Dynamic processes 216
6.5 Historical evolution of the tourism destination network 221
6.5.1 A network based assessment of the destination evolution 223
6.6 Conclusions: a summary of the main results 227

7 Discussion and interpretation 229

7.1 The Elba tourism destination network: a static characterisation 230
7.1.1 Assessing collaboration and cooperation characteristics 232
7.1.2 Modularity features 236
7.2 The technological network 238
7.2.1 The virtual and the real network, a comparison 240

7.3 Dynamic processes ... 243
7.4 Simulated evolution of the Elba destination ... 247
7.5 Network thinking: implications for management and governance ... 252
7.6 An answer to the research questions ... 254

8 Conclusions ... 257

8.1 A summary of the thesis ... 257
8.1.1 Research limitations and boundaries ... 258
8.2 Contributions of this thesis and further research ... 259
8.3 A concluding remark ... 262

References ... 267

Appendix 1: Main definitions ... 295

Appendix 2: Ethical clearance ... 299

List of figures

Figure 2.1 The Königsberg bridges (source: Euler, 1736) 25

Figure 2.2 Königsberg (from a 1894 map online at http://www.koenigsberg-stadtplan.de.vu/) 26

Figure 2.3 An undirected graph. 31

Figure 2.4 Graph types: a, b and c are simple graphs, (a) is undirected and unweighted, (b) is directed, (c) is weighted; (d) is a multigraph containing a loop. 31

Figure 2.5 An undirected graph and its adjacency matrix 33

Figure 2.6 A directed graph and its adjacency matrix 34

Figure 2.7 The formation of an ER random graph 43

Figure 2.8 Degree distribution for a graph with N=10 000 and p=0.0015 compared with the expected value of a Poisson distribution (after Albert & Barabási, 2002) 45

Figure 2.9 Generation of a small-world network obtained by rewiring (a) existing links or adding (b) new edges (after Dorogovtsev & Mendes, 2002) 48

Figure 2.10 Clustering coefficient and average distance in SW networks (after Watts & Strogatz, 1998) 49

Figure 2.11 Crossover size for different dimensions of the lattice (after Albert & Barabási, 2002) 50

Figure 2.12 Clustering coefficient vs *p* for networks with different *N* (after Barrat & Weigt, 2000) 51

Figure 2.13 Degree distribution in SW networks for *k=3*, *N=1000* and for a corresponding ER network (full circles) (after Albert & Barabási, 2002) 52

Figure 2.14 Local and global efficiencies as functions of the rewiring probability (after Latora & Marchiori, 2001) 54

Figure 2.15 Simulation of E_{glob}, E_{loc} and Cost (after Latora & Marchiori, 2001) 55

Figure 2.16 Examples of degree distributions of real world networks (after Albert & Barabási, 2002) 56

Figure 2.17 Distributions: Poisson, exponential and power-law (drawn on a log-log diagram) 57

Figure 2.18 Stochastic generation of a SF network 58

Figure 2.19 Deterministic generation of a SF network 58

Figure 2.20 Average distance *l* as function of *N*, a comparison between SF and ER models with same size and mean degree (after Albert & Barabási, 2002) 59

Figure 2.21 Clustering coefficient in SF and ER networks (after Albert & Barabási, 2002) 60

Figure 2.22 P(k) behaviour for assortative networks at different degree correlation probabilities (after Catanzaro et al., 2004) 61

Figure 2.23 Betweenness distribution and relation with *k* (after Catanzaro et al., 2004)...62

Figure 2.24 Spectral density for a random graph (N = 1024). The peak corresponds to the principal eigenvalue (after de Aguiar & Bar-Yam, 2005) 63
Figure 2.25 Spectral densities for a SW network with *p=0*, *p=0.01*, *p=0.3* and *p=1* (after Albert & Barabási, 2002) 64
Figure 2.26 SF network spectral density (after de Aguiar & Bar-Yam, 2005)....64
Figure 3.1 Size of the components and of the giant cluster in an ER network (after Newman, 2003b) 74
Figure 3.2 Power-law degree distributions for networks with different sizes (adapted from Baiesi & Manna, 2003) 75
Figure 3.3 Deviation from a pure power-law (solid line) of the degree distribution due to age (a) or cost (b) constraints (after Amaral et al., 2000) 77
Figure 3.4 Fit-gets-rich phenomenon and Bose-Einstein condensation analogy (after Bianconi & Barabási, 2001) 80
Figure 3.5 Maximum degree in a SF network as function of temperature T (after Bianconi & Barabási, 2001) 80
Figure 3.6 Simulation of the effects of malfunctions (random removals, solid line) and targeted attacks (dotted line) for random (exponential and ER) networks and scale-free networks (after Dorogovtsev & Mendes, 2002)....83
Figure 3.7 Local and global efficiency as a function of random and targeted removal of a fraction *p* of nodes in different networks (after Crucitti et al., 2003)....84
Figure 3.8 Infection model and critical threshold (after Pastor-Satorras & Vespignani, 2003)....85
Figure 3.9 Diffusion (infection) in SF networks (solid line) compared to homogeneous random networks (ER, dotted line) (after Pastor-Satorras & Vespignani, 2003)....86
Figure 3.10 Immunisation techniques for SF networks at a fixed spreading rate (after Pastor-Satorras & Vespignani, 2002) 87
Figure 3.11 Survival probability of an infection for λ=0.065 and N from ~10^3 to ~10^6 (top to bottom) (after Pastor-Satorras & Vespignani, 2003) 87
Figure 3.12 The sandpile giving the name to self-organised criticality model (drawing by E. Wiesenfeld in Bak, 1996: 2)....88
Figure 3.13 Size and lifetime distributions for avalanches in small-world networks different sizes (L) and rewiring probability p=0.3 (after de Arcangelis & Herrmann, 2002) 90
Figure 3.14 Largest component G as function of α in a SF network (squares: random node removal, asterisks: removal of highly connected nodes, circles: removal of high load nodes) (after Motter & Lai, 2002)....91
Figure 3.15 G (α) for an ER network (after Motter & Lai, 2002)....92
Figure 3.16 The window of possibilities for avalanches as function of the average network degree and the tolerance parameter. The area included in the two curves is the one of the values for which an avalanche is possible in SF and ER networks (after Watts, 2002)....93

Figure 3.17 Density (A), energy (B), clustering coefficient (C) and distance (D) as function of λ. (C shows the clustering coefficient for an ER network, D the normalised distance for a star network d=6/(N+1) and for a poissonian network $d=\log N/\log\langle k\rangle$) (after Ferrer i Cancho & Solé, 2003)....96

Figure 3.18 Entropy H as function of λ. Different network topologies are shown (after Ferrer i Cancho & Solé, 2003)....97

Figure 3.19 Link length distribution for optimised networks (after Manna & Kabakçıoğlu, 2003)....98

Figure 3.20 In-degree and out-degree distributions for the copying mechanism (after Kumar et al., 2000b)....102

Figure 3.21 The bow-tie, a general model for the Web space (after Broder et al., 2000)....103

Figure 4.1 The life cycle of an industrial cluster....117

Figure 4.2 Determinants of cluster competitiveness (after van der Linde, 2003)....118

Figure 4.3 Most important reasons for cluster establishment (after van der Linde, 2003)....119

Figure 4.4 Marginal and average costs and benefits for new incumbents in a cluster (after Maggioni, 2004)....122

Figure 4.5 A model of a tourism district comprising clusters of diverse industries, associations and public government bodies interlinked and embedded in a 'social' territory endowed with different resources. Virtual links can tie some of the internal clusters to external groups....133

Figure 4.6 Tourism destination models (adapted from a: J. R. B. Ritchie & Crouch, 2003; b: Heath & Wall, 1992)....138

Figure 4.7 Tourism destination models (adapted from c: Dwyer & Kim, 2003; d: Viganò, 2004)....139

Figure 4.8 McKercher's model of a tourism system (adapted from McKercher, 1999)....140

Figure 4.9 Butler's model of a tourism area life cycle (after Butler, 1980)....144

Figure 4.10 Life cycle of an adaptive system (after Holling & Gunderson, 2002)....146

Figure 4.11 Influence of different scales (left) and hierarchical evolution of adaptive life cycles (adapted from Holling & Gunderson, 2002)....147

Figure 4.12 The tourism panarchy (after Farrell & Twining-Ward, 2004)....147

Figure 4.13 Schematic diagram of the theoretical background for the thesis....152

Figure 5.1 A schematic depiction of the methodology used in this work....157

Figure 5.2 The geographical position of Elba....170

Figure 5.3 Elba in a map (dated 1791) of the Napoleonic period (after Godlewska, 2002)....171

Figure 5.4 Tourism in Elba: overnight stays for the period 1954-2006 (source: Elba Tourism Board)....173

Figure 5.5 Overnight stays seasonality: 1996-2006 average (reference is year average) (source: Elba Tourism Board)....174

Figure 5.6 Effects of random removals (errors) and targeted attacks (attacks) for random (ER) and scale-free (BA) networks (f is the fraction removed) on average path length (L) and efficiency (E) of the system (after Boccaletti et al., 2006)....178
Figure 5.7 Effects of the sampling fraction (p) on the degree distribution of an exponential (ER) network (after Stumpf & Wiuf, 2005)....180
Figure 5.8 Topological entropy (H) and average path length (l) for different network topologies. The networks have same number of nodes (N = 100) and edges (M = 200) (after Demetrius & Manke, 2005)....187
Figure 5.9 Schematisation of the Metropolis algorithm used for network optimisation....192
Figure 5.10 Overnight stays at Elba. Possible structural breaks in the time series are indicated by vertical lines and marked a, b and c (data source: Elban Tourism Board)....194
Figure 6.1 The network of Elban tourism operators....202
Figure 6.2 A close-up of the central region shown in Figure 6.1....203
Figure 6.3 A network with a random distribution of links....203
Figure 6.4 Degree distribution of the Elba network....206
Figure 6.5 Cumulative degree distribution of the Elba network....206
Figure 6.6 The adjacency matrix eigenvalues spectrum of the Elba network....207
Figure 6.7 The Laplacian matrix eigenvalues spectrum of the Elba network....208
Figure 6.8 The Elban web network....211
Figure 6.9 In- and out-degree cumulative distributions for the Elba web network....212
Figure 6.10 Cumulative degree distributions for the tourism stakeholders (TN) and the tourism websites (WN) networks....214
Figure 6.11 The NetLogo diffusion model....216
Figure 6.12 Cumulative diffusion curves (Rnd = random network, EDiff = differential actors capabilities, EN = original Elba network, RW = rewired optimised network)....218
Figure 6.13 Differential diffusion curves (Rnd = random network, EDiff = differential actors capabilities, EN = original Elba network, RW = rewired optimised network)....219
Figure 6.14 Local efficiency variation as a function of random removal of edges....220
Figure 6.15 Global efficiency variation as a function of random edges removal....221
Figure 6.16 Elba overnight stays for the period 1954-2006. The three evolution phases a, b, c are discussed in the text (data source: Elba Tourism Board)....222
Figure 6.17 Evolution of the hospitality infrastructure at Elba (1970 = 100, data source: Elba Tourism Board)....224
Figure 6.18 Cumulative degree distributions for the original labelled as (E_100) and the reconstructed 1990s networks labelled as (E_90)....226
Figure 6.19 Cumulative degree distributions for the different network reconstructions (lower parts)....227

List of tables

Table 2.1 Main characteristics of some real networks from the literature (adapted from Albert & Barabási, 2002; Boccaletti et al., 2006; Caldarelli, 2007; Dorogovtsev & Mendes, 2003; Pastor-Satorras & Vespignani, 2004). Type: B = biological, S = social, T = technical; N = number of nodes, M = number of edges, δ = link density, <k> = mean degree, L = average path length, C = clustering coefficient, Eglob = global efficiency, Eloc = local efficiency, r = assortativity coefficient, γ = exponent of power-law degree distribution (undirected networks), γ-in = indegree distribution exponent, γ-out = outdegree distribution exponent (directed networks). More extensive definitions and explanations of these measurements are discussed in Chapters 3 and 4 70

Table 3.1 Average distance between web pages 104

Table 5.1 Elba hotels and bed places in the last years (source: Elba Tourism Board)....172

Table 5.2 Elba tourism statistics since 2000 (source: Elba Tourism Board) 173

Table 5.3 Effects of sampling methods on SF network properties. As the sampling fraction decreases (⇩), the quantity may increase (⇧), decrease (⇩), stay the same (⇔), or behave according to the specific network (⇳) (adapted from Lee et al., 2006) 179

Table 5.4 Components of a tourism destination 181

Table 5.5 Elba destination network actors collected for the study 181

Table 5.6 Summary of the methods used...199

Table 6.1 The values of the main metrics for the Elba network (*random* values are averages over 10 realisations)...204

Table 6.2 Modularity for the Elba network 209

Table 6.3 The main characteristics of the Elba web network compared with common values for the whole WWW...211

Table 6.4 Bow-tie components calculated for the Elba Web and general values for the WWW (as reported by Broder et al., 2000; Kumar et al., 2000b) 212

Table 6.5 Comparison between the main network metrics for the socio-economic (TN) and the web (WN) networks of the Elba tourism destination 214

Table 6.6 The Kolmogorov-Smirnov D-statistics for the different network metrics (WN = web network, TN = real tourism destination network, RN = random network)..215

Table 6.7 Size distribution of Elban tourism operators 217

Table 6.8 Network parameters used in the simulation..217

Table 6.9 Speed improvements in diffusion over different network topologies (Rnd = random network, EDiff = differential actors capabilities, EN = original Elba network, RW = rewired optimised network)..219

Table 6.10 Foundation dates for the Elban associations and consortia (Note 1: these organisations although formally founded as shown, were greatly extended and revitalised in the early 1990s; source: public records, see text) 225

1 Introduction

The last 50 years have seen a growing interest in tourism and related topics by the academic world. Having developed into the most predominant sector in world's economy, tourism could not have been ignored by the community of researchers. Explaining the phenomenon of tourism, its effects, its influence and relationships with other sectors of human activity, and attempts at forecasting future developments and behaviours have increased in importance with significant numbers of people involved in such research.

The position of researchers in the tourism field is at times difficult. Practitioners of other disciplines charge them with being too *soft*, too application oriented and of not having been able (yet) to build up a rational and uniform theoretical framework (Hall & Butler, 1995). On the other hand, people involved in day by day operational activities accuse them of *flying too high* and wasting time in fooling around with models and conjectures without producing much of practical use in helping with the problems they face. As a very recent field of investigation, tourism is still trying to find a reasonable compromise between these two extremes (Farrell & Twining-Ward, 2004; Leiper, 2000b; Tribe, 1997).

The boundaries of the tourism and travel industry are indefinite (Cohen, 1984; Cooper et al., 2005; McIntosh & Goeldner, 1990). Tourism brings together segments from a number of different activities with a wide variety of products and services exhibiting little homogeneity and with different technologies used in the production process. It may be questioned whether it should even be classified as an industry by itself in the traditional sense of manufacturing or trade (Leiper, 2000b; Mill & Morrison, 1992; Nordin, 2003). Moreover, reflecting changes in wider society, in the last few years tourism has become an extremely dynamic system. Introduction of flexible organisational structures, fast changing customer behaviour and strong impacts from the development of transportation technologies have exerted a formidable pressure on the whole sector.

Many researchers have contributed and are contributing to the growth of knowledge in the tourism domain. They are bringing into the field their multiform and diverse experiences and backgrounds. Geographers, sociologists, economists, mathematicians, ecologists and historians, are all giving their interdisciplinary and multidisciplinary contributions and trying to shape the paradigms (in the Kuhnian sense) that may raise the status of this area of study to an accepted scientific discipline (Echtner & Jamal, 1997; Farrell & Twining-Ward, 2004; Faulkner & Russell, 1997; Leiper, 2000b; Tribe, 2005).

In the vast catalogue of expertise that plays a role in this *exploration*, one was missing so far: physics.

In recent years the ideas, concepts and techniques of physics have been applied to different disciplines such as biology, economics and sociology. The results of this interdisciplinary endeavour are interesting and have helped in improving the general understanding of these fields. In particular, in the study of economics, physicists are having an increasingly important role and a new 'science' has been born: *econophysics* (Mantegna & Stanley, 2000).

Central to this work is the idea that meaningful insights can be derived by considering social actors (individuals, groups of people, companies etc.) as 'particles' of a physical system and by studying their behaviour and the effects on the whole system under investigation. This idea is far from new. Since the 17th century many scholars have taken into account the statistical properties of the elements of an economic or a social system to build the theories and models that constitute our current understanding of these (Ball, 2002, 2003). More recently, the usage of physical methods has provided important results such as the modelling of crowd behaviour (Helbing & Molnar, 1995; Henderson, 1971), traffic flows (Kerner & Rehborn, 1996; Nagel & Schreckenberg, 1992) or political elections (Bernardes et al., 2002; Costa Filho et al., 1999), the formation of business alliances (Axelrod et al., 1995; Castellano et al., 2000) and the behaviour of economic markets (Saari, 1995; Sornette, 2003). The application of the most recent developments in the field of complex systems modelling to social systems is also starting to receive 'institutional legitimacy' (Henrickson & McKelvey, 2002).

The characteristics of tourism make it a very difficult subject to define with reasonable accuracy. Terms such as complex, dynamic, networks, information intensive and others are very often used to describe the area (Baggio, 2006a; McKercher, 1999; Mill & Morrison, 1992; Werthner & Klein, 1999), so that a characterisation of tourism as *a complex dynamic* *information network business* is very likely to be generally accepted without major discussion. These words bring to a physicist's mind a series of concepts and considerations that form the main subject of this work.

The general framework in which this study is conducted is known as 'complexity science' (Lewin 1999; Waldrop, 1992). This is a rather recently formed corpus of multidisciplinary methods, and shares with tourism the common characteristic of being vaguely defined, not formalised and with disputed outcomes (Kauffman, 1995; Levin, 2003). Nonetheless, the latest results show enormous possibilities in improving our general understanding of social, economic, biological and technological phenomena (Amaral & Ottino, 2004; Arthur et al., 1997; Durlauf, 1997; Henrickson & McKelvey, 2002; Ottino, 2004; Stacey, 1996).

1.1 Preliminary framework to the research questions

The concept of a tourism destination is central to tourism.

Paraphrasing Ritchie and Crouch (2003: x) we may say that:

> While many other paradigms have been the basis of studies on tourism, we believe that, from a management perspective, the destination is the fundamental unit on which all the many complex dimensions of tourism are based. … While all of the different views of tourism are very valuable, we believe that a focus on the destination provides the integrated perspective that a destination must have if we are to comprehend, pull together and manage all the many elements that determine the success of a tourism destination.

The nature of development and evolution of a tourism destination (TD) has profound implications not only for public authorities, as an aid in control and planning, but also for all the destination stakeholders, the beneficiaries of the outcomes. The basic hypothesis is that one of the main determinants of a tourism destination's competitiveness is a balanced progress of all its components (Lozano et al., 2005; Ritchie & Crouch, 2003; Vanhove, 2005). From a structural point of view, a tourism destination can be seen as an ensemble of elements of many different sorts, many of which are connected by some kind of relationship.

A number of recent works have started to consider the characteristics of a tourism system with such a systemic approach. They make use of the ideas and the paradigms of the study of complex systems as a framework for understanding the many and different phenomena in the tourism sector (Farrell & Twining-Ward, 2004; Faulkner, 2000; Faulkner & Russell, 1997; McKercher, 1999).

Complexity is increasingly accredited as a key characteristic of the world we live in and of the way we perceive it. Although no common definition of complexity exists, it is possible to state that a system is complex when its properties cannot be derived as a simple composition (sum) of those characterising its components (Pavard & Dugdale, 2000).

Many components bounded by a series of non-linear relationships give rise to behaviours that seem unpredictable, or at least look so on the basis of simplified (linearly approximated) models of components' interactions. Basically, something is complex when equations that describe its progress over time cannot be solved analytically.

Complex systems therefore are those in which local interactions between many elements give rise to complex behaviour at the level of the whole system. Understanding these systems is a challenge faced by many different scientific disciplines, from neuroscience and ecology to linguistics and geography. Complex adaptive systems (CAS) are called adaptive because their elements respond or adapt to events around them (Levin, 2003; Lewin 1999). Also, they may form structures that somehow maintain their integrity in the face of continuing change. The components of a complex adaptive system may follow simple rules and yet produce complex patterns that often change over time. They may become more elaborately structured and develop multiple levels layered on top of each other. In other words, complex adaptive systems may evolve (Bar-Yam, 1997).

The tourism sector, as an economic activity, shares many of the characteristics that are used to define a complex adaptive system. Accidents of history, positive feedback, increasing returns, social multipliers, lock-in effects, non linearities, path dependency, evolution, self-organisation, emergence, outbreaks and catastrophes are all phenomena that have been used to explore social interaction in the emerging field of complexity. These phenomena have direct applications to tourism.

A number of tools have been developed in recent years to analyse a complex system. In their review, Amaral and Ottino (2004) identify three main classes of tools belonging to areas well known to physicists and mathematicians:

- nonlinear dynamics;
- statistical physics; and
- network theory.

Among these, the techniques belonging to what is called ‘network science’ are probably the most promising series of methods that can be used to model a complex system (Watts, 2004).

Many complex systems can be described in terms of networks of interacting elements. A number of researchers have shed light on some topological aspects of many kinds of social and natural networks (see the reviews by: Albert & Barabási, 2002; Boccaletti et al., 2006; Newman, 2003b). As a result, we know that the topology of a network is not just a curiosity, but a predictable property of some types of networks that affects their overall dynamic behaviour and explains a number of processes, from the diffusion of ideas in a social system, to the robustness to external attacks for a technological system such as the Internet, to the optimisation of the relationships among the network components and their effects on the transfer of information or knowledge.

These investigative techniques can be considered a diagnostic method for collecting and analysing data about the patterns of relationships among people in groups or among organisations. They provide a view into the network of relationships that may give tourism organisation managers leverage to improve the flow of information and to target opportunities where this flow may have a crucial impact on regulatory or business activities.

Network theory has been used only in a limited number of tourism related works (Cooper & Scott, 2005; Pavlovich, 2003a; Pavlovich & Kearins, 2004; Pforr, 2001, 2002, 2006), but none of the methods proposed here have been applied, up until now, to the study of tourism destinations.

1.2 Objectives and research questions

The purpose of the present work is to apply network analysis investigation methods to the field of tourism studies. The general theoretical framework that will be used is known as the 'science of complexity' (Waldrop, 1992). The objective is to attain a better understanding of the general behaviour and dynamic evolution of a tourism destination and to analyse the implications for the general management of a tourism destination. This understanding may support not only tourism managers in control and planning, but also the public authorities responsible for developing the infrastructure and maintaining environmental quality and sustainability.

By applying methods and techniques of 'network science', this work examines the evolution of a tourism system to give explanations of the occurrence and possible optimisations of dynamic processes such as information and knowledge diffusion.

In summary (a deeper discussion will be conducted in Chapter 5), the main research question this thesis addresses is 'How do tourism destination networks evolve over time?' Subquestions are:

- *What is the structure of the tourism destination network and how do the real and technological networks compare in a tourism destination?*
- *How does the network structure affect dynamic processes?*
- *How can the network be optimised?*
- *How can we build a model for network evolution that reproduces the present structure of the network?*
- *How does this model relate to the history of the destination?*

1.3 The approach of this thesis

The investigation presented in this thesis deals with the structure of a tourism destination network. Structure, in the social sciences literature, is seen as an enduring pattern in the components of a system and how they are assembled (Giddens, 1979; Radcliffe-Brown, 1940; Wellman, 1988, 2002). Here structure is defined (and this term will be used throughout this work in this meaning) as a combination of the components of the system (individual elements, the relationships among them) and how they are connected (the number and distribution of both elements and relationships). In this definition, structure refers both to the static elements and the dynamic processes that govern the system. The form of such a structure (its topology) is typified by the metrics described in Chapter 2; they mainly relate to the connectivity characteristics of the network.

Although it would be possible, as has been done in many other studies, to use computer simulations to model an ensemble of tourism destination networks formed by stakeholders (see for example Boccaletti et al., 2006; Dorogovtsev & Mendes, 2002; Watts, 2004), the approach followed in this work is to analyse a real case. A study conducted on a *real* destination is obviously more interesting and can give more useful findings in many cases: for example, by connecting the outcomes of the 'experiments' with real situations or by posing constraints, derived from actual conditions, to the number and type of simulations that can be implemented.

The first step in this research is the choice of a suitable tourism destination, as this underpins the quality and the validity of the results. It is preferable to deal with the smallest possible size of destination in order to ease the data collection process and for 'computational' reasons; some of the algorithms involved could take a long time to run for very large networks. On the other hand, the size should be large enough to give statistically significant outcomes.

The destination chosen as the 'test bench' is the island of Elba (Italy). The island is located three nautical miles (5 km) west of the coast of Tuscany, in the central part of the western Mediterranean Sea. It is a typical seaside summer destination. The island's economy is almost exclusively dependent on tourism. Elba is one of the most significant Italian destinations in terms of tourist flows, both domestic and international. It counts an average (past five years) of 450 000 arrivals and 3 000 000 overnights per year, with very strong seasonality (peak in July-August). The destination offers its visitors (70% of whom are domestic) more than 600 accommodation establishments (hotels plus other establishments) providing about 12 000 rooms and 36 000 bed places.

Data collection was performed by identifying, enumerating and surveying the possible types of relationships among the organisations acting in this tourism destination. This data was subsequently enhanced with any possible public sources giving similar information. Furthermore, a literature scan allowed gathering of data on comparable social systems (industrial districts, networks of corporations, etc.) which could serve as a reference, thus allowing a type of data triangulation to be used.

Public sources were also integrated with a series of field surveys whose objective was to verify the reliability of the information gathered and to discover missing elements (nodes or links in the network). Sampling issues were addressed by studying how the characteristics of the whole network were affected by missing information on nodes and links and what confidence levels could be attained in considering a reduced subset of the whole network (Marsden, 1990).

Simulation algorithms were used here that took into account available graph theoretical computational methods and the principles of statistical mechanics. Dynamic processes (information spreading, robustness with respect to external and/or internal crises, fragility to structural modifications, etc.) were then studied and the network was optimised with respect to efficiency of information or knowledge diffusion or other structural parameters. The comparison between the network topology of the tourism destination and more general social network communities made it possible to improve the optimisation algorithms in order to adapt it to better meet the expectations of stakeholders, local community and tourists.

The characteristics of the local level interactions among actors allowed us to infer, at a statistical level, some general trends or behaviours in the evolution of the tourism destination system that could be validated by looking at available data. These, along with other information useful for characterising the conditions and evolution of the destination (statistical series on tourism and basic economy, historical data etc.) were gathered and organised.

The application of the results of theoretical studies at this statistical level has several potential benefits:

- the definition of an appropriate growth model for a given network can be used to predict its extent and structure at later times;
- changes to the network's structure can be designed to reduce its vulnerability and improve its functionality by changing specific parameters through management action; and
- relevant mechanisms and parameters could be inserted into a specific network's design, by creating new structures, in order to encourage into the systems, solutions able to prevent (and to cope with) abrupt component faults, errors, attacks or external disruptions.

Policy issues are very important in guiding the development of a tourism destination, both from the 'managing body', usually termed a destination management organisation, and from the individual stakeholders' point of view (Buhalis, 2000; Framke, 2002b; Ritchie & Crouch, 2003). The analysis of the dynamic network evolution suggests the possibility of finding strategies to guide the evolution of such systems in order to meet some requirements of economic or social utility.

Another application is that the wide diffusion of the Internet may lead us to think of a correlation between the topology of the World Wide Web (WWW) and the topology of the destination. This possible relationship can be verified. By comparing the 'real' network with its 'virtual' counterpart it is possible to obtain hints to improve the efficiency and effectiveness of the investments that tourism operators make in the field of information and communication technologies (ICTs). It is also possible to examine the interorganisational network of the destination's websites (the hyperlinks among the organisations' websites) and to assess the apparent level of collaboration achieved.

Most of the results and the models on the structure of a tourism destination are derived by the analysis, mainly qualitative, of single cases. The importance of network structure, on the other hand, promotes the usage of the techniques initially tuned by sociologists and, more recently, by so-called 'network science'. Formal analysis techniques are deemed able to uncover patterns that a casual inspection may not be capable of highlighting. Network analysis also plays an important role in determining the effectiveness of the nature of the relationships (see section 4.2.4).

However, although network analysis methods are quite 'old' and tourism is a network business, little research has been conducted so far in this field in tourism (Pavlovich & Kearins, 2004; Pforr, 2001, 2002, 2006). Indeed, the network approach has proved useful in both political and strategic tourism issues (Scott et al., 2008a) and in suggesting solutions for 'practical' problems (Shih, 2006).

The proposal made here, to employ rigorous quantitative methods for the analysis of a tourism network, is thought to have important consequences for the study of a destination. It is not only an intriguing intellectual problem, but also a means to improve our ability to understand the functioning mechanisms of a tourism destination in order to manage it effectively and efficiently.

1.4 Delimitations of the study

A tourism destination system can be seen and analysed in many different ways, as can the network which can be drawn to represent it. This thesis sets forth a number of hypotheses that delimit the scope of the research.

Specific assumptions will be discussed, when needed, in Chapter 5, dedicated to methodological issues. Nonetheless, the following premises have a global validity in this study:

- the network nodes are organisations, companies, associations and the links studied are considered to be connecting these elements;
- all connections are supposed to be undirected (symmetrical) and unweighted unless explicitly stated in specific parts or for specific analyses;
- the general topological characteristics of the networks are studied. All the elements (nodes) will be considered as equivalent. Possible specific functions of single network elements are not taken into account; and
- the network and the system are studied at a macro level, and therefore all characteristics derived have the same macro scale. In other words, considering only topological characteristics means abstracting from the specific features of single

actors and connections and examining the structure of the system as a whole and the relationships it has with the dynamic processes that may unfold over it.

These are almost standard assumptions when performing an initial study of complex networks (Albert & Barabási, 2002; Boccaletti et al., 2006; Newman, 2003b). In our case they are employed to ease the computations in order to focus on the interpretation of the results and their implications for the management of a tourism destination. These assumptions, although representing a crude simplification when dealing with a socio-economic system, are acceptable for the exploratory nature of the study presented here.

1.5 Outline of the thesis

The remainder of this work is organised as follows.

Chapters 2, 3 and 4 contain a discussion of the theoretical background. In Chapters 2 and 3, the general framework of complex systems theories, and the main findings of network analysis methods are presented. These include the theories based on statistical mechanics as well as those derived from the social network analysis models. Both static and dynamic models are analysed. Next in Chapter 4 the key models of the structure of a tourism destination are described. A destination is pictured as a cluster of companies, various organisations and communities, and the analysis of its structure may draw upon the theory of industrial clusters, their formation mechanisms and their evolution. A review of the main models related to industrial clusters is therefore presented as a natural introduction to the discussion on the structure of a tourism destination.

Chapter 5 is dedicated to the methods employed for this study. First of all a more detailed description of the research questions and objectives is given. These are followed by a summary depiction of the epistemological framework in which this work was conducted. The research design, the discussion of the case chosen and the destination used as a 'test bench' are then presented.

The chapter continues with the discussion of the methods and issues related to the collection of the data needed: network elements (nodes and links) and historical statistical data. The survey methods and the questionnaire used are examined along with the main concerns about completeness of the data and possible issues pertaining to sampling in cases such as this. The methods and the tools used to analyse the network data and to perform the processes' simulations are presented and, finally, the main limitations of these methods are examined.

Chapter 6 presents the main results. It contains the findings from the static analysis of the network and gives the topological characteristics of the destination at the present time (2007). After that, the results of the simulations are discussed. The chapter continues with the results of the dynamic analysis and the reconstruction of the destination evolution in terms of network dynamics.

The next chapter (Chapter 7) contains the discussion and interpretation of the results. It closes with an examination of the methods used and a discussion of the possible methodological implications for a tourism destination in terms of policy setting, planning and stakeholder management.

The thesis ends with a further, final chapter (Chapter 8) containing a summary of the main conclusions of the study and the contributions of this study to the tourism field. Final considerations on the applicability and transferability of the approach examined in this work and possible further research directions conclude the thesis.

2 Theoretical background: models of complex networks

This and the following two chapters describe the general theoretical background for the work presented in this thesis. As discussed in Chapter 1, the main aim of this research is to provide an interdisciplinary contribution to the study of a tourism system, namely a tourism destination, by simulating and analysing the main characteristics and the possible evolutionary paths of the network formed by its stakeholders and the relationships that bind them. As a consequence, the background for this work comes from three main fields of study: one concerning complexity theory and network models, one dealing with the economic characteristics of a cluster of industries and a third pertaining to tourism destinations.

Network analysis methods are embedded into the literature of complex and chaotic systems, therefore this chapter starts with a brief introduction to these topics, highlighting the main lines of research and the methodologies used. This provides the theoretical framework in which network theories are embedded and justified. Then a review of the most important mathematical models of the topological characteristics of networks is presented. This review aims at providing a conceptual architecture and theoretical support for this thesis by grounding it on prior knowledge. As Boote and Beile (2005: 3) remark, a detailed and systematic literature review is a necessary foundation for the advancement of the collective understanding in any discipline. The analysis of strengths and weaknesses of previous studies is a necessary step to generate significant research. A critical reading of the achievements attained by those who have come before us, the process called generativity by Shulman (1999) and considered to be one of the hallmarks of scholarship, is used to grant this thesis integrity and sophistication (Shulman, 1999). It also serves the purpose of avoiding the idea, well known since Galileo (1623: 237), “that in philosophizing one must support oneself upon the opinion of some celebrated author, as if our minds ought to remain completely sterile and barren unless wedded to the reasoning of some other person”.

The recent literature on network analysis has produced a large array of metrics, each with a number of possible computational mechanisms (Boccaletti et al., 2006; Caldarelli, 2007; da Fontoura Costa et al., 2007). Those presented have been chosen either because they are generally considered to be the most effective in characterising the topological structure of a network or because they are particularly suitable for the type of analysis conducted here. The aim of this section is to present the results obtained so far and to give the basic definitions and measurement methods which will be used in the rest of the thesis.

In these two areas this chapter deals with the static investigation of a network (static here means that the network that is considered to be a an ensemble of nodes and links fixed in time and space). Chapter 3 contains the description and the discussion of the main dynamic processes occurring in complex networks and their relationships with the network topology. Growth mechanisms are described first, along with their effects on the network structure. Networks here represent dynamic systems, subject to continuous modifications in the number of elements and in the patterns of the connections among them. Therefore, a crucial issue for this thesis is the modelling of this evolutionary growth and the resulting topology. A discussion on the robustness of a network with respect to possible external events follows. A socio-economic system such as the tourism destination studied is not a closed system. It interacts, at various levels, with the surrounding environment. External (and sometimes internal) circumstances can have a great impact on the structure of the system, especially if the magnitude of the events is big enough to risk causing disruptive changes in the system. It is therefore important to investigate under which conditions the system is able to absorb the shocks and still preserve a structure able to allow its 'normal' functioning. The second part of Chapter 3 analyses the behaviour of processes such as information or knowledge diffusion and the implications that a specific topology has on the evolution of the phenomenon. Changes in the topology of the system, obtained by numerical simulations, and their effects on the processes are finally discussed.

To complete the theoretical background of the thesis, Chapter 4 contains a review of the literature of economics related to industrial clusters as an introduction to the discussion on the structure of a tourism destination.

Apart from this general framework, more specific theories and models will be discussed in the different sections they apply to.

2.1 Complexity and complex systems

In natural language, the concept of complexity has several meanings, usually related to the size and number of components in a system. There is still no universally accepted definition, nor a rigorous theoretical formalisation, of complexity. Nonetheless, it is currently a much investigated research topic.[1]

Intuitively we may characterise a complex system as “a system for which it is difficult, if not impossible to reduce the number of parameters or characterising variables without losing its essential global functional properties” (da Fontoura Costa et al., 2007; Pavard & Dugdale, 2000). The parts of a complex system interact in a non-linear manner. There are rarely simple cause and effect relationships between elements, and a small stimulus may cause a large effect, or no effect at all. The non-linearity of the interactions among the system’s parts generates a series of specific properties that characterise its behaviour as complex.

It is important to highlight the difference between a complicated and complex system. A complicated system is a collection of a number of elements (often very high) whose collective behaviour is the cumulative sum of the individual behaviours. In other words, a complicated system can be decomposed into sub-elements and understood by analysing each of them. On the contrary, a complex system can be understood only by analysing it as a whole, almost independently of the number of parts composing it.

For example, as Amaral and Ottino (2004: 147) note:

> A Boeing 747-400 has more than 3×10^6 parts. In complicated systems, such as the Boeing, parts have to work in unison to accomplish a function. One key defect (in one of the many critical parts) brings the entire system to a halt. This is why redundancy is built into the design when system failure is not an option. More importantly, complicated systems have a limited range of responses to environmental changes. […] a Boeing without its crew is not able to do much of anything to adjust to something extraordinary.

On the other hand, a ‘simple’ object made of only two elements, a double pendulum, a pendulum hanging from another pendulum, is well known to any physics student for its

---

[1] This and the following sections are partial edited versions from Baggio, R. (2006a).

totally unpredictable, chaotic behaviour (under the basic Newtonian laws of motion). A ‘simple’ school of fish, composed of a few dozen elements, is able to adapt its behaviour to the external conditions without apparent organisation but following a few simple rules regarding local interaction, spacing and velocity (Reynolds, 1987).

Generally, as Bar-Yam (1997) notes, a complex system is a mesoscopic structure, composed of a number of interacting elements which is neither too low nor too high (but even this distinction is rather confused).

In one special class of complex systems, the *complex adaptive system (CAS)*, interactions among the elements are of a dynamic nature and are influenced by, and in turn influence, the external environment. In this type of system, the parts

> interact with each other according to sets of rules that require them to examine and respond to each other’s behaviour in order to improve their behaviour and thus the behaviour of the system they comprise (Stacey, 1996: 10).

For a CAS, the main characterising features may be summarised as follows (Baggio, 2008; Levin, 2003; Waldrop, 1992):

- *non-determinism*. It is impossible to anticipate precisely the behaviour of such systems even knowing the function of its elements. The dependence of a system’s behaviour from the initial conditions is extremely sensitive and appears to be extremely erratic; the only predictions that can be made are probabilistic;
- *presence of feedback cycles (positive or negative)*. The relationships among the elements become more important than their own specific characteristics, and the feedback cycles can influence the overall behaviour of the system;
- *distributed nature.* Many properties and functions cannot be precisely localised, in many cases there are redundancies and overlaps; it is a *distributed* system;
- *emergence and self-organisation*. A number of *emergent* properties are not directly accessible (identifiable or foreseeable) from an understanding of its components. Very often, in a CAS, global structures emerge over a critical threshold of some parameter. Typically, a new hierarchical level appears that reduces the complexity. In continuing the evolution, the system evolves, increasing its complexity up to the next self-organisation process;

- *limited decomposability*. The dynamic structure is studied as a whole. It is difficult, if not impossible, to study its properties by decomposing it into functionally stable parts. Its permanent interaction with the environment and its properties of self-organisation allow it to functionally restructure itself;
- *self-similarity*. It implies that the system considered will look like itself on a different scale, if magnified or made smaller in a suitable way. The self-similarity is evidence of a possible internal complex dynamic. The system is at a critical state between chaos and order, a condition that has been also called a self-organised critical state. A self-similar object, described by parameters N and z, has a power-law relationship between them: $N = z^k$. The best known of these laws is the rank-size rule which describes objects as varied as population in cities, word frequencies, and incomes. A power-law means that there is no 'normal' or 'typical' event, and that there is no qualitative difference between the larger and smaller fluctuations.

Examples of CAS include the patterns of birds in flight or the interactions of various life forms in an ecosystem, the behaviour of consumers in a retail environment, people and groups in a community, the economy, the stock market, the weather, earthquakes, traffic jams, the immune system, river networks, zebra stripes, sea-shell patterns, and many others.

Complexity is a multidisciplinary concept derived from mathematics and physics that has been applied to the world of economics. As Saari (1995: 222) writes, "even the simple models from introductory economics can exhibit dynamical behavior far more complex than anything found in classical physics or biology."

To describe this approach, we may identify a number of features of an economy that present difficulties for the 'linear' mathematics usually employed in economics (derived from J. Holland, the father of genetic algorithms, and quoted in Arthur et al. (1997: 4):

- *Dispersed Interaction* What happens in the economy is determined by the interaction of many dispersed, possibly heterogeneous, agents acting in parallel. The action of any given agent depends upon the anticipated actions of a limited number of other agents and on the aggregate state these agents co-create.

- *No Global Controller* No global entity controls interactions. Instead, controls are provided by mechanisms of competition and coordination between agents. Economic actions are mediated by legal institutions, assigned roles, and shifting associations. Nor is there a universal competitor—a single agent that can exploit all opportunities in the economy.
- *Cross-cutting Hierarchical Organization* The economy has many levels of organization and interaction. Units at any given level—behaviors, actions, strategies, products—typically serve as 'building blocks' for constructing units at the next higher level. The overall organization is more than hierarchical, with many sorts of tangling interactions (associations, channels of communication) across levels.
- *Continual Adaptation Behaviors*, actions, strategies, and products are revised continually as the individual agents accumulate experience—the system constantly adapts.
- *Perpetual Novelty* Niches are continually created by new markets, new technologies, new behaviors, new institutions. The very act of filling a niche may provide new niches. The result is ongoing, perpetual novelty.
- *Out-of-Equilibrium Dynamics* Because new niches, new potentials, new possibilities, are continually created, the economy operates far from any optimum or global equilibrium. Improvements are always possible and indeed occur regularly.

A tourism system, involving such economic activity, shares many of these characteristics. The theoretical work in this field is still in its infancy and just a handful of researchers have started to consider the complex systems approach as a more effective framework for understanding the many and different phenomena (Farrell & Twining-Ward, 2004; Faulkner, 2000; McKercher, 1999). Complexity theory offers the hope of being able to understand, for example, how crises, disasters or turbulent changes may influence the sector, or why, after major crises such as 9/11, the tourism sector is able to show a rapid and almost unexpected recovery (UNWTO, 2002b).

Chaos theory is considered by many researchers as part of the wider 'complexity sciences' field (Lewin, 1999). Basically, a system may be considered as evolving from a completely ordered phase to one in which its behaviour is so strongly dependent on very small variations of the initial conditions to appear to be almost unpredictable: a ‘chaotic’ phase. In this, still governed by deterministic laws, the system may tend to certain specific configurations. These, the *attractors* and the regions close to them (their basins), can be fixed equilibrium points, or follow orbits or more complicated patterns. It is also possible to have a system that never returns to the same place (in these cases we speak of strange attractors).

The region at the boundary of these phases, known as the ‘edge of chaos’, is a region of complexity (Crutchfield & Young, 1990; Waldrop, 1992). Chaos theory essentially studies non-linear effects in deterministic systems, while complexity theory studies definite patterns in non-deterministic systems (Gleick, 1987; Kauffman, 1995).

In tourism, one model that explicitly adopts a chaos approach is the one proposed by McKercher. The main components of this system are (McKercher, 1999: 429):

- *The traveller*, who is the essential player in tourism, for without people travelling no tourism would occur.
- *The communication vectors* used to connect the traveller to the destination.
- *The considerations* or factors that influence the effectiveness of the *communication vectors used.*
- *The destination* or internal tourism community consisting of all businesses involved in tourism at the destination.
- *External tourism agencies* (public and private sector) that try to influence tourism.
- *Other tourism-related externalities*, such as alternative tourism destinations that affect a destination’s ability to attract travellers.
- *Non-tourism-related externalities*, or macro-environmental forces, such as changing political, economic or social conditions, war, natural disaster, that affect people’s ability to travel.
- *Outputs* from the system both desired and undesired.
- *Rogues or chaos makers* who can push a system to the edge of chaos.

The model describes the operation of complex tourism systems by representing the elements that influence tourism on a wide range of possible scales: national, regional, local and, possibly, at an enterprise level. It supposes that even if the number of actors influencing the system changes at each level, the relationships between the different elements are similar.

This way, the author provides a framework able to provide an analogy for the failure of many well designed, controlled and sustainable tourism development plans. By using this approach we may be also able to better understand the impact of the usage, and the uses, of technologies like the Internet on parts of the system and how the system may react to more developments (although only at a conceptual level).

The general picture can be summarised in the following way. Let us consider a destination as a representative tourism subsystem. It is a system whose elements can be thought of as belonging to a number of groups (such as those proposed by McKercher, 1999). These groups are connected with relationships that exhibit a non-linear dynamic behaviour, producing outcomes that cannot be simply explained as 'summing up' the individual characteristics. In the evolution of the system it is possible to find many of the features indicated by Arthur et al. (1997) as portraying a CAS.

A central property of a CAS is the possible emergence of unforeseen properties or structures termed self-organisation. This is one of the most striking features characterising a complex system. A consequence of this is the robustness or resilience of the system to perturbations (or errors); the system is relatively insensitive and has a strong capacity to return to a stable behaviour in the absence of external inputs. For tourism destinations and for the tourism sector in general, this property is the one which may be considered to have been exhibited on several occasions after crises and disruptions. Recent events such as the 9/11 and Madrid terrorist attacks, the Bali bombings, the SARS epidemics, the Iraq war and others, have greatly affected the sector. However, almost unpredictably, the recovery to pre-event levels (see, for example, the tourist arrivals statistics provided by the WTO) was accomplished in a relatively short period of time, typically a few months. No simple 'linear' model could have explained fully this behaviour.

2.2 Tools for the analysis of complex systems

A number of tools have been developed in recent years to cope with the task of describing a complex system. Many of them originate from the work of 19th century scientists, but only modern computational facilities have made them amenable to calculation. In their review, Amaral and Ottino (2004) identify three main classes of analysis tools each belonging to an area well known to physicists and mathematicians:

- nonlinear dynamics;
- statistical physics; and
- network theory.

2.2.1 Nonlinear dynamics

A striking characteristic of complex systems is the nonlinearity of the interactions among the components. The main consequence is that the equations describing a system's behaviour (provided they exist) can be solved only in very rare cases.

The work of Poincaré (1883; 1884) on the three body problem, at the end of 19th century, showed that even 'simple' Newtonian systems involving more than two bodies may exhibit very complicated dynamics with almost unpredictable results arising from small variations of the initial conditions.

Since then, a number of mathematical techniques have been developed to approximate the solutions of the differential equations used to describe such systems, but only the advent of modern powerful computers made it possible to find solutions (which, in nearly all cases, can be obtained by numerical approximations). Much of the mathematics of chaos theory involves the repeated iteration of simple mathematical formulas, which would be impractical to do otherwise. Nonlinear dynamic systems are capable of exhibiting self-organisation and *chaos*. This mechanism is called *deterministic chaos*, since the equations of motion which generate such erratic, and apparently unpredictable, behaviour do not contain any random terms.

Deterministic chaos refers to the irregular (chaotic) motion generated by a system whose evolution is governed by dynamic laws that uniquely determine the state of the system at all times from a knowledge of the system's previous history. The source of irregularity is the exponential divergence of initially close trajectories in a bounded region of phase-space. This divergence can be measured with the aid of the theory proposed by the 19th century Russian astronomer Aleksandr Mikhailovich Lyapunov (see Kantz & Schreiber, 1997). In this sense, chaotic behaviour can be regarded as very complex dynamics.

This sensitivity to initial conditions is sometimes popularly called the *butterfly effect*, suggesting the idea that chaotic weather patterns can be altered by a butterfly flapping its wings. A practical implication is that it is essentially impossible to formulate long-term predictions about the behaviour of a dynamic system; even if it were possible to fix the initial conditions to a predetermined, finite accuracy, their errors would increase at an exponential rate.

Examples of systems exhibiting nonlinear (chaotic) behaviour include the atmosphere, the solar system, plate tectonics, turbulent fluids, mixing of coloured dyes, economies, stock markets, population growth or the 'simple' double pendulum.

2.2.2 Statistical physics

Statistical physics (or statistical mechanics), one of the fundamental fields of physics, uses statistical methods for addressing physical problems. A wide variety of issues, with an inherently stochastic nature, are treated with these methods. It provides a framework for relating the microscopic properties of individual atoms and molecules to the macroscopic properties of materials observed in every day life. It is possible to explain thermodynamics, and thermodynamic properties, as a natural result of statistics and mechanics (classical and quantum).

The main result, and power, of this approach is in the bypassing of some classical mechanics problems, such as the impossibility of solving the three-body problem, by dealing with systems composed of a large number of elements, and reasoning in terms of statistical ensembles. Moreover, it introduced the idea of discrete models and agent-based models (Wolfram, 2002).

In recent years, our understanding of phase transitions and critical phenomena has led to the development of two important new concepts: universality and scaling (Amaral &

Ottino, 2004). Many physical systems exhibit universal properties that are independent of the specific form of the interactions among their constituents. This, for the sake of analogy, may suggest the hypothesis that universal laws or results may also show up in other types of complex systems, whether they be social, economic or biological.

The scaling hypothesis, born in the framework of the study of critical phenomena, has provided the idea that a set of relations, called scaling laws, may help in relating the various critical-point exponents characterising the singular behaviour of an order parameter and of response functions. The predictions of the scaling hypothesis are supported by a wide range of experimental work, and also by numerous calculations on model systems.

The concept of universality in statistical physics and complex systems has the basic objective of capturing the essence of different systems and classifying them into distinct classes. The universality of critical behaviour pushes the investigations on the features of the microscopic relationships important for determining critical-point exponents and scaling functions.

Statistical approaches can be very effective in systems when the number of degrees of freedom (and elements described by a number of variables) is so large that an exact solution is not practical or possible. Even in cases where it is possible to use analytical approximations, most current research utilises the processing power of modern computers to simulate numerical solutions.

One more important outcome of the use of statistical physics methods is the use of discrete models. The fundamental assumption is that some phenomena can be modelled in terms of computer programs (algorithms) rather than in terms of analytical expressions. Cellular automata (see for example Wolfram, 2002) are examples of discrete time and space models developed for a computer utilisation.

Cellular automata are computer programs that model dynamic structures, discrete in space and time, that operate on a uniform, regular lattice, characterised by ‘local’ interactions. They are made up of many cells, each of which may be in one of a finite number of states. A cell may change state only at fixed, regular intervals, and only in accordance with fixed rules that depend on their own values and the values of neighbours within a certain distance.

Applications of such models exist in many fields of physical, chemical, biological and social sciences; the propagation of fire, predator-prey models or the evolution of artificial organisations can all be represented with cellular automata.

2.2.3 Network theory

Most complex systems can be described as networks of interacting elements. In many cases these interactions lead to global behaviours that are not observable at the level of the single elements and that share the characteristics of emergence typical of a complex system. Moreover, the collective properties of dynamic systems composed of a large number of interconnected parts are strongly influenced by the topology of the connecting network. This concept will be developed further in the next sections.

2.3 Elementary graph theory

The mathematical models of network structures have been developed in graph theory. A graph is a generalisation of the concept of a set of points (vertices, nodes), connected by links (edges, arcs). These, depending on the specific situation, may or may not have a direction (the graph is directed or undirected). In a directed graph it is possible to track a route from some vertex to another, but not in the opposite direction. Links may be associated with numeric values (they may represent distances, costs, energies, information exchanges etc.) called weights. In the last few years, a number of researchers have shed light on some topological aspects of many kinds of social and natural networks, (the WWW, power grids, collaboration networks, networks of words, metabolic networks, economic agents).

2.3.1 A brief historical sketch of graph and network theories

Graph theory is one of the few areas in mathematics with a definite date of birth.[2] The paper that fixes the start is *Solutio problematis ad geometriam situs pertinentis* written in

[2] Part of this section forms the introduction to the book: Network Analysis and Tourism: From Theory to Practice, by N. Scott, C. Cooper and R. Baggio, Channelview (2008b)

1736 by the Swiss mathematician Leonhard Euler. In it, Euler proposes a mathematical formulation of the renowned Königsberg Bridge Problem:

> Is it possible to plan a walk through the town of Königsberg which crosses each of the town's seven bridges once and only once?

The puzzle problem, proposed to Euler by the Königsberg inhabitants, is stated in this way (the English translation of the original Latin paper can be found in Biggs et al., 1976):

> The problem, which I am told is widely known, is as follows: in Königsberg in Prussia, there is an island A, called the Kneiphof; the river which surrounds it is divided into two branches, as can be seen in Figure 2.1, and these branches are crossed by seven bridges, a, b, c, d, e, f and g. Concerning these bridges, it was asked whether anyone could arrange a route in such a way that he would cross each bridge once and only once. I was told that some people asserted that this was impossible, while others were in doubt: but nobody would actually assert that it could be done. From this, I have formulated the general problem: whatever be the arrangement and division of the river into branches, and however many bridges there be, can one find out whether or not it is possible to cross each bridge exactly once?

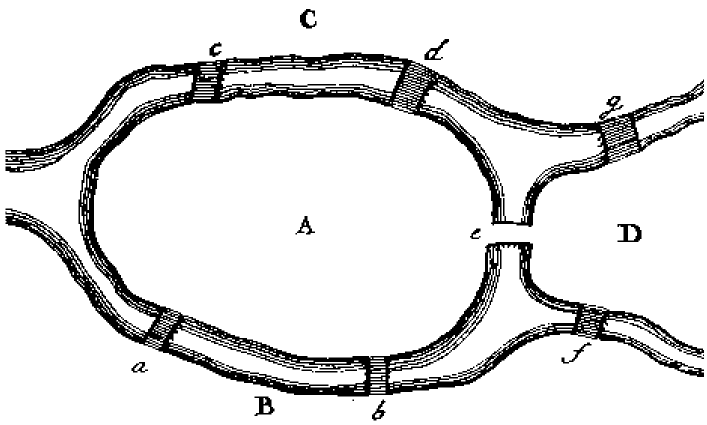


Figure 2.1 The Königsberg bridges (source: Euler, 1736)

The solution of the general problem is well known. In Euler's words:

> So whatever arrangement may be proposed, one can easily determine whether or not a journey can be made, crossing each bridge once, by the following rules: If there are more than two areas to which an odd number of bridges lead, then such a journey is impossible. If, however, the number of bridges is odd for exactly two areas, then the journey is possible if it starts in either of these areas. If, finally, there are no areas to which an odd number of bridges leads, then the required journey can be accomplished starting from any area. With these rules, the given

problem can always be solved. When it has been determined that such a journey can be made, one still has to find how it should be arranged. For this I use the following rule: let those pairs of bridges which lead from one area to another be mentally removed, thereby considerably reducing the number of bridges; it is then an easy task to construct the required route across the remaining bridges, and the bridges which have been removed will not significantly alter the route found, as will become clear after a little thought. I do not therefore think it worthwhile to give any further details concerning the finding of the routes.

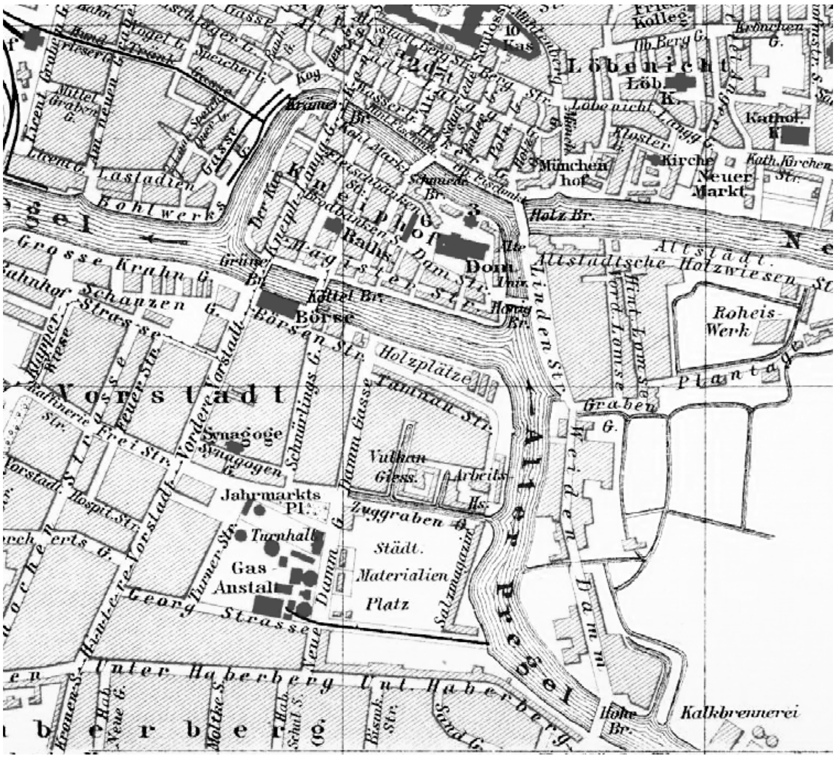


Figure 2.2 Königsberg (from a 1894 map online at http://www.koenigsberg-stadtplan.de.vu/)

The importance of Euler's paper for the history of mathematics does not lie, obviously, in the solution of the game. It is related to the approach taken, the one stated in the very first paragraph of the paper:

In addition to that branch of geometry which is concerned with magnitudes, and which has always received the greatest attention, there is another branch, previously almost unknown, which Leibniz first mentioned, calling it the geometry of position. This branch is concerned only with the determination of position and its properties; it does not involve measurements, nor calculations made with them. It has not yet been satisfactorily determined what kind of problems are relevant to this geometry of position, or what methods should be used in solving them. Hence, when a problem was recently mentioned, which seemed geometrical but was so

> constructed that it did not require the measurement of distances, nor did calculation help at all, I had no doubt that it was concerned with the geometry of position - especially as its solution involved only position, and no calculation was of any use. I have therefore decided to give here the method which I have found for solving this kind of problem, as an example of the geometry of position.

G*eometria situs*, as Leibniz had called it[3] is today known with the name of topology, and Euler's solution is the first of this kind formally stated and solved.

Despite the numerous but sparse works on this topic in the second part of the 18th and in the 19th centuries (Cauchy, Kirchoff, Hamilton, Poincaré, to quote just the most famous authors), a formal setting of these theories came only exactly 200 years after the Königsberg bridges paper. In 1936, the German mathematician Dénes König (1884-1944) published in Leipzig the first systematic study of what he called *graphs* in his *Theorie der endlichen und unendlichen Graphen* (König, 1936).

In the same period, ideas and techniques developed for the study of these abstract objects were applied to a completely different field. Realising that a group of individuals can be represented by enumerating the *actors* of the group and their mutual relationships, sociologists started using graph theory and methods to describe and analyse the patterns of social relations (Freeman, 2004; Wasserman & Faust, 1994). Jacob L. Moreno (1934) introduced *sociometry*. By using a *sociogram* (a diagram of points and lines used to represent relations among persons) he aimed at identifying the structure of relationships around a person, group, or organisation in order to study how these configurations may affect beliefs or behaviours.

Today, the term *social network analysis* has superseded the earlier *sociometry*, but both refer to the analysis of social networks in part utilising graphical methods. Friendships among individuals, business relationships between companies, and trade agreements

---

[3] In a letter to C. Huygens (dated 8 September 1679), Gottfried W. Leibniz (1646-1716) writes: *I am not content with algebra, in that it yields neither the shortest proofs nor the most beautiful constructions of geometry. Consequently, in view of this, I consider that we need yet another kind of analysis, geometric or linear, which deals directly with position, as algebra deals with magnitude.* (quoted by Biggs et al., 1976: 20; see also Leibniz, 1693).

among nations are all examples of networks which have been studied by using these techniques.

Different lines of research stem from this origin. As Scott et al. (2000: 7) states:

> A number of very diverse strands have shaped the development of present-day social network analysis. These strands have intersected with one another in a complex and fascinating history, sometimes fusing and other times diverging on to their separate paths. A clear lineage for the mainstream of social network analysis can, nevertheless, be constructed from this complex history. In this lineage there are three main traditions: the sociometric analysts, who worked on small groups and produced many technical advances with the methods of graph theory; the Harvard researchers of the 1930s, who explored patterns of interpersonal relations and the formation of 'cliques'; and the Manchester anthropologists, who built on both of these strands to investigate the structure of 'community' relations in tribal and village societies. These traditions were eventually brought together in the 1960s and 1970s, again at Harvard, when contemporary social network analysis was forged.

Common terms such as *weak ties* in a social context, or the *smallness* of the world of our acquaintances, come directly from seminal works in this area such as that by Mark Granovetter (1973) or the Stanley Milgram (1967) experiments.

The next major breakthrough in the history of network theories is a series of three papers published in the early 1960s by the Hungarian mathematicians Paul Erdős and Alfréd Rényi (1959; 1960; 1961) on *random graphs*. The problem addressed was a fundamental question in the quest for understanding graphs, networks and interconnection phenomena: how do these objects form? and how do they evolve over time?

The approach used is statistical and probabilistic. The Erdős-Rényi (ER) model has, since then, become a standard model, able to explain many of the characteristics of the networks encountered in the real world (see section 2.4.1). For almost 30 years the ER model, the only available model of this kind, has been used, investigated and developed by many authors.

In the last years of the 1990s, the Internet revolution had a tremendous impact on almost all aspects of our life. One of the crucial influences the Internet has had is in the birth of a completely new approach to the studies of networks. For the first time a huge mass of data was available to researchers. Data were easily accessible and usable for sophisticated statistical analyses, and the Internet was the primary source, or the medium, for this availability.

The beginning of the *'New' Science of Networks,* as it has been named (Watts, 2004), lies in three papers written in 1998-1999:

- *Collective dynamics of 'small world' networks*, by Watts and Strogatz (1998);
- *On power-law relationships of the internet topology* by M. Faloutsos, P. Faloutsos, and C. Faloutsos (1999); and
- *Emergence of scaling in random networks* by Barabási, and Albert (1999).

These works have provided evidence that the ER model was simply a crude approximation of only a special class of networks, and that many of those found in the real world, technological, physical, biological or social, exhibited characteristics and properties of a different nature. Since then, a vast amount of work has been carried out and numerous phenomena have received explanation and have been modelled. Furthermore, it has strongly reinforced the idea that the collective properties of dynamic systems composed of a large number of interconnected parts are strongly influenced by the topology of the underlying network (see the bulky reviews by Albert & Barabási, 2002; Boccaletti et al., 2006; Dorogovtsev & Mendes, 2002; Newman, 2003b).

One more aspect of this work is also worth noting: the contributions to this *new science* are, probably for the first time in the history of science, truly and absolutely interdisciplinary. Physicists, mathematicians, computer scientists, biologists, economists, and sociologists are all equally contributing to the growth of the knowledge in this field.

2.3.2 Graphs: main definitions

A network of elements (such as people, computers, firms, or roads) is usually represented by a diagram consisting of a number of dots (nodes or vertices) and a number of lines (arcs or edges) connecting certain pairs of dots. From a mathematical point of view, in such diagrams the important thing is whether or not two given points are connected by a line, and the nature of the connection is disregarded. This abstraction is what we call a *graph*.

Formally[4] a *graph* G is an ordered pair of disjointed sets (V, E), where $V = \{v_1, \dots v_n\}$ is the set of vertices and $E = \{(v_1,u_1), \dots, (v_i,u_j)\}$ is the set of arcs. E is a subset of the Cartesian product $V \times V$. In other words, E is a binary relation on V and its elements are pairs of elements belonging to V.

If E is symmetrical, i.e.:

$$\forall v_i, v_j : e_{ij} = (v_i, v_j) \in E \Leftrightarrow e_{ji} = (v_j, v_i) \in E$$

the graph is said to be *undirected* (Figure 2.3).

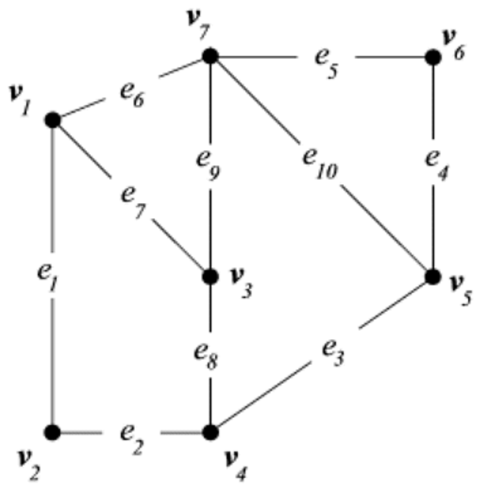
If E is not symmetrical the graph is said to be *directed* (the term *digraph* is sometimes used to denote a directed graph).

In a directed graph, edges are ordered pairs, connecting a source vertex to a target vertex. In an undirected graph, edges are unordered pairs and connect the two vertices in both directions, hence in an undirected graph (u,v) and (v,u) are two ways of denoting the same edge.

The cardinality $|V|$ of the vertex set (the number of vertices: n) is called *order* of the graph G; the cardinality $|E|$ of the edges set (the number of edges: m) is called *size* of the graph G (it must be noted here, however, that, quite often, the term size is used in the meaning of number of vertices of a graph).

The different elements of a graph may assume different names: vertices are also called points, nodes, or actors, and the links among them are called edges, arcs, connections, or lines (usually, the term arc is used for a directed link between two nodes, edge for an undirected one).

[4] This and the following sections are derived from standard graph theory texts (Bollobás, 1998; Bondy & Murty, 1982; Diestel, 2005; Godsyl & Royle, 2001).



Graph $G = (V, E)$

vertices:

$$V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$$

edges:

$$E = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8, e_9, e_{10}\} = \\ = \begin{Bmatrix} (v_1,v_2),(v_2,v_4),(v_4,v_5),(v_5,v_6),(v_6,v_7), \\ (v_7,v_1),(v_1,v_3),(v_3,v_4),(v_3,v_7),(v_7,v_5) \end{Bmatrix}$$

Figure 2.3 An undirected graph.

A graph is called *weighted* if each edge has been assigned a numeric value *w* (weight); the weight can represent a distance, a time, a cost etc. (Figure 2.4).

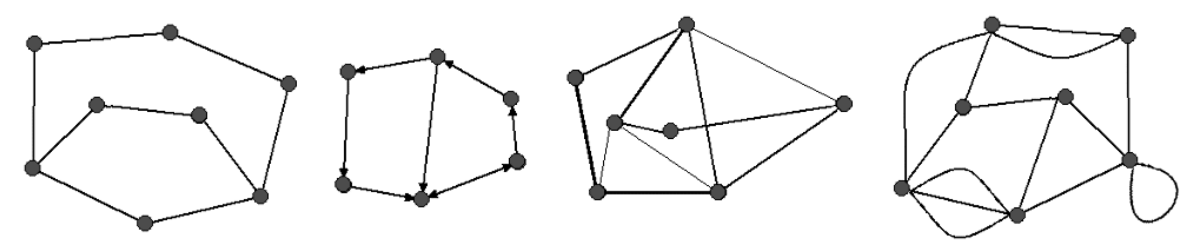


Figure 2.4 Graph types: a, b and c are simple graphs, (a) is undirected and unweighted, (b) is directed, (c) is weighted; (d) is a multigraph containing a loop.

Two edges of a graph are called *adjacent* if they share a common vertex. Two vertices are called *adjacent* if they share a common edge, i.e.: $(u,v) \in E$ (if the graph is directed, the converse may not be true). A vertex *v* adjacent to *u* is called *neighbour*, or *first neighbour* of *u*; the set of vertices adjacent to a certain vertex *v*: $\Gamma(v) = \{v \in V: v \text{ adjacent to } u\}$ is called *neighbourhood* of *u*.

A complete graph K_n of order *n* is a simple graph with *n* vertices in which every vertex is connected (adjacent) to every other. It has $n(n-1)/2$ edges (all possible choices of pairs of vertices).

A subgraph of a graph *G* is a graph whose vertex and edge sets are subsets of those of *G*. If all the vertices are connected the subgraph is called *clique*. Since any subgraph induced by a clique is a complete subgraph, the two terms and their notations are usually used interchangeably. A *k-clique* is a clique of order *k* (a *3-clique* is called a *triangle*). The clique number *ω(G)* of a graph *G* is the order of the largest clique in G.

If the set of vertices *V* of a graph *G* can be divided into two disjointed non-null sets *A* e *B*, such that each edge of *G* connects a vertex in *A* with a vertex in *B*, the graph *G* is called *bipartite* (there are no 'internal' edges among the vertices in *A* or *B*).

The edge's *density* is the ratio between the actual number of links and the maximum possible number:

$$\delta = \frac{m}{\frac{1}{2}n(n-1)}$$

The number of edges connected to a vertex *u* is called vertex *degree deg*(*u*) (also indicated as d_u). In a directed graph the degree is usually divided into the *in-degree* and the *out-degree* (whose sum is the degree of the vertex in the underlying undirected graph). The in-degree of a vertex *u* is the number of edges with *u* as their terminal vertex. The out-degree of a vertex *u* is the number of edges with *u* as their initial vertex. A vertex may be said to be *even* or *odd* if its degree is such.

Average (mean) and maximum degrees are the arithmetic mean and the maximum values for a vertex degree in a graph. It can be easily demonstrated that the average degree is: $k = 2m/n$, where m = *number of edges* and n = *number of vertices*. The maximum possible degree in a simple graph is ½ *n(n-1)* (each vertex is connected to all other vertices). A *k-regular* graph is a graph in which all the *n* vertices have the same degree *k*.

In a graph *G*, the *degree distribution P(k)* is a function which gives the probability that any vertex $u \in V$ has degree *k* (has *k* neighbours): *Pr[deg(u) = k]*. The sum of the degrees of all its vertices is called *volume*:

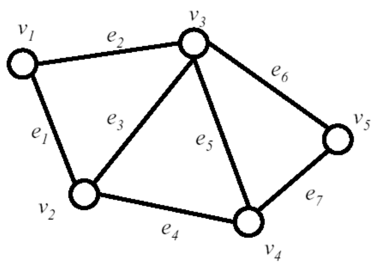
$$vol(G) = \sum_{v_i \in V} deg(v_i)$$

An interesting parameter describing the connections in a neighbourhood of a certain vertex is the *clustering coefficient.* In an undirected graph, the number of all the possible connections between a vertex *i* and all its neighbours is E_i. The ratio between this value and the maximum possible number of edges in the graph $[k_i(k_i\text{-}1)]/2$ is called the *clustering coefficient*:

$$C_i = \frac{2E_i}{k_i(k_i - 1)}$$

C_i varies in the interval (0, 1) and represents a local density of edges. The average (arithmetic mean) of the clustering coefficients of all the vertices in a graph is the *clustering coefficient* of the graph:

$$C = \frac{1}{N}\sum_{i=1}^{N} C_i$$



$$A_G = \begin{array}{c|ccccc} & v_1 & v_2 & v_3 & v_4 & v_5 \\ \hline v_1 & 0 & 1 & 1 & 0 & 0 \\ v_2 & 1 & 0 & 1 & 1 & 0 \\ v_3 & 1 & 1 & 0 & 1 & 1 \\ v_4 & 0 & 1 & 1 & 0 & 1 \\ v_5 & 0 & 0 & 1 & 1 & 0 \end{array}$$

Figure 2.5 An undirected graph and its adjacency matrix

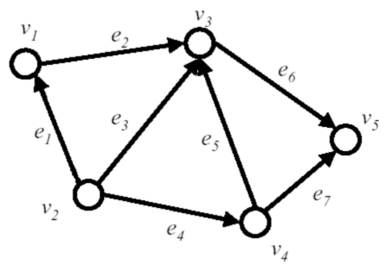
A standard way to characterise a graph is the *adjacency matrix* (Figure 2.5). An undirected graph $G = (V, E)$ with *n* vertices can be represented by an $n \times n$ symmetric Boolean matrix A_G whose elements are defined by:

$$A_G[i, j] = \begin{cases} 1 & if\ \{i, j\} \in E \\ 0 & otherwise \end{cases}$$

If the graph is directed, A_G is defined by:

$$A_G[i,j] = \begin{cases} 1 & \textit{if } \langle i,j \rangle \in E \\ 0 & \textit{otherwise} \end{cases}$$

In this case the matrix is not symmetric ($\langle i, j\rangle$ represents an ordered pair, Figure 2.6).



$$A_G = \begin{array}{c|ccccc} & v_1 & v_2 & v_3 & v_4 & v_5 \\ \hline v_1 & 0 & 0 & 1 & 0 & 0 \\ v_2 & 1 & 0 & 1 & 1 & 0 \\ v_3 & 0 & 0 & 0 & 0 & 1 \\ v_4 & 0 & 0 & 1 & 0 & 1 \\ v_5 & 0 & 0 & 0 & 0 & 0 \end{array}$$

Figure 2.6 A directed graph and its adjacency matrix

The adjacency matrix is symmetric if the graph is undirected. If the graph is weighted, the arrays contain integer or real numbers (instead of Boolean values), where the number indicates the weight of an edge.

The adjacency matrix is unique for each graph and each graph has a unique adjacency matrix. This biunivocal relationship allows us to swap between the geometric and the algebraic (matricial) representation of a graph and to use the results coming from both areas interchangeably. Many interesting results derived from linear algebra techniques can be used. Some of them will be described in section 2.3.4.

In a graph G, a *path* from vertex u to vertex u' with length k is a sequence $(v_0, v_1, \ldots, v_k)$ of vertices such that $u = v_0$, $u' = v_k$ and $(v_1, v_i) \in E$, for $i=1, \ldots, k$. A path is the series of edges that connect $(v_0, v_1, \ldots, v_k)$. If a path from u to v exists, v is said to be *reachable* from u. A path is *simple* if it contains no vertex more than once, if some vertices are repeated, it is called a *walk*. A *cycle* is a path from some vertex to that same vertex. A cycle is *simple* if it contains no vertex more than once, except the start (and end) vertex, which only appear as the first and last vertex in the path. A cycle of length 3 is called a *triangle*, a cycle of length 1 is called a *loop*. If G has no cycles it is said to be *acyclic*. A *tree* is a connected acyclic.

In a weighted graph, the length *d(u,v)* of a path *p* is the sum of the weight *w* on the relevant edges:

$$d(p) = \sum w(v_{i-1}, v_i)$$

If more than one path exists between two vertices *u* and *v*, the *shortest path* is the one for which the total weight $\delta(u, v) = w(p)$ is the minimum possible. If there is no connection between two vertices, it is customarily assumed to be: $\delta(u, v) = \infty$. If the graph is unweighted, obviously, the length of a path is calculated by putting $w = 1$ for all the edges of the path. A shortest path (also called geodetic) is called, and interpreted as, the distance between two vertices. Paths and cycles and their characteristics are important structural features of a graph.

The maximum value of the *distance d(u,v)* from a vertex *u* to each other $v \in V$ is the *eccentricity* of *u.* The maximum eccentricity, i.e. the maximum *d(u,v)* for each pair of graph vertices, is the graph *diameter diam(G)* (or D_G). The minimum graph eccentricity is called the graph *radius r(G)*.

A first example of practical usage of the adjacency matrix A_G of a graph *G* is in the calculation of the paths of a graph. It can be easily demonstrated (by induction) that the number of paths of length *n* from v_i to v_j is the (*i, j*) entry of A_G^n. By raising A_G to a power *m* so that all the elements a_{ij} of A_G are positive, *m* is the diameter of *G*.

One more important parameter which describes a graph is its *characteristic (*or *average) path length L(G).* It is the arithmetic mean of the distance values calculated by considering all possible pairs of vertices. Since the number of pairs is *½(n(n-1)*, *L(G)* is defined by:

$$L(G) = \frac{2}{n(n-1)} \sum_{\{u,\, v\} \subseteq V} d(u, v)$$

A *connected* graph is an undirected graph that has paths between any two vertices. *Connected components* of a graph are the equivalence classes of vertices defined by the relation *reachable from*. A directed graph is *strongly connected* if it has a path from each vertex to every other vertex.

The maximum possible length for a path between any two vertices in a connected graph is *n-1* (*n* order of the graph); the number of cycles is: $I = m - n + 1$. If a graph contains no closed cycles (it is a *tree*) and is connected, then: $m = n - 1$.

2.3.3 Graphs of social networks

In the study of *social* networks, graph analysis techniques are the basis on which models are built. The discipline has developed its own language (typically a network node is called an *actor* or *agent* and the edges are denoted as *ties* or *relationships*), and a number of characteristics, relevant for the type of considerations existing in investigating groups of individuals or organisations, have developed with a peculiar terminology (Hannemann, 2001; Scott, 2000; Wasserman & Faust, 1994).

In particular, focus is given to the concept of *centrality* of a network node (actor) with respect to the others: the individual (vertex in a graph) most popular in his group, the one at the centre of the stage.

Two main types of centrality measures exist: local and global. The first one quantifies, basically, how well a node is connected with its neighbourhood. The simplest measure of local centrality is the node degree:

$$C_D(v_i) = deg(v_i)$$

Global centrality takes into account the structure of the whole network in determining the importance of the position of a certain actor. The common measures used are closeness and betweenness.

Closeness is expressed in terms of the distance (geodetic) between the different nodes. A vertex is globally central if its distance from many other vertices is small. In other words, the shorter the distances between node *i* and other nodes, the more central node *i* is. Formally:

$$C_C(v_i) = \frac{1}{\sum_{j=1}^{n} d(v_i, v_j)}$$

where ($d(v_i,v_j)$ is the distance between v_i and v_j).

Betweenness assesses the probability that a path from actor j to actor k takes a particular route through an *intermediary i*. It can be interpreted as a local dependency of an actor from the others. Betweenness is defined as:

$$C_B(v_i) = \sum_{j<k} g_{jk}(v_i) / g_{jk}$$

where g_{jk} is the probability that, for each pair j and k, vertex i belongs to the geodetic jk. In other terms, $g_{jk}(v_i)$ is the number of geodetics between j and k that go through i, while g_{jk} is the total number of geodetics between j and k.. It is thus possible to measure the centrality of i with respect to its possible role as an intermediary between j and k.

These measures depend on the size of the network. Therefore, to allow comparisons, they are standardised: degree centrality and closeness are divided by $n-1$; betweenness is divided by $(n-1)(n-2)$.

2.3.4 Graph spectra

Let A be a square matrix and x a non-null vector. It is possible to find a scalar λ such that (Lang, 1970):

$$Ax = \lambda x$$

λ is called *eigenvalue* for A and x is the corresponding *eigenvector*; together they form the *eigenpair*. The eigenvalue satisfies the equation:

$$(A - \lambda I)x = 0$$

which has nontrivial solutions if and only if:

$$det(A - \lambda I) = 0.$$

This equation[5] is known as the characteristic equation of A, and the left-hand side is known as the characteristic polynomial.:

$$p(\lambda) = det(A - \lambda I)$$

There exist exactly n roots (not necessarily distinct) of a polynomial of degree n. If A is an $n \times n$ real matrix, then its n eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the real and complex roots of the characteristic polynomial. Two similar matrices have the same characteristic polynomial and therefore the same eigenvalues (the same spectrum).

The ordered set of the eigenvalues for A is called the *spectrum* of A.

$$sp(A) = \lambda_1, \lambda_2, \ldots, \lambda_n \text{ with } \lambda_1 \geq \lambda_2 \geq, \ldots, \geq \lambda_n$$

Since a graph is completely described by its adjacency matrix, expressions such as *eigenvalue* (or spectrum) *of the adjacency matrix of the graph* and *eigenvalue* (or spectrum) *of the graph* are equivalent. The number of eigenvalues of a graph corresponds to its order (number of vertices). For an undirected graph the adjacency matrix is real symmetric, hence its eigenvalues are all real numbers.

Eigenvalues and spectrum of a graph are strictly connected to its structural characteristics and do not depend on the order of the elements of the matrix; they represent the topology of the graph (Seary & Richards, 2003).

The *spectral density* $\rho(\lambda)$ of a graph is the spectral density of its eigenvalues. For a finite system it is defined as the sum of δ functions:

$$\rho(\lambda) := \frac{1}{n}\sum_{j=1}^{n} \delta(\lambda - \lambda_j)$$

where δ is:

[5] The determinant of a square matrix, denoted with det(A), is a function that associates with each matrix A a single scalar number. It is calculated as the sum over the product of all permutations of all elements within any row, multiplied by either +1 or -1, depending on whether the respective permutation is even or odd.

$$\delta(\lambda - \lambda_j) = \begin{cases} \infty & \text{if } \lambda = \lambda_j \\ 0 & \text{otherwise} \end{cases}$$

$\rho(\lambda)$ converges to a continuous function for $n \to \infty$.

In a complete graph the maximum eigenvalue (the *principal eigenvalue*) corresponds to the common degree of the graph, otherwise it is at least equal to the average graph degree (it is an indicator of the edge density):

$$\lambda_1 \leq \max_i \deg_G(i)$$
$$\lambda_1 \geq \frac{1}{n} \sum \deg_G(i) \quad \text{[mean degree]}$$

From these it is possible to derive:

$$\lambda_1 = \max_i \deg_G(i) \text{ and G is connected} \Rightarrow \text{G is regular}$$
$$\lambda_1 = \frac{1}{n} \sum \deg_G(i) \Rightarrow \text{G is regular}$$

If G is an unordered k-regular graph, $k = \lambda_1$, and the multiplicity of k as eigenvalue of G equals the number of connected components of G.

A connected graph with diameter D has at least $D+1$ distinct eigenvalues and $\lambda_1 \geq 1/(D\ volG)$; where $vol\ G = \sum deg(v)$ (vol is the volume of the graph).

If d_{max} is the maximal degree of a graph G ($|E| = m$ number of edges), we have:

$$|\lambda_i(G)| \leq \min(d_{\max}, \sqrt{|E|})$$

The second smallest eigenvalue is a measure of graph compactness; a large value indicates a compact network, a small one is an indicator of an 'elongated' topology.

The eigenvector I_{n-1} corresponding to the second smallest eigenvalue λ_{n-1} is used to solve the *minimum cut* problem: dividing a graph into two subgraphs of (almost) the same order with the minimum possible number of connections between them.

The Perron-Frobenius theorem (Godsyl & Royle, 2001) states that there exists a unique positive eigenvalue λ_A whose value is the largest ($\forall$ *eigenvalue* $\kappa : |\kappa| \leq \lambda_A$); its logarithm is used as a measure of the entropy of a graph:

$$h(G) = log\ (\lambda_A)$$

It is also possible to define an energy for a graph G as the sum of the absolute values of all of its eigenvalues:

$$\mathrm{E}(G) = \sum_i |\lambda_i|$$

If the degree distribution of a graph follows a power law with exponent α ($P(k) \propto k^{-\alpha}$), the highest eigenvalues are distributed according to a power law with exponent $\alpha/2$.

The degree matrix D is a diagonal matrix associated to the adjacency matrix A_G, whose elements are defined as:

$$D_G[i,j] = \begin{cases} \sum deg(i) & for\ i = j \\ 0 & otherwise \end{cases}$$

From D two more important matrices can be derived:

$$Laplacian\ matrix{:}\ L = D - A$$

$$Normal\ matrix{:}\ N = D^{-1} A$$

L defines a set of prime differences among the vertices of G and can be interpreted as being analogous to the continuous operator gradient ∇; its square LL^T as the continuous operator Laplacian ∇^2. L is a symmetric $n \times n$ matrix (n = order of the graph G):

$$L_{ij} = \begin{cases} deg(i) & if\ \ i = j \\ -w_{ij} & if\ i \neq j \end{cases}$$

(w_{ij} = weight of the edge ij, for an unweighted graph: $w=1$ for all the edges).

L has a number of interesting properties:

- *L* is a real and symmetric matrix, therefore all its eigenvalues are real;
- all the eigenvalues of *L* are non-negative;
- the eigenvalues of *L* are in the interval [0,2];
- the second smallest eigenvalue indicates the connectedness of the graph;
- the multiplicity of the null eigenvalue equals the number of the connected components in *G*.

The spectrum of the Laplacian matrix of a graph can be seen as a *normalised spectrum;* this makes easier the comparison of different spectra and of different topologies. For this reason the Laplacian spectrum is increasingly used for the study of the network properties.

Besides the capacity of revealing the structure of a network, the spectrum can be used to define local metrics able to take into account the whole structure of the network.

One example is the *eigenvector centrality* measure (Bonacich, 1972, 1987). This measure recognises the idea that not all connections are equal. Higher weight is given with a reciprocal process in which the centrality of each node is proportional to the sum of the centralities of those to whom it is connected. If x_i is the centrality of vertex *i* , which has *j* neighbours, the centrality is given by:

$$x_i = \frac{1}{\lambda} \sum_{j=1}^{n} A_{ij} x_j$$

where λ is the principal eigenvalue of the graph. The formula can be rewritten as:

$$\lambda x = Ax$$

where x is the eigenvector corresponding to λ (the principal eigenvector of the graph).

A variant of eigenvector centrality is employed by the Google search engine to rank web pages (the PageRank described by Brin & Page, 1982).

2.4 Random networks

This section contains the discussion of the main network models and their characteristics, and the most important dynamic processes that influence, or are influenced by, the different topologies of networks.

The standard references for the whole section (unless otherwise noted in specific parts) are the reviews by Boccaletti et al. (2006), Watts (2004), Newman (2003b), Albert and Barabási (2002), Dorogovtsev and Mendes (2002), and the books by Bornholdt and Schuster (2002), Dorogovtsev and Mendes (2003), and Pastor-Satorras and Vespignani (2004).

2.4.1 The Erdős-Rényi model

As discussed above, the Erdős-Rényi (ER) model is the first which has attempted to provide an explanation of the behaviour of a large real network. Up to the end of the last century, this model was the only means of representing all those situations in which a regular lattice could not reasonably explain a certain phenomenon.

An ER random network (graph, Figure 2.7) is formed by N nodes connected by L links taken at random from the $N(N-1)/2$ possible links (Erdös & Rényi, 1959). This way we build a statistical ensemble of graphs, all having the same probability. The number of graphs in the ensemble is:

$$K = C^{L}_{\frac{N(N-1)}{2}}$$

As an alternative, it is possible to use a binomial model. Starting with N nodes, two at a time are connected with probability p; the number of edges L is a stochastic variable with expected value:

$$E(L) = p\frac{N(N-1)}{2}$$

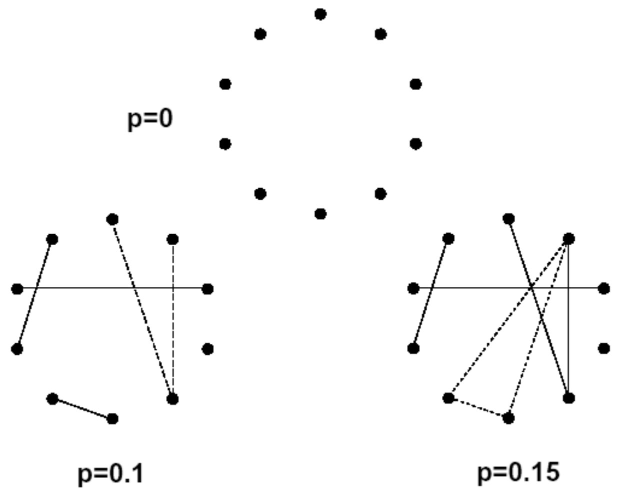


Figure 2.7 The formation of an ER random graph

The probability of obtaining a graph G_0 with N vertices and L edges is:

$$P(G_0) = p^L (1-p)^{\frac{N(N-1)}{2}-L}$$

G_0 is a connected graph without duplicate edges if $N - 1 \leq L \leq N(N - 1)/2$. A relation between L and N exists, even if the two parameters are considered to be independent of each other.

The construction of a graph is sometimes called a process of evolution. Starting from N isolated nodes, the graph develops with subsequent additions of random edges (with probability p). In different stages we obtain increasing connection probabilities up to a fully connected graph ($p \rightarrow 1$). ER random graphs are indicated as $G(N,p)$ or $G_{N,p}$

The model studies the features of graphs for $N \rightarrow \infty$ and the connection probability p which generates some property Q. The important result is that, given a probability p, Q exists in (almost) all graphs generated or it does not exist at all. That is, there is a critical threshold p_c beyond which Q appears; if $p(N)$ grows faster than $p_c(N)$, (almost) all the graphs have Q, otherwise no graph has Q.

In other words, the probability that a graph with N vertices and L edges connected with probability $p = p(N)$ has Q is:

$$\lim_{N\to\infty} P_{N,p}(Q) = \begin{cases} 0 \text{ if } \dfrac{p(N)}{p_c(N)} \to 0 \\ 1 \text{ if } \dfrac{p(N)}{p_c(N)} \to \infty \end{cases}$$

The critical probability $p_c(N)$ behaves like the similar parameter (percolation threshold) of the percolation theory[6] (Stauffer & Aharony, 1992). In a physical system, usually, N is fixed, so the different regimes are identified by values of p greater or smaller than p_c.

In random graph theory the connection probability is a function of the graph order: p is the fraction of edges actually present with respect to the total $N(N-1)/2$. This means that, for many properties Q, there is no single threshold p_c independent from N; rather, it is defined as a function of the network size and for the *percolation threshold* the relationship $p_c(N\to\infty)\to 0$ holds.

The mean number of edges per node (mean nodal degree): $\bar{k} = 2L/N = p\cdot(N-1) \cong pN$ also has a critical value which depends on N.

Degree distribution

The degrees k_i of the vertices i follow a binomial distribution with parameters p and $N-1$:

$$P(k_i = k) = C_{N-1}^{k} p^k (1-p)^{N-1-k}$$

P is the number of ways by which k edges can be connected to a certain node. The probability of having k edges is p^k and there are C_{N-1}^{k} equivalent ways of choosing k nodes at the ends of an edge. If i and j are different nodes, $P(k_i=k)$ and $P(k_j=k)$ are independent stochastic variables.

For high N values, with good approximation, the distribution $P(k)$ has, as a limit, the Poisson distribution (Figure 2.8):

---

[6] Percolation theory deals with fluid flow (or any other similar process) in random media. The medium can be represented as a set of regular lattice points. The percolation threshold is the critical fraction of lattice points that must be filled to create a continuous path of nearest neighbours from one side to another of the medium.

$$P(k) \approx e^{-pN} \frac{(pN)^k}{k!} = e^{-\bar{k}} \frac{\bar{k}^k}{k!}$$

The mean value $\bar{k}$ is: $\bar{k} = NC_{N-1}^{k} p^k (1-p)^{N-1-k}$ and $\sigma = \sqrt{\bar{k}}$. If the vertices are independent, the number of them having degree k is $X_k \approx N\,P(k_i = k)$.

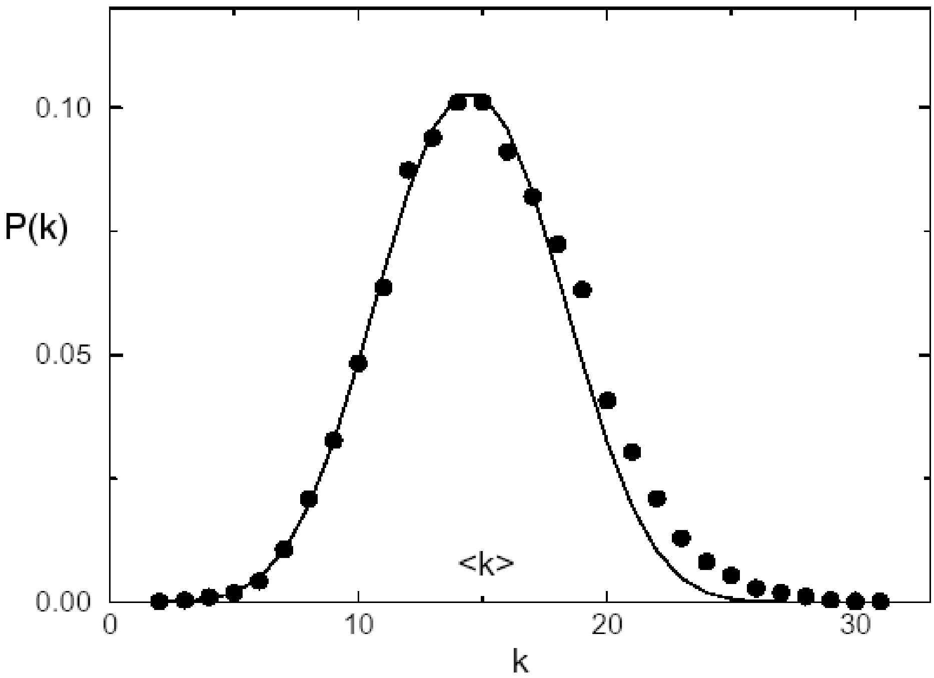


Figure 2.8 Degree distribution for a graph with N=10 000 and p=0.0015 compared with the expected value of a Poisson distribution (after Albert & Barabási, 2002)

Connection and diameter

The graph diameter D is the maximum possible distance between any two vertices. If the graph is not connected, and there are a certain number of isolated components, D = ∞. In this case it is possible to use the diameter of the greatest component as the graph diameter.

Random graphs tend to have small diameters if p is large enough. There is a good probability that the number of nodes with distance l from a given node is not too much smaller than $\bar{k}^l$. The diameter D depends on the logarithm of the node number:

$$D \propto \frac{\ln N}{\ln \bar{k}} = \frac{\ln N}{\ln pN}$$

Clustering coefficient

Given a random network *G(N,p)*, if we consider a vertex and one of its first neighbours, the probability of their connection is equal to the probability that any two nodes are connected. The clustering coefficient for such a network is therefore:

$$C_{G(N,p)} = p = \frac{\bar{k}}{N}$$

The ratio $C/\bar{k}$ as a function of N, drawn on a log-log paper, is a line with slope -1. This ratio is a constant for a given network (given N).

Generation of a graph with given degree distribution

A network with an assigned degree distribution can be built in the following way. A degree distribution p_k is defined, so that p_k is the fraction of nodes having degree k. The series of degrees will be the set of n values k_i for the nodes $j = 1, \ldots, n$. Each node j is assigned k_i outgoing 'segments'. Finally, n pairs of segments are randomly joined.

This process produces all possible topologies of a graph with a given degree distribution and equal probability. In the ensemble thus obtained, all graphs have the same weight.

2.5 Small-world networks

Despite the large size of a network, two vertices have often a relatively 'short' distance between them; in this case we speak of *small-world networks* (SW networks). A random graph (ER-like) is the simplest network which may show this effect; the typical distance between any two of the N nodes scales as $\ln(N)$.

The phenomenon of SW networks was first observed in social networks. Stanley Milgram (1967) studied the problem of estimating the number of acquaintances needed to pass a letter from one side to the other of continental USA. His conclusion was that the chain needed, on average, six persons: a remarkably low number for such a great distance and possible number of people involved.

The phrase ‘six degrees of separation’, and the expression ‘small world’ (the title of Milgram’s paper) have become commonly used to signify that, in a social context, any two persons are bound by a relatively small number of ‘steps’. In fact, there seems to be a high probability that even in a very wide social group, if two individuals are friends of a third person, there is a higher probability they are friends than any other two persons taken at random. This way, connected clusters are formed inside a larger network.

Just like a social network, a small-world graph is characterised by a relatively short average distance and a high clustering coefficient (the network is made up of dense clusters). In a random ER graph, the distance between two vertices is small on the average, but there is no tendency to create clusters (all edges are equiprobable). The modification of the degree distribution into a non uniform one does not imply the generation of clusters (vertices with high degree do not necessarily pile up in dense subgraphs).

The model by Watts and Strogatz (1998) describes the generation of a small-world network. The initial graph is a regular *circulant graph* $G_{C:N,k}$ in which a vertex v (of the N) is connected to $2k$ neighbours (k in each direction). A certain number of edges $\{v,w\}$ chosen at random are reconnected (*rewired*) with probability p to other vertices of the graph (multiple edges are excluded). The process is repeated k times.

To avoid obtaining an unconnected graph, the condition $1 << \ln N << 2k << N$ is imposed ($2k >> \ln N$ guarantees that the resulting graph is connected). The reconnection can be obtained by substituting the chosen edge or by adding new edges to the original graph. In this case nk pairs $\{v,w\}$ of nodes are taken randomly and nk edges with probability ϕ are added, getting $nk\phi$ new shortcuts.

The two techniques can be combined (Figure 2.9).

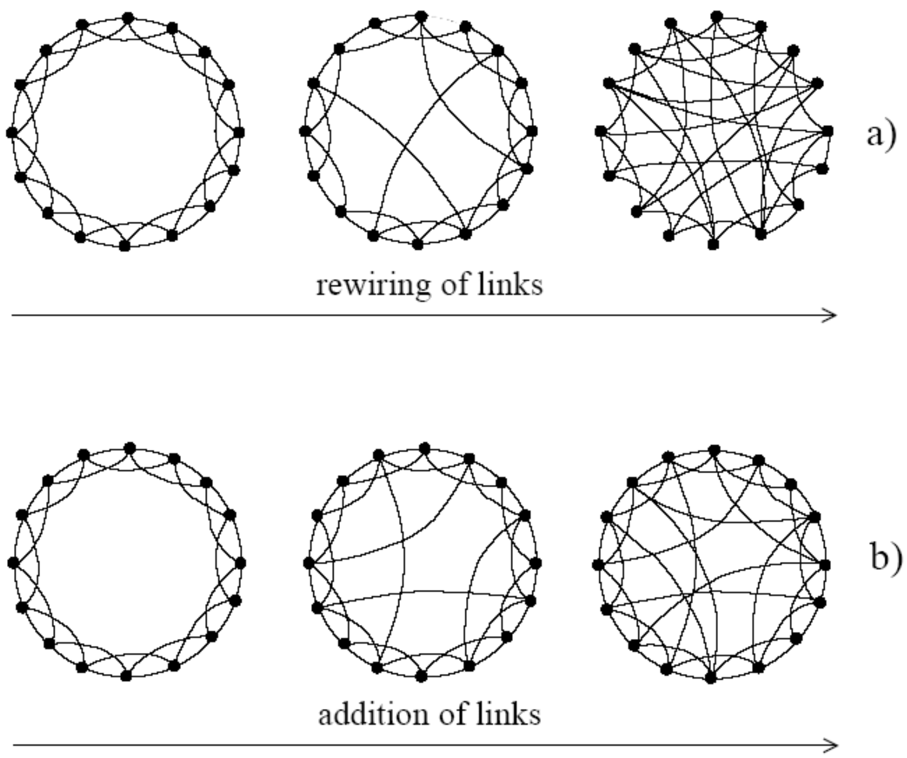


Figure 2.9 Generation of a small-world network obtained by rewiring (a) existing links or adding (b) new edges (after Dorogovtsev & Mendes, 2002)

In the starting graph $G_{C:N,k}$ the clustering coefficient is $C = 3(k - 2)/4(k - 1)$, which converges to $C=3/4$ for large k values. The average distance in such a graph is proportional to $N^{1/d}$ (d = dimension), which grows faster than ln(N).

If we study $C(p)$ e $l(p)$ (clustering and average distance as a function of the rewiring p), we find that, for the initial regular lattice $l(0) \sim N/2k >> 1$ *e* $C(0) \sim 3/4$, l scales linearly with the order of the graph and C is rather large. For $p \to 1$ the graph converges to a random graph in which: $l(1) \sim ln(N)/ln(k)$ and $C(1) \sim k/N$ (total number of edges is constant), therefore l has a logarithmic behaviour and C decreases with N. In other words, a large C is associated with a large l and a small C with a small l.

However, numerical simulations show that there is a range of p values in which $l(p) \sim l(1)$ but $C(p) >> C(1)$. In this regime the graphs have clustering and short average distances (Figure 2.10); in this range we speak of SW networks.

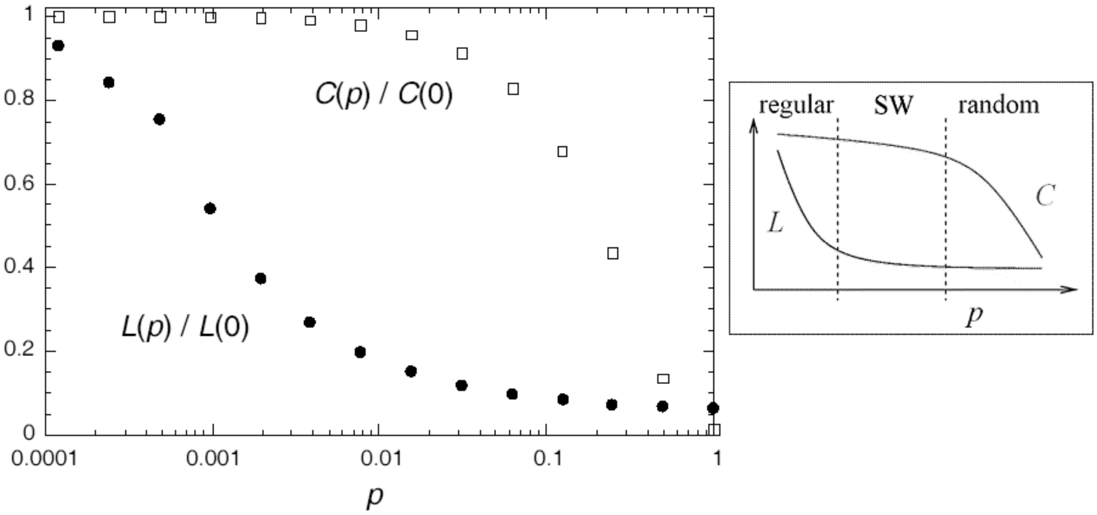


Figure 2.10 Clustering coefficient and average distance in SW networks (after Watts & Strogatz, 1998)

2.5.1 Features of small-world networks

Characteristic distance

The appearance of *shortcuts* between distant nodes causes a sudden fall in the characteristic distance l, while locally, the network stays ordered. The transition does not happen until $p \geq 2N/k$ (or at least a shortcut exists). The critical probability p_c depends on N and on a *crossover size* N^* (Figure 2.11).

$$l(N,p) \approx N^* F(N/N^*) \quad \text{where} \quad F(u) = \begin{cases} u & \text{if } u << 1 \\ \ln u & \text{if } u >> 1 \end{cases}$$

From numerical simulations $N^* \sim p^{-\tau}$ with $\tau = 1/d$ (d = lattice dimension) for a circle $d=1$. Therefore, the small-world behaviour appears at $p_c \sim 1/N$. The situation is similar to the critical probability at which a giant component is formed in an ER network or to the bond percolation critical probability.

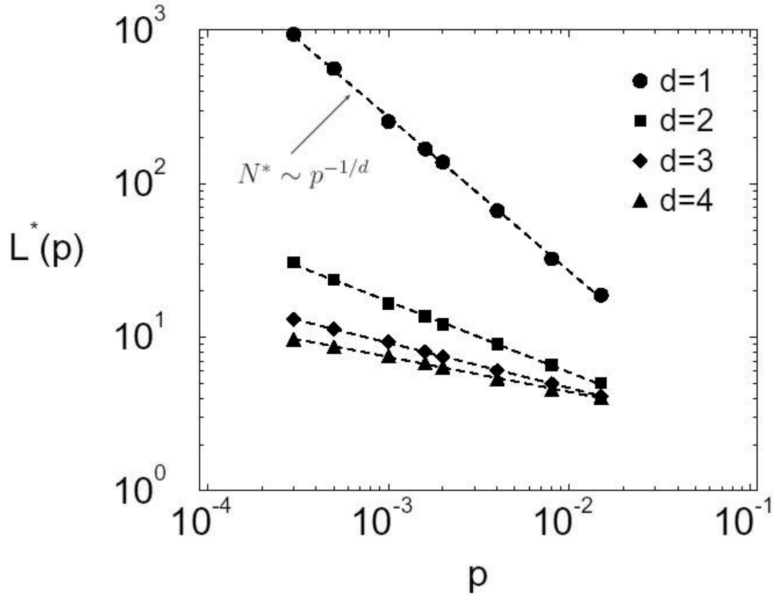


Figure 2.11 Crossover size for different dimensions of the lattice (after Albert & Barabási, 2002)

In summary:

$$l(N,p) \approx \frac{N}{k} f(pkN^d)$$

where

$$f(u) = \begin{cases} \text{constant} & \text{if } u << 1 \\ (\ln u)/u & \text{if } u >> 1 \end{cases}$$

For the one-dimensional case, for values of u far from *1* (mean field approximation):

$$f(u) = \frac{4}{\sqrt{u^2+4u}} \tanh^{-1} \frac{u}{\sqrt{u^2+4u}}$$

The variable pkN^d represents the double of the average number of shortcuts in the graph (given p) and $f(u)$ is the average of the fraction with which the distance between two vertices reduces for a certain value u.

In an SW network with N nodes and L edges, the average distance l depends on the average $\langle s \rangle$ of the Euclidean shortest distances between two diametrically opposed vertices and on its square $\langle s^2 \rangle$:

$$\frac{l}{N} = \frac{\langle s \rangle}{N-1} - \frac{\langle s^2 \rangle}{L(N-1)}$$

Clustering coefficient

SW networks, besides having a short characteristic distance, have a relatively high clustering coefficient (see the definition formula in section 2.3.2).

For both versions of SW networks (with and without rewiring), the clustering coefficient is (Figure 2.12):

$$\textit{with rewiring:}\ C = \frac{3(k-1)}{2(2k-1)}(1-p)^3$$

$$\textit{without rewiring:}\ C = \frac{3(k-1)}{2(2k-1)+4kp(p+2)}$$

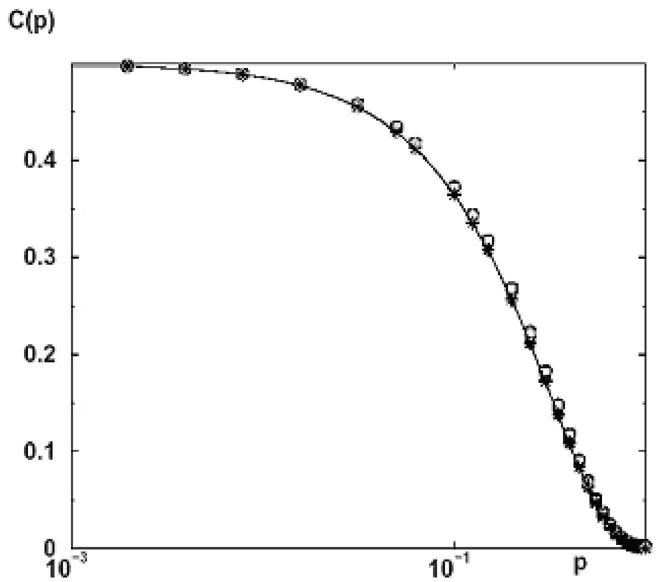


Figure 2.12 Clustering coefficient vs p for networks with different N (after Barrat & Weigt, 2000)

Degree distribution

The degree distribution derives from the generation of an SW network. At the beginning ($p=0$) all vertices have the same degree, the distribution is a δ centred on k, then, with probability $p \neq 0$, shortcuts are added. They introduce some 'noise' which widens the distribution. For $k>2$, unlike in ER networks, there are no isolated nodes.

The SW degree distribution is given by:

$$P(j) = \sum_{n=0}^{\min(j-k,k)} \binom{k}{n} (1-p)^n p^{k-n} \frac{(pk)^{j-k-n}}{(j-k-n)!} e^{-pk}$$

for $j \geq k$; $P(j) = 0$ for $j < k$.

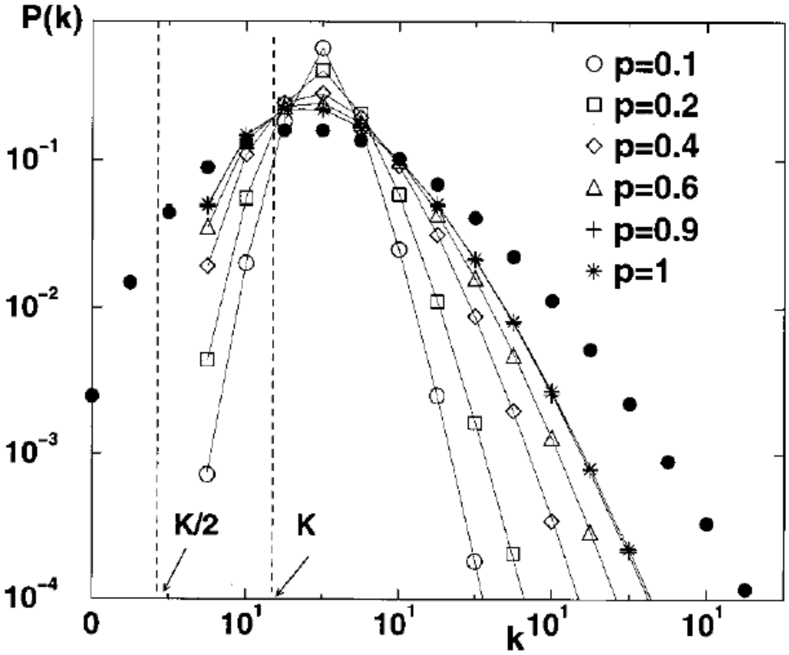


Figure 2.13 Degree distribution in SW networks for $k=3$, $N=1000$ and for a corresponding ER network (full circles) (after Albert & Barabási, 2002)

The shape of the distribution is similar (Figure 2.13) to that of a random network (ER) with a peak at $\bar{k} = K$. The topology of an SW network is homogeneous, and most nodes have about the same number of links.

Other characteristic parameters

Proximity ratio can be seen as a small-worldness index. It is defined as the ratio between clustering coefficient and characteristic distance normalised to the values the same network would have in the hypothesis of full randomness (ER):

$$\mu = \frac{C\, l_{rand}}{C_{rand}\, l}$$

Connectivity length is obtained by generalising to weighted networks and introducing a separation distance $d_s(u,v)$ as the smallest physical (weighted) distance in the set of paths between two vertices u and v (if the network is not connected, some distances may be infinite). The connectivity length D is the harmonic mean of the d_s:

$$D(G) = \frac{N(N-1)}{\sum d_s(u,v)^{-1}}$$

The behaviour of D is similar to that of l if calculated globally, to $1/C$ on a local scale.

Efficiency: Besides the adjacency matrix A, it is possible to define a distance matrix L whose elements lij are the physical distances between i and j (i.e. total weight of the connection, strength of the interaction etc.). If the graph is unweighted, L is the unit matrix I. The cost matrix is: $D = A \cdot L$, where $d_{ij} = a_{ij} \cdot l_{ij}$.

The efficiency of a pair ij of distinct vertices is (Latora & Marchiori, 2001):

$$\varepsilon_{ij} = 1 / d_{ij}$$

The average efficiency of the graph G is the average of the efficiencies of its vertices:

$$\mathrm{E}(G) = \frac{1}{N(N-1)} \sum_{i \neq j} \varepsilon_{ij} = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}}$$

This *global efficiency* can be normalised with respect to that a complete graph K would have in which $d_{ij} = l_{ij}\ \forall i,j \in V$ (the maximum possible):

$$\mathrm{E}(K) = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{l_{ij}}$$

therefore:

$$E_{glob}(G) = E(G) / E(K)$$

E_{glob} varies in $[0,1]$. At a local level, similar quantities can be calculated. If we consider the subgraph G_u of the neighbours to u, normalising to $K_{deg(u)}$, the *local efficiency* is:

$$\mathrm{E}_{loc} = \frac{1}{N} \sum_{u} \mathrm{E}_u$$

The efficiency of a network is an approximation of C and l (normalised, Figure 2.14); it has the advantage of being computable even for an unconnected graph (for which $\varepsilon = 0$ instead of having an infinite value).

In systems with small differences in the node-to-node distances $E_{glob} \sim 1/l$, if the graph has many dense subgraphs, $E_{loc} \sim C$.

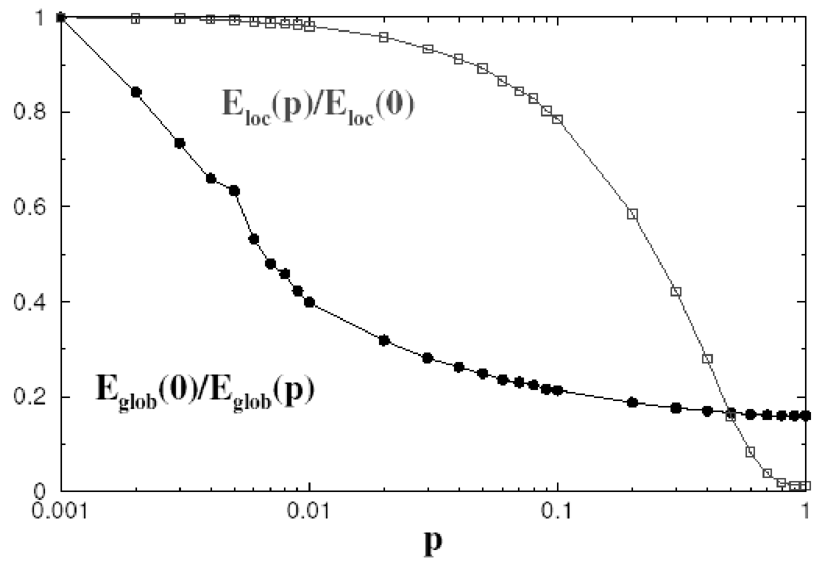


Figure 2.14 Local and global efficiencies as functions of the rewiring probability (after Latora & Marchiori, 2001)

It is possible to introduce a cost evaluation function $\gamma(l_{ij})$ that represents the cost associated with the generation of a link. The total graph cost *Cost(G)* is (it includes a normalisation to *K* and varies on [*0,1*]):

$$Cost(G) = \frac{\sum_{i \neq j} a_{ij}\gamma(l_{ij})}{\sum_{i \neq j} \gamma(l_{ij})}$$

If $\gamma(x) = x$ (identity function), the cost could be considered proportional to the connection length. An *economical* SW network is a network with *high* E_{glob} and E_{loc} and low *Cost* (Figure 2.15).

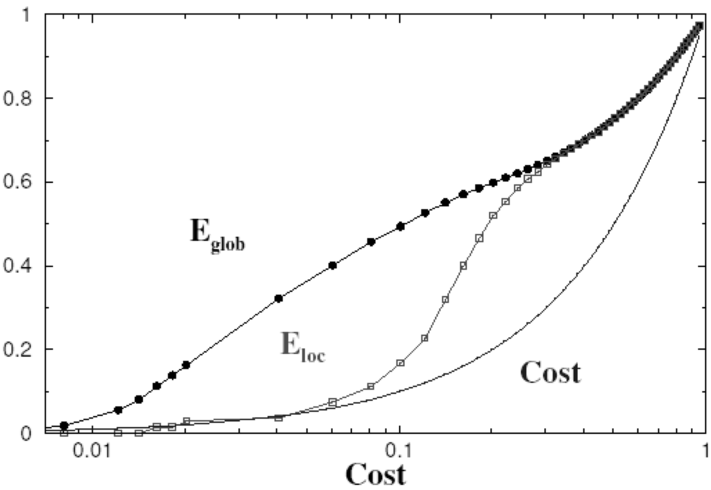


Figure 2.15 Simulation of E_{glob}, E_{loc} and Cost (after Latora & Marchiori, 2001)

As for the proximity ratio μ, an *efficiency index* can be defined:

$$\eta = E_{glob}(rand)\ E_{loc} \,/\, E_{loc}(rand)\ E_{glob}$$

The values calculated for real networks are in good agreement with the corresponding μ values.

2.6 Scale-free networks

Many real networks have a degree distribution which is significantly different from the expected Poisson or exponential distribution (typical of the ER random networks).

Even cases in which an exponential tail exists show a power-law behaviour such as $P(k) \sim k^{-\gamma}$ (Figure 2.16 and Figure 2.17) . These networks do not have a characteristic scale parameter $\bar{k}$ (the mean of a Poisson distribution). They are, therefore, denoted as *scale-free* (SF) or scale invariant.

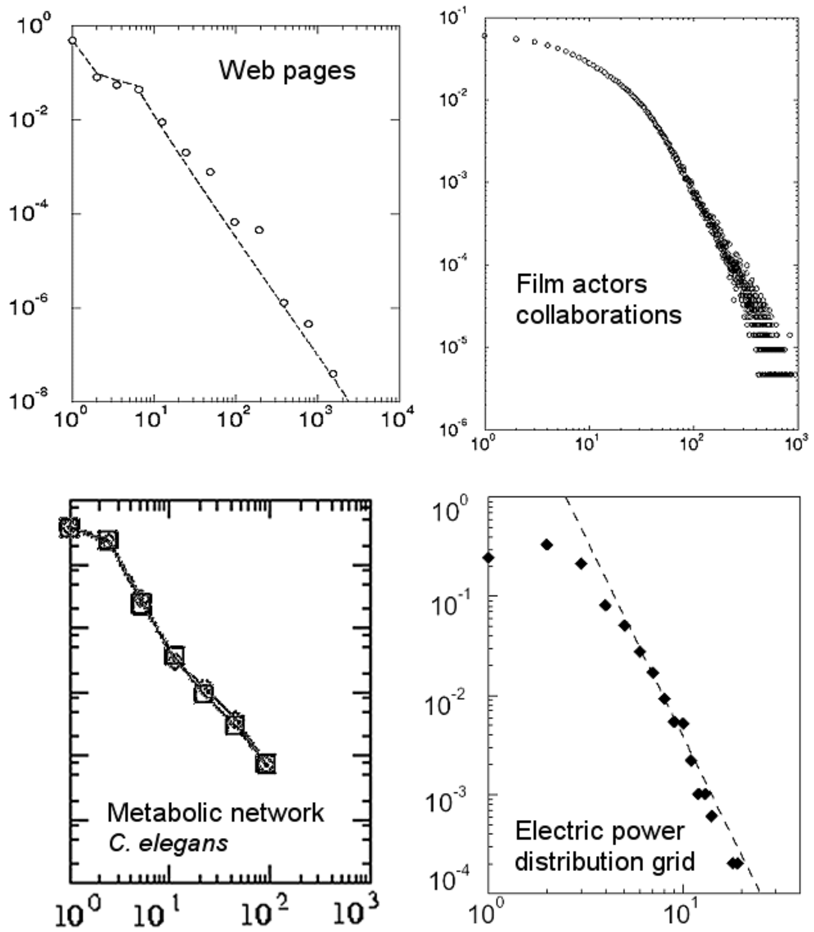


Figure 2.16 Examples of degree distributions of real world networks (after Albert & Barabási, 2002)

The ER and SW models described above are not able to explain the generation and the evolution of these networks. The problem was first challenged by Barabási and Albert (1999), whose model renders the main characteristics of such networks. Its importance is in the accuracy with which it is capable of simulating open dynamic systems, that is, systems that change their shape by adding (or removing) elements and connections among them. Moreover, SF networks are able to model the relationships among network elements that, in many cases, are not random at all, but follow definite rules.

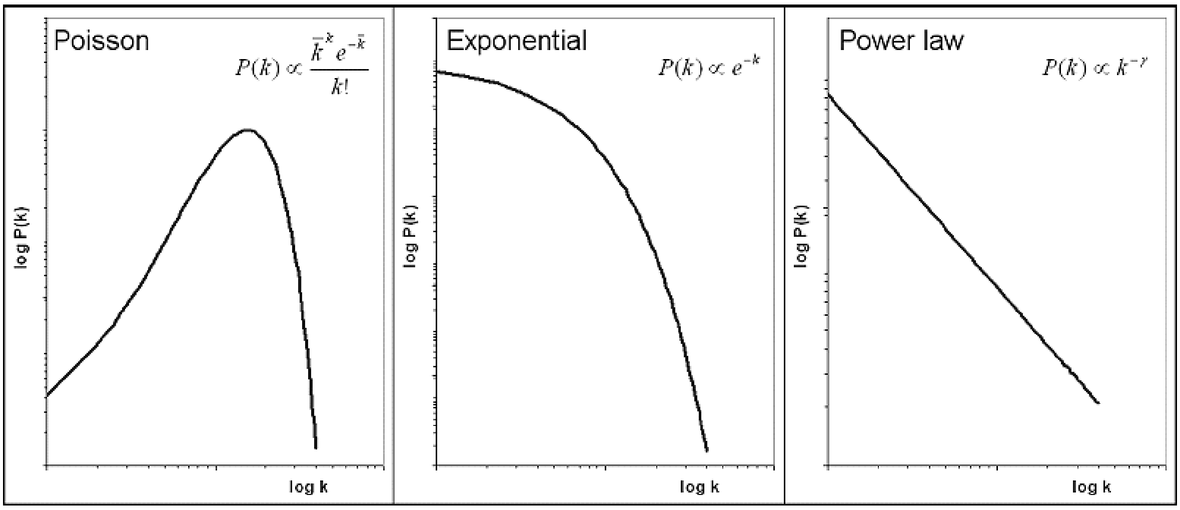


Figure 2.17 Distributions: Poisson, exponential and power-law (drawn on a log-log diagram)

The SF model is a dynamic one. In contrast to ER and SW models, which allow the generation of an ensemble of graphs with the desired topological characteristics, the SF model tries to explain the evolution of a complex network over time. The main underlying hypothesis is that, by explaining the temporal evolution, it is also possible to obtain a correct description of the topology of the resulting network.

Two main methods for the generation of a SF network exist: stochastic and deterministic (Albert & Barabási, 2002; Dorogovtsev & Mendes, 2003).

The stochastic method consists of the application of two rules (Figure 2.18):

- *growth*: starting with a small number of nodes m_0, at each time interval t a new node (or more than one) is added. This one connects with $m \leq m_0$ edges to m existing vertices;
- *preferential attachment*: the choice of the connections is not random, but the probability with which a new node attaches to an old one i depends on the degree k_i of i so that:

$$\Pi(k_i) = \frac{k_i}{\sum_j k_i}$$

After t time intervals, the network has $N = t + m_0$ nodes and mt edges. With numerical simulations it is shown that the probability that a vertex has k edges follows a power law with exponent $\gamma_{SF} = 3$ (i.e.: $P(k) \sim 2m^2\ k^{-3}$). The scaling exponent is independent from m and it is the only parameter of the model.

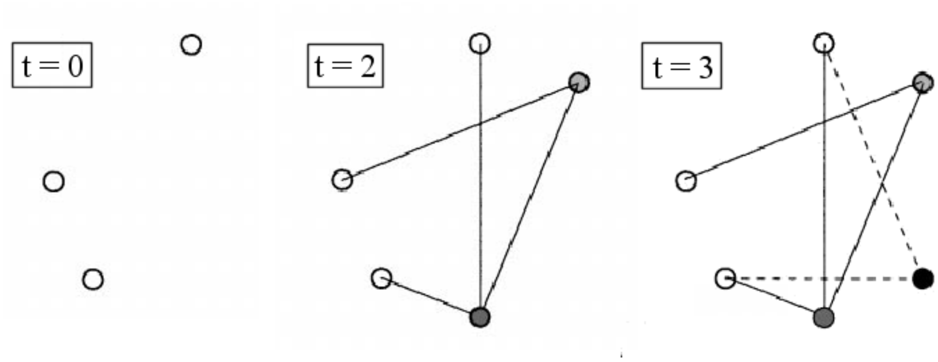


Figure 2.18 Stochastic generation of a SF network

The second method (deterministic) follows a hierarchical rule: at each step the elements of the previous one are reused. We start from a single node, the root of the graph, and at each interval the existing elements are tripled and all vertices are connected to the root. This way it is possible to obtain a distribution $P(k) \sim k^{-\gamma}$ with $\gamma = \ln 2 / \ln 3$ with an average number of connections, at step n (for a sufficiently large n), equal to $\bar{k} = 4n/3$ (Figure 2.19).

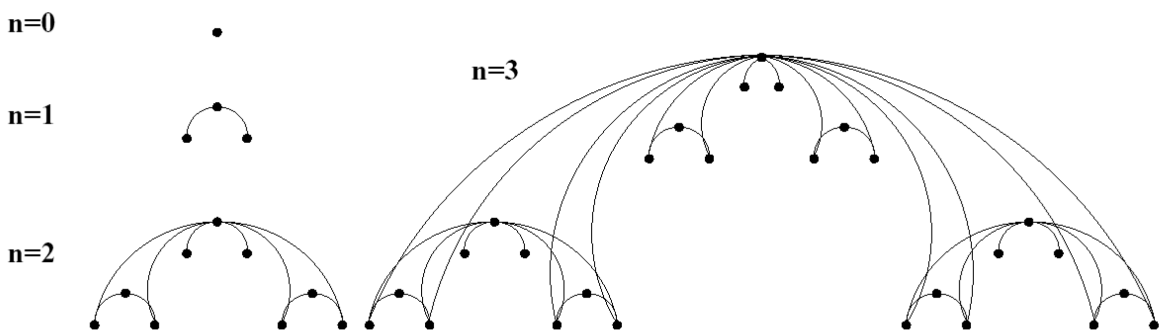


Figure 2.19 Deterministic generation of a SF network

2.6.1 Features of scale-free networks

Characteristic distance

The characteristic distance of an SF network grows logarithmically with the number of nodes N; a generalised form for the relation is (Figure 2.20):

$$l = A\ ln(N - B) + C$$

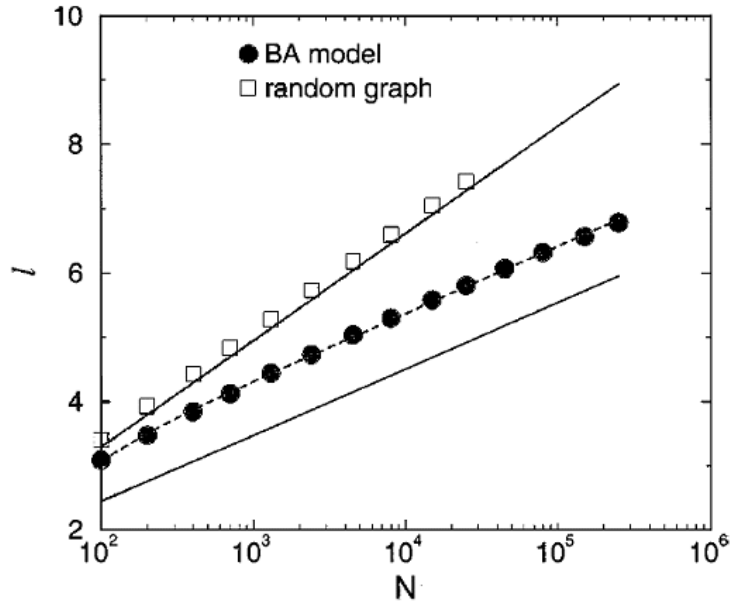


Figure 2.20 Average distance l as function of N, a comparison between SF and ER models with same size and mean degree (after Albert & Barabási, 2002)

Node-degree correlation

In a random graph there is no correlation between vertices and degrees. In a SF network these correlations are present.

Let us consider nodes with degree k and nodes with degree l, n_{kl} is the number of pairs connected with degrees k and l. We have: $n_{kl} \neq n_k \cdot n_l$. That is, correlations exist, while for random networks $n_{kl} = n_k \cdot n_l$. The only case in which a simple factorisation exists is the one in which $1 << k << l$ for which $n_{kl} \approx k^{-2}\ l^{-2}$ (in the absence of correlations it would be: $n_{kl} = k^{-3}\ l^{-3}$). The SF generation process creates non trivial correlations among the network vertices.

Clustering coefficient

There is no analytical expression for C.

The coefficient decreases with the order N of the network approximately as $C \sim N^{-0.75}$. In case of a random graph we would expect $C = \bar{k} N^{1}$. C is independent from N in a small-world (Figure 2.21).

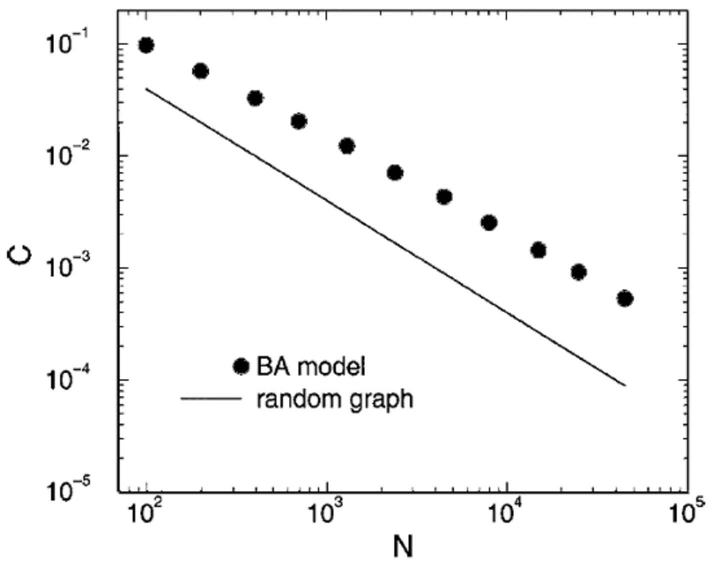


Figure 2.21 Clustering coefficient in SF and ER networks (after Albert & Barabási, 2002)

Assortative networks

In social network analysis, *assortative mixing* (or linking) refers to selective connections in which actors of a similar degree are connected among them.

Let *P(k'|k)* be the joint probability that a neighbour of a vertex with degree *k* has degree *k'*. A correlation measure is:

$$K_{nn}(k) = \sum_{k'} k' P(k'|k) .$$

If this value is growing, the network is said to be assortative; if it is constant the network is not assortative; if it is decreasing the network is disassortative (there is a tendency for a node with high degree to connect to nodes with low degrees). An assortativity coefficient is given by the Pearson correlation coefficient between the degrees k_i and k_j of the nodes (Newman, 2002a):

$$r \propto \overline{k_i k_j} - \overline{k_i}\ \overline{k_j}$$

with the meanings: $r>0$ for assortative networks, $r=0$ for non assortative networks and $r<0$ for disassortative networks.

Numerical simulations with joint probabilities of the type (inversely proportional to *k*):

$$P(k'|k) \propto \frac{1}{|k-k'|+1}$$

bring to a degree distribution with exponent:

$$\gamma(p) = 2 + \frac{p}{2-p}$$

for $p \to 1$, $\gamma \to 3$ (as for the standard SF model).

With an exponential probability:

$$P(k'|k) \propto e^{-|k-k'|}$$

there are two different behaviours. For high p values, the distribution is still scale-free. At low p we find a power law if k is small, a peaked distribution if k is high. The transition happens at $p \sim 0.5$ (Figure 2.22).

The assortativity of a network is directly connected to its robustness (resilience) properties; the more assortative a network is, the higher its resilience (Newman, 2002a, 2002b).

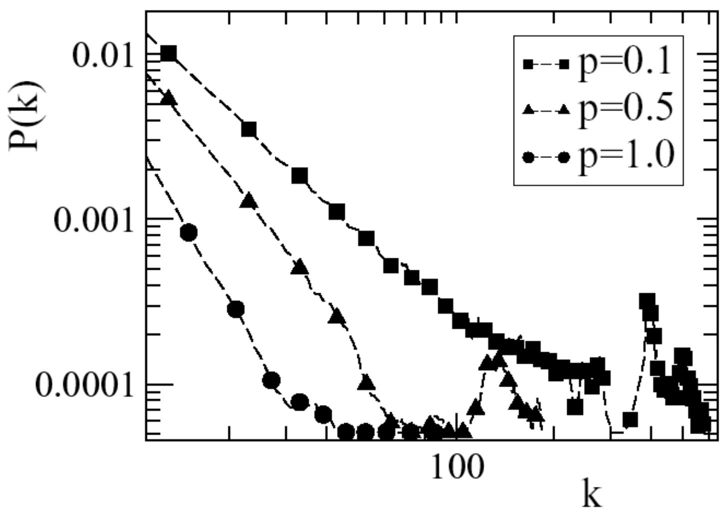


Figure 2.22 P(k) behaviour for assortative networks at different degree correlation probabilities (after Catanzaro et al., 2004)

The betweenness distribution has a similar power-law behaviour $P(b) \propto b^{-\eta}$ ($\eta = 2.2$). It depends on the degree k in an analogous way: $b(k) \propto k^{-(\gamma-1)(\eta-1)}$ (k exponent = 1.31).

The two relations are shown in Figure 2.23 for probabilities with inverse proportionality (inv), exponential (esp) and for data on scientific collaborations for physics papers published in the *Arxiv/cond-mat* online archive (cm).

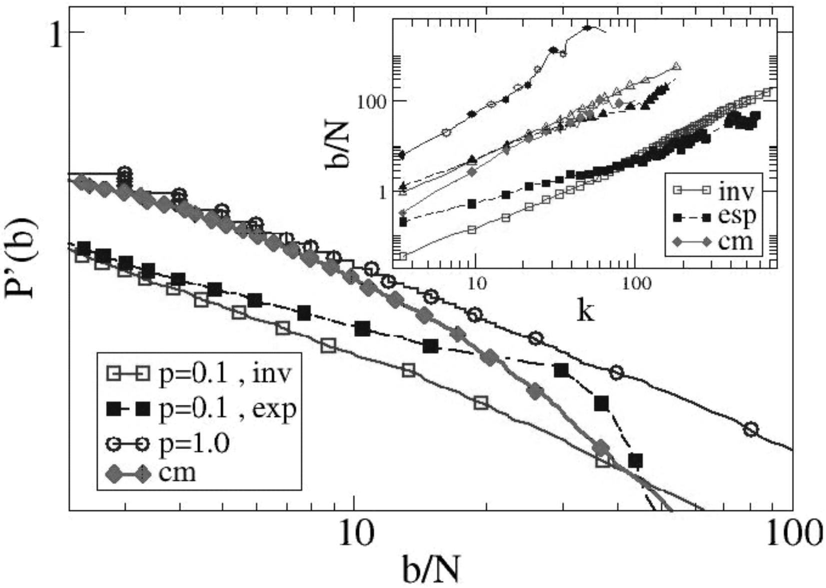


Figure 2.23 Betweenness distribution and relation with *k* (after Catanzaro et al., 2004)

2.6.2 Graph spectra

The spectrum of the connectivity (adjacency) matrix of a graph is an important element in the study of networks as it contains much information about a network's topology (see section 2.3.4).

A graph *G(N,p)* with *N* vertices has *N* eigenvalues. Let us consider a random graph in which $p(N) = cN^{-z}$. For $z < 1$ there is an infinite cluster and for $N \to \infty$ every node belongs to it. In this case the spectral density $\rho(\lambda)$ converges to a semicircular distribution (de Aguiar & Bar-Yam, 2005; Farkas et al., 2001).

This distribution (see Figure 2.24) is commonly referred to as Wigner's Law (after the papers by Eugene Wigner, 1955; 1958):

$$\rho(\lambda) = \begin{cases} \dfrac{\sqrt{4\mathrm{Np}(1-\mathrm{p}) - \lambda^2}}{2\pi Np(1-p)} & \text{if } |\lambda| < 2\sqrt{Np(1-p)} \\ 0 & \text{otherwise} \end{cases}$$

The largest eigenvalue (principal) λ_1 is separated from the main part of the spectrum and grows with the size of the network as pN. When $z>1$ the spectral density is different from a semicircle. The odd moments of $\rho(\lambda)$ are null; this indicates that the only way to 'go back' to a node is to follow the same path, i.e. there are no closed paths, therefore the graph is composed of trees.

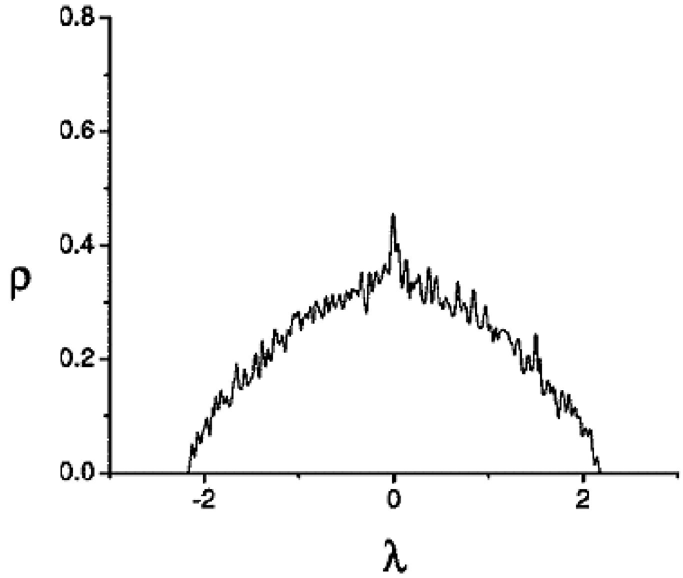


Figure 2.24 Spectral density for a random graph (N = 1024). The peak corresponds to the principal eigenvalue (after de Aguiar & Bar-Yam, 2005)

In a small-world network, the spectrum depends on the rewiring probability p (Figure 2.25).

At $p=0$ the network is a regular periodic lattice, and the spectral density $\rho(\lambda)$ has many singularities. At values of p between 0 and 1 the singularities go 'out of focus' and when $p \rightarrow 1$ it gets closer to the semicircle typical of ER networks.

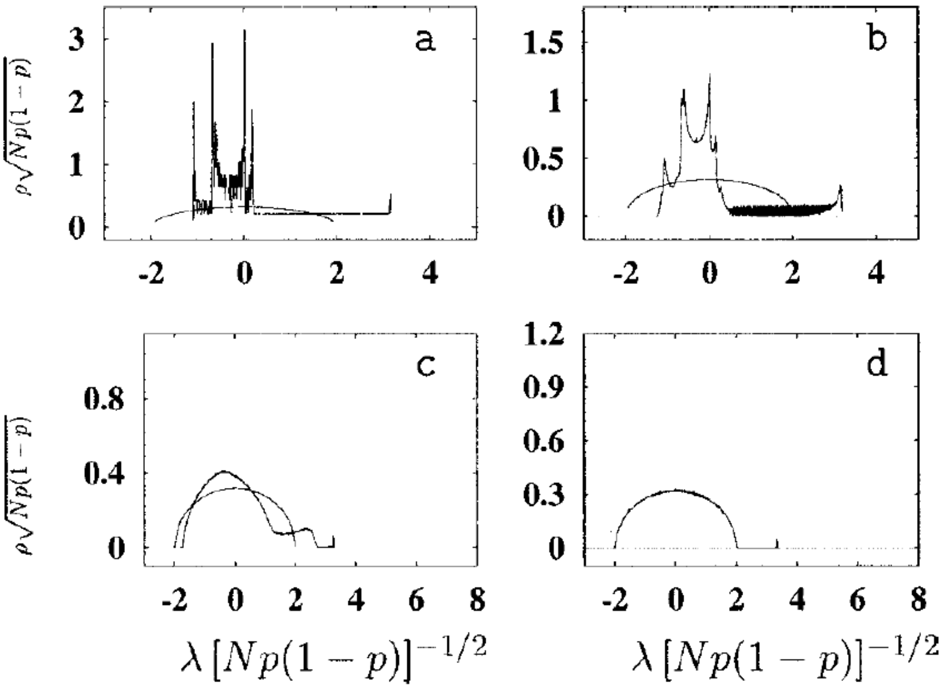


Figure 2.25 Spectral densities for a SW network with *p=0*, *p=0.01*, *p=0.3* and *p=1* (after Albert & Barabási, 2002)

The spectral density of a scale-free network has a continuous behaviour, and it is remarkably different from the semicircular shape typical of ER networks. Numerical simulations (de Aguiar & Bar-Yam, 2005) show that most of the density has a triangular shape (Figure 2.26).

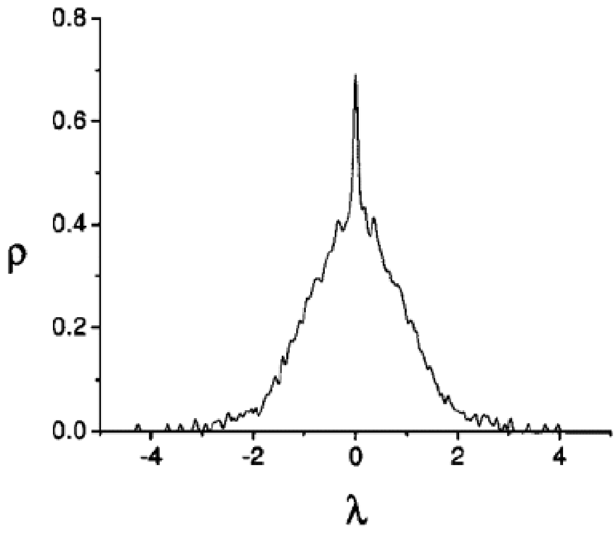


Figure 2.26 SF network spectral density (after de Aguiar & Bar-Yam, 2005)

The tails follow a power law, and the principal eigenvalue λ_1 is clearly separated from the rest of the distribution. It plays an important role in the moments of $\rho(\lambda)$ determining a loop structure in the network. The fraction of loops with more than four edges increases with N, that of triangles decreases in the thermodynamic limit ($N \rightarrow \infty$). A lower bound for λ_1 is $\sqrt{k_1}$ (k_1 = maximum degree in the network); since in an SF network the degree values follow $N^{1/2}$, we have $\lambda_1 \propto N^{1/4}$.

2.7 Conclusions

The observed non-linearity of the relationships among the numerous components of social and economic environments and the difficulty in stating simple laws to study the behaviour of such systems calls for a number of methods and techniques assembled in what is known as *science of complexity.*[7]

A number of tools have been developed in recent years to describe a complex system and they can be grouped in three main areas:

- nonlinear dynamics;
- statistical physics; and
- network theory.

The last one has recently received attention and has provided useful results in many fields, since many complex systems can be described in terms of networks of interacting elements (sections 2.1 and 2.2).

Intuitively, the concept of network is a spatial one. However, since the very beginning of the study of graphs, it has been realised that the description of such objects, and the solution to the problems associated with them, did not correspond well with their metric properties. Graph theory, the mathematical framework for the science of networks, was born and has evolved as a topological discipline. This means that many of the characteristics of a network and of the systems that can be represented by a network depend exclusively on the network form. Systems sharing similar characteristics will have

[7] This section is an edited version of the introduction to the book: Network Analysis and Tourism: From Theory to Practice, by N. Scott, C. Cooper and R. Baggio, Channelview (2008b)

similar properties and behaviours, independent of the nature of their constituents (section 2.3).

The distribution *P(k)* of the vertex degrees (number of connections), the clustering coefficient *C* (interpretable as a local density) and the average length *L* of the shortest paths between any two vertices have been found to be characterising parameters for the topology of a network.

The first mathematical model, used for many years to describe almost all kinds of networks, come from Erdős and Rényi. According to the ER model, a graph is a set of nodes connected two at a time with a certain probability. The random distribution of the node degrees *P(k)* follows, for large sizes, a Poisson law. This implies that most vertices have about the same number of links (the average degree $\langle k \rangle$), while nodes with a degree that deviates significantly from the average are extremely rare. The *tail* (high *k* region) of the degree distribution *P(k)* decreases exponentially.

The clustering coefficient *C* depends on the average degree of the network $\langle k \rangle$. It is therefore constant for a given graph. The mean path length *L* is proportional to the logarithm of the network order, $L \propto \log n$. An ER graph is in equilibrium; the number of vertices is fixed. A possible evolution is obtained by varying the connection probability. One of the most interesting results of this model is that a number of peculiar characteristics depend strongly on special values of *p* (section 2.4).

At the end of the 1990s, empirical studies confirmed that, in many cases, ER random graphs are quite different from *real* world networks. Subsequently, numerous investigations have been done and a great variety of theoretical works have been published. The first evolution of the ER model is the one proposed by Watts and Strogatz (1998).

Watts and Strogatz (1998) noticed that a number of examples exhibited clustering coefficients significantly higher and mean shortest path lengths lower than expected. *Small-world* (SW) networks arise as the result of a random replacing (rewiring) of a fraction *p* of the links of a regular lattice with new random connections. In this evolutionary process they position themselves between the two limiting cases of a regular lattice and a random graph. An SW network is still characterised by a Poissonian degree distribution; the local neighbourhood is preserved and the diameter increases logarithmically with the number of vertices. This is why they are called small-world networks: it is possible, on the average, to connect any two vertices through just a few links (section 2.5).

The analysis of other real world networks (Internet routers and web pages, in particular) led to the discovery that such systems had a peculiar characteristic. Their degree distribution approximates a power law: $P(k) \propto k^{-\gamma}$. The distribution is largely uneven and there is no characteristic mean nodal degree (the mean of a Poissonian ER or SW distribution). A few nodes act as very connected hubs, having a very large number of ties, while the majority of nodes have a small number of links. The absence of a characteristic average degree $\langle k \rangle$, the characteristic *scale* of the network, has gained these networks the name *scale-free* (SF) , or scale-invariant networks.

SF networks are dynamic systems; they grow with the addition of new nodes and new links that are not distributed randomly, but follow specific mechanisms. The most commonly invoked is a *preferential attachment* in which a new node has a higher probability to attach to one of the most connected ones. Deviations from pure power laws (kinks, cut-offs etc.), which can be observed, are generally explained by introducing corrections or nonlinear terms in the expression of the preferential attachment, by considering the limitations given by the finiteness of the network size or by assigning special properties (age, fitness or attractiveness) to some of the network actors.

The network thus created does not have an intrinsic modularity (the clustering coefficient C is independent of the degree k). Scale-free networks with degree exponents $2 < \gamma < 3$, values found in most real cases, have very small average path lengths (typically $L \propto log\ log\ N$, significantly shorter than the one typical of SW networks).

Other complex networks, mainly if they represent social or biological systems, exhibit multiple characteristics such as modularity, high local densities and scale-free topology. The nodes are part of highly clustered areas, with few hubs that are responsible for connecting the different neighbourhoods. Hierarchical models of network formation have been devised, assuming that clusters mix in an iterative way and create a hierarchical structure. The resulting network has a power-law degree distribution and a large average clustering coefficient. More importantly, C follows a power law $C(k) \propto k^{-1}$. This latter characteristic is thought to be the signature of a hierarchical network (section 2.6).

Table 2.1 shows some examples, reported in the literature, of real networks along with some of the main parameters measured (adapted from Albert & Barabási, 2002; Boccaletti et al., 2006; Caldarelli, 2007; Dorogovtsev & Mendes, 2003; Pastor-Satorras & Vespignani, 2004).

Type	Network	N	M	δ	<k>	L	C	E_{glob}	E_{loc}	r	γ	γ-in	γ-out
B	Caenorhabditis elegans	282			14.0	2.65	0.28	0.46	0.47				
B	Escherichia coli (metabolic)	778			7.4	3.2					2.2		
B	Food web Littlero Lake*	92	997	0.2382	10.8	1.9	0.087			-0.326			
B	Food web Silwood park	154			4.8	3.4	0.15				1.1		
B	Food web Ythan estuary	134			8.7	2.43	0.22				1.1		
B	Metabolic network	765	3686	0.0126	9.6	2.56	0.67			-0.240	2.2		
B	Protein interactions	2115	2240	0.0010	2.1	6.8	0.071			-0.156	2.4		
B	Saccharomyces cerevisiae (neural network)	1870			2.4							2.4	2.4
S	Coauthors, MEDLINE	1500000			18.1	4.6	0.066						
S	Coauthors, physics	52909	245300	0.0002	9.3	6.19	0.56			0.363			
S	Coauthors, SPIRES	56627			173.0	4	0.726				1.2		
S	Company directors	7673	55392	0.0019	14.4	4.6	0.88			0.276			
S	Email messages	59912	86300	0.0000	1.4	4.95	0.16					1.5	2
S	Karate Club	34	78	0.1390	4.6					-0.476			
S	Movie Actors 1	449913	3E+07	0.0003	113.4	3.48	0.78			0.208	2.3		
S	Movie Actors 3	225226			61.0	3.65	0.79	0.37	0.67		2.3		
S	Sexual contacts	2810				3.2						3.4	3.4
S	Telephone calls	4.7E+07	8E+07	0.0000	3.2						2.1		
S	Word co-occurrence	460902	2E+07	0.0002	70.1	2.67	0.44				2.7		
S	Words Roget's Thesaurus	1022	5103	0.0098	5.0	4.87	0.15			0.157			
T	Electronic circuits	24097	53248	0.0002	4.3	11.05	0.03			-0.154	3		

Type	Network	N	M	δ	<k>	L	C	E_{glob}	E_{loc}	r	γ	γ-in	γ-out
T	Internet routers	10697	31992	0.0006	6.0	3.31	0.39			-0.189	2.5		
T	Internet, AS 2	6474			4.1	3.76	0.32	0.29	0.26				
T	Peer-To-Peer network	880	1296	0.0034	1.5	4.28	0.011			-0.366	2.1		
T	Power grid	4941	6594	0.0005	2.7	18.99	0.08			-0.003			
T	Software packages	1439	1723	0.0017	1.2	2.42	0.082			-0.016		1.4	1.6
T	WWW	325729			4.5	11.2		0.28	0.36		2.5	2.1	2.5
T	WWW (Altavista)	2E+08	2E+09	1.0E-07	10.5	16.18						2.1	2.7
T	WWW pages (nd.edu)	269504	1E+06	4.1E-05	5.6	11.27	0.29			-0.067		2.1	2.4
T	WWW, sites	153127			35.2	3.1	0.108					1.94	

Table 2.1 Main characteristics of some real networks from the literature (adapted from Albert & Barabási, 2002; Boccaletti et al., 2006; Caldarelli, 2007; Dorogovtsev & Mendes, 2003; Pastor-Satorras & Vespignani, 2004). Type: B = biological, S = social, T = technical; N = number of nodes, M = number of edges, δ = link density, <k> = mean degree, L = average path length, C = clustering coefficient, Eglob = global efficiency, Eloc = local efficiency, r = assortativity coefficient, γ = exponent of power-law degree distribution (undirected networks), γ-in = indegree distribution exponent, γ-out = outdegree distribution exponent (directed networks). More extensive definitions and explanations of these measurements are discussed in Chapters 3 and 4

3 Theoretical background: network dynamics

A complex adaptive system such as a tourism destination is a dynamic entity. It undergoes a number of different evolutionary processes due to internal mechanisms or to external events during which its structure may change. New elements may be added to the system, different sets of relationships may be established, and parts of the system may be deleted or inhibited. Moreover, various processes may occur which use the web of connections among the system's components as a medium, and the topology of this web can influence their development (Boccaletti et al., 2006; Ebel et al., 2003). Therefore, it is important to study the evolution and the behaviour of a network (representing the system) from a dynamic point of view. This chapter examines the main results on this subject.

This study of network dynamics is a major departure from the 'traditional' social network analysis and may be considered a main advance of contemporary network science. As Moody et al. (2005: 1209) remarks, "much research on social networks is filled with static nouns, such as 'roles', 'relations', 'obligations', and so forth". The structural focus of much social network research has been criticised by some scholars (Emirbayer, 1997; Emirbayer & Goodwin, 1994; Moody et al., 2005) also think this approach limits the understanding of the dynamic nature of social relations. On the other hand, this bias can lead to a risk of overly simplistic theorisation, while the investigation of networked systems could be "far richer and more theoretically nuanced when combined with theoretical insights arising from research on network dynamics" (Hoang & Antoncic, 2003: 181).

Two types of dynamics are discussed in this chapter: one internal, and concerning the time evolution of the number of nodes and links, and one concerning the relationships with the external environment. This latter type regards the overall behaviour of a dynamic process occurring in a system and its reactions to the different possible topologies of the underlying network. The basic models of network growth are discussed first, then the reactions of a networked system to external events are described along with the methods to assess its resilience (capacity to absorb external or internal shocks while maintaining its basic functioning characteristics focusing on knowledge transfer). Thus the following part deals with the diffusion of information and knowledge and the effects of the network's topology on the evolution of this process.

The final section describes the World Wide Web (WWW). The importance of the Internet and related technologies in many fields and in tourism in particular is well known. In this part the most accepted models of the WWW are seen from the *network science* point of view. This will allow comparison of the structural characteristics of the real and virtual networks of the tourism destination under study in order to derive insights into the technological side of the destination activities.

3.1 Basic graph evolution

The classical ER model is a static model with a fixed number of nodes N. A random graph with N nodes and $p(N) \propto N^{z}$ ($-\infty < z < 0$) contains isolated trees and no cycles until $z < -1$ ($\bar{k} = pN \to 0$ for $N \to \infty$). When z passes through -1 ($\bar{k}$ =cost) the asymptotic probability to have cycles, of every order, passes from 0 to 1; for $z \to 0$ the graph contains complete subgraphs of all finite orders.

There is a sudden change in the structure of the clusters when $\bar{k} \to 1$. For $0 < \bar{k} < 1$ the clusters are trees or contain one and only one cycle. The average number is N-L, the largest is a tree whose size is proportional to ln$N$. When $\bar{k}$ assumes the value $\bar{k}_c = 1$, the largest has about $N^{2/3}$ nodes; for $\bar{k} > 1$ it has a finite fraction of nodes and $S = 1-f(\bar{k})$ ($f(x)$ varies exponentially from $f(1)=1$ to 0 for $x \to \infty$). When $p_c \approx 1/N$, the network suddenly changes its topology from a collection of small clusters to a system, in which a *giant component* (cluster) exists. This is defined in different ways in the literature, but the commonly accepted definition is that a giant cluster is formed by a connected subnetwork containing the majority of the nodes so that there is a definite path among all of them.

The transition to a giant cluster happens if and only if:

$$\sum_k k(k-2)P(k) > 0$$

The supercritical phase is characterised by this giant component whose size S grows proportionally to the difference between p and p_c:

$$S \propto (p - p_c).$$

This has a striking analogy with the percolation threshold probability (see section 2.4.1).

A graph can be used as a representation of a complex system (see section 2.2); as such it is possible to study its evolution during time. Let us suppose we have a random network in which, for each time increment, the number of vertices grows by one unit and randomly connects to an 'old' vertex (Dorogovtsev & Mendes, 2002)

Starting with two nodes at time $t = 1$, at time t the network will have $t+1$ nodes and t edges; the total degree is $2t$. The average shortest path length (characteristic distance) is $l \sim \ln t$. If $p(s,k,t)$ is the probability that a vertex s has degree k at time t, the total degree distribution is:

$$P(k,t) = \frac{1}{t+1}\sum_{s=0}^{t} p(s,k,t)$$

The equation, for $t \to \infty$ has the solution:

$$P(k) = 2^{-k}$$

These networks are called *exponential* (in a pure random network the degree distribution $P(k)$ follows a Poisson behaviour).

For $s, t >> 1$ the average degree of a single node is:

$$\bar{k}(s,t) = 1 - \ln(s/t)$$

For a fixed s/t ratio, the degree distribution is:

$$p(s,k,t) = \frac{s}{t}\frac{1}{(k+1)!}\ln^{k+1}\left(\frac{t}{s}\right)$$

The function decreases rapidly for large values of k. Similar results can be obtained if the new node forms a fixed number > 1 of connections with the older nodes.

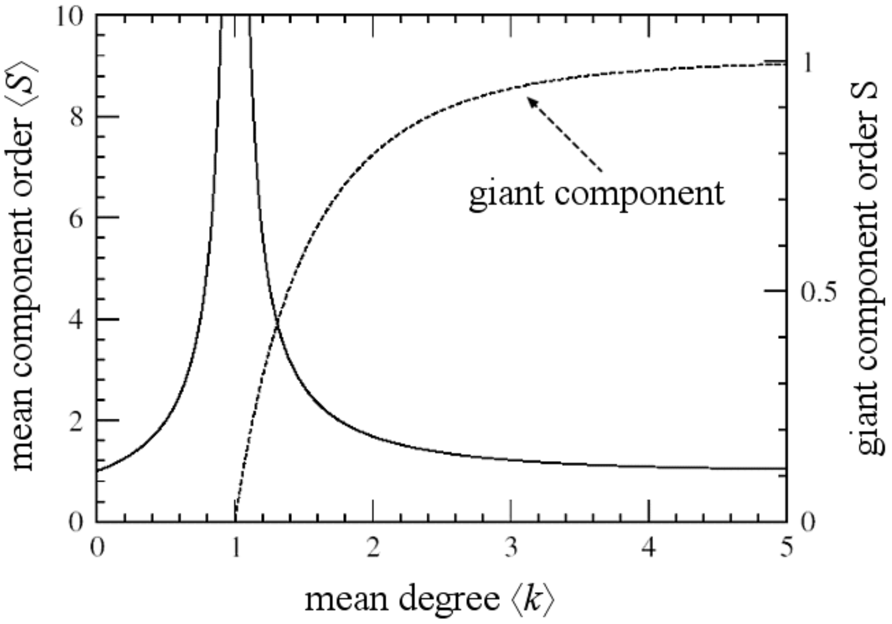


Figure 3.1 Size of the components and of the giant cluster in an ER network (after Newman, 2003b)

In general, during the evolution of a random graph, the following applies:

- if $\bar{k} = pN < 1$, the graph is unconnected and composed of isolated trees; D is the maximum diameter of its subtrees. The critical probability for which almost every graph has a subgraph with z nodes and l edges is: $p_c(N)=cN-z/l$; the one for which a tree of order z exists is $p_c(N)=cN-z/(z-1)$; the one for which a cycle of order z exists is $p_c(N)=cN-1$ and the one for which a complete subgraph of order z exists (with all the vertices connected by the all possible $z(z-1)/2$ edges) is $p_c(N)=cN-2/(z-1)$ (c is a proportionality constant);
- if $\bar{k} = pN > 1$, a giant component appears, if $\bar{k} \geq 3.5$, its diameter is proportional to $ln(N)/ln(\bar{k})$;
- if $\bar{k} = pN \geq ln\ N$, the graph is completely connected and its diameter is close to the value $ln(N)/ln(\bar{k})$.

The average path length l_{rand} depends on N in a similar manner to the diameter:

$$l_{rand} \approx \frac{\ln N}{\ln \bar{k}}$$

A different expression for its value is:

$$l_{rand} = \frac{\ln(N/z_1)}{\ln(z_2/z_1)} + 1$$

where $z_1 = \bar{k}$ and z_2 are the average numbers of first and second neighbours of a vertex.

3.1.1 Dynamic scale-free networks

A scale-free network model is a dynamic one. The network assumes its characteristics through a process of addition of links which is, essentially, governed by a preferential attachment mechanism (see section 2.6). This mechanism can have a number of variations, all giving, as an outcome, a network with a power-law distribution in the degrees, although some deviations may be present.

Finite size effects

The SF model presented in section 2.6, is valid in the *thermodynamic limit* (i.e. for a number of nodes $N \rightarrow \infty$). The finite size of a real network generates deviations from the 'pure' power-law behaviour of the degree distribution.

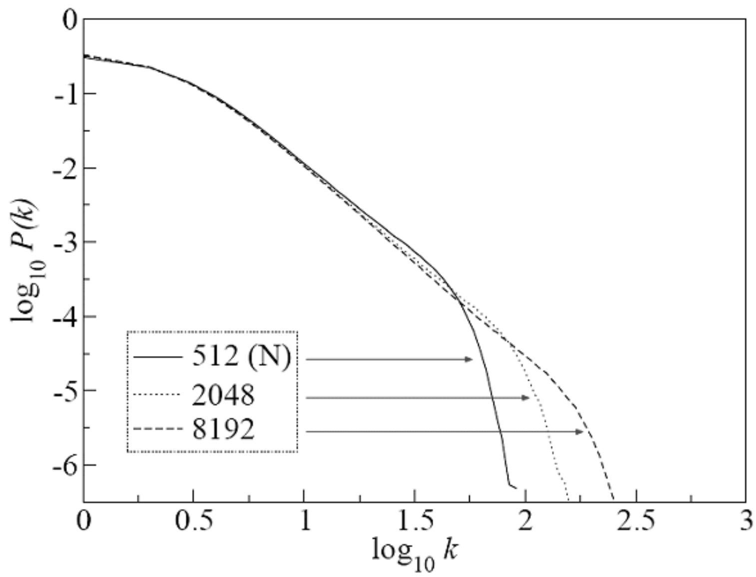


Figure 3.2 Power-law degree distributions for networks with different sizes (adapted from Baiesi & Manna, 2003)

From a formal point of view this can be obtained by thinking of a correction factor to the basic power-law distribution (Baiesi & Manna, 2003), i.e.: $P(k) \sim k^{-\gamma} \cdot f(k/N^{\nu})$, where $f(x)\to 0$ for $N\to\infty$. The deviation shows up as the appearance of curvatures in the shape of the function. An example of what happens in this case is given in Figure 3.2. By increasing the network size, the cut-off becomes less pronounced and smoother.

Non-linear preferential attachments

Preferential attachment rules can be non-linear, i.e. with the form $\Pi(k) \sim k^{\alpha}$. In these cases, typically, the scale-free nature of the network is destroyed. The only case in which the topology is preserved is when the preferential attachment is asymptotically linear: $\Pi(k) \sim \alpha_{\infty} k$ for $k\to\infty$. In this case:

$$P(k) \approx k^{-\gamma} \quad \text{con} \quad \gamma = 1 + \frac{\mu}{\alpha_{\infty}}$$

Any exponent between 2 and ∞ can thus be obtained.

Attractiveness and accelerated growth

A further characteristic of $\Pi(k)$ in many real networks is that $\Pi(0)\neq 0$. It has a non-null probability that a new node connects an isolated one. In general we have:

$$\Pi(k) = A + k^{\alpha}$$

where A is an initial *attractiveness* for a node and represents the probability that an element is 'discovered' and connected by new nodes, even if it does not have any existing connections. The initial form of $\Pi(k)$ would explicitly exclude that. The degree distribution becomes $P(k) \sim k^{-\alpha}$ with $\alpha = 2 + A/m$, (m = number of edges added to each new node). An initial attractiveness maintains the scale-free topology.

If, then, in a directed network at each t a new vertex adds $c_0 t^{\theta}$ links (accelerated growth with a preferential attachment such as $\Pi(k_{in}) \sim A + k_{in}$), the exponent becomes $\alpha = 1 + 1/(1+\theta)$, still preserving the topology.

Age and cost limitations

The basic SF model can be enriched with limitations due to age or cost considerations. Vertices with age or degree higher than a certain value can refuse new connections from the most recently generated nodes. Alternatively, if the network is weighted, nodes whose total weighted degree is greater than a critical value are prevented from adding new connections.

In all these cases, the effect for the degree distribution is the presence of knees or cut-offs that modify the 'linear' (on a log-log plot) behaviour. These effects are more visible if the cumulative degree distribution is plotted (*P(>k)* instead of *P(k)*).

Examples are given in Figure 3.3.

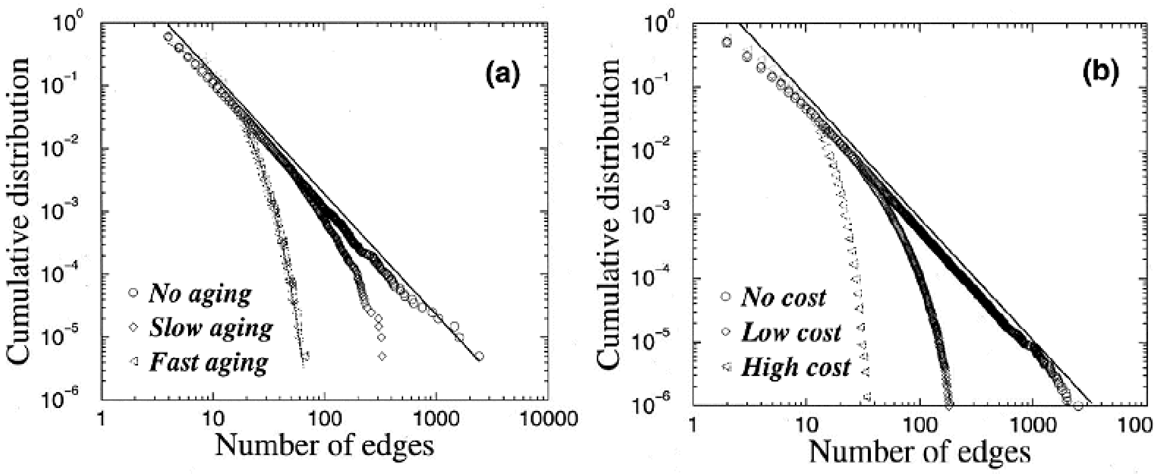


Figure 3.3 Deviation from a pure power-law (solid line) of the degree distribution due to age (a) or cost (b) constraints (after Amaral et al., 2000)

Limited information

One more mechanism able to modify the 'pure' power-law shape of the degree distribution is a limitation in the information a node has on the overall characteristics of the network. In other words a node may establish a connection with only a subset of the other nodes of the network. This fact, as reported by some authors (Mossa et al., 2002; Štefančić & Zlatić, 2005) is responsible for truncations in the power-law behaviours of the degree distributions.

Fitness

The competition characterising many aspects of different kinds of systems can also be traced also in a complex network. It reveals itself as an intrinsic capability of a node to obtain more connections at the expense of others. To realise that, it is possible to introduce a *fitness* parameter η, with distribution $\rho(\eta)$, which modifies the preferential attachment expression as follows:

$$\Pi_i = \frac{\eta_i k_i}{\sum_i \eta_i k_i}$$

A new node will have a probability to create connections that depends on its fitness. This fact can explain the reason why, in many situations, recent elements can achieve a greater success by increasing their connections at a higher speed than older nodes.

The time evolution of k is:

$$k_{\eta_i}(t, t_i) = m\left(\frac{t}{t_i}\right)^{\beta(\eta_i)}$$

where $\beta(\eta) = \eta / C$ and $C \sim \int \rho(\eta) d\eta$.

For a uniform $\rho(\eta)$ distribution $C=1.255$ and the degree distribution becomes:

$$P(k) \approx \frac{k^{-C-1}}{\ln k}$$

This situation leads to a very interesting conclusion. Bianconi and Barabási (2001) have shown that there is remarkable analogy between this phenomenon and those particle systems, known in quantum mechanics, that satisfy the Bose-Einstein statistics.

Let us assign each node in a network a quantity: $\varepsilon_i = -(1/\beta) \ln \eta_i$. It may be interpreted as an energy and $1/\beta=T$ as a temperature.

A link between two vertices i and j with energies ε_i and ε_j corresponds to two interacting particles, each at its own energy level. The addition of a new element means adding a new energy level l and $2m$ new particles to the system. One half of these will be at energy l, the others will be distributed on the energy levels of their correspondents. The probability that a particle has energy at level i can be written as follows (see for example Landau & Lifshitz, 1980):

$$\Pi_i = \frac{e^{-\beta\varepsilon_i} k_i}{\sum e^{-\beta\varepsilon_i} k_i}$$

Following the time evolution, for $t \to \infty$ the number of particles with a given energy is:

$$n(\varepsilon) = \frac{1}{e^{\beta(\varepsilon-\mu)} - 1}$$

which is the expression for the Bose-Einstein statistics. The occupation of lower energy levels corresponds to the number of links possessed by the nodes with higher fitness.

The topology of such a network is affected by the energy (fitness) distribution $p(\varepsilon)$. It has been shown (Bianconi & Barabási, 2001) that for a distribution such that $p(\varepsilon) \to 0$ as $\varepsilon \to 0$, there is a critical temperature T_{BE} at which there is a topological transition. If $T < T_{BE}$, the node with highest fitness acquires a finite fraction of links and keeps it indefinitely. This corresponds, in quantum mechanics, to the Bose-Einstein condensation. If the temperature $T > T_{BE}$, fitter nodes will get connections faster, but, by increasing the network size, the proportion of connections to vertices with higher degrees decreases with time. This phase is called *fit-gets-rich* (FGR, see Figure 3.4).

Figure 3.5 shows the behaviour of maximum degree in a system with energy distribution $\rho(\varepsilon) = C\varepsilon^{\theta}$ (with $\varepsilon \in (0,1)$; $m=2$ and $\theta=1$).

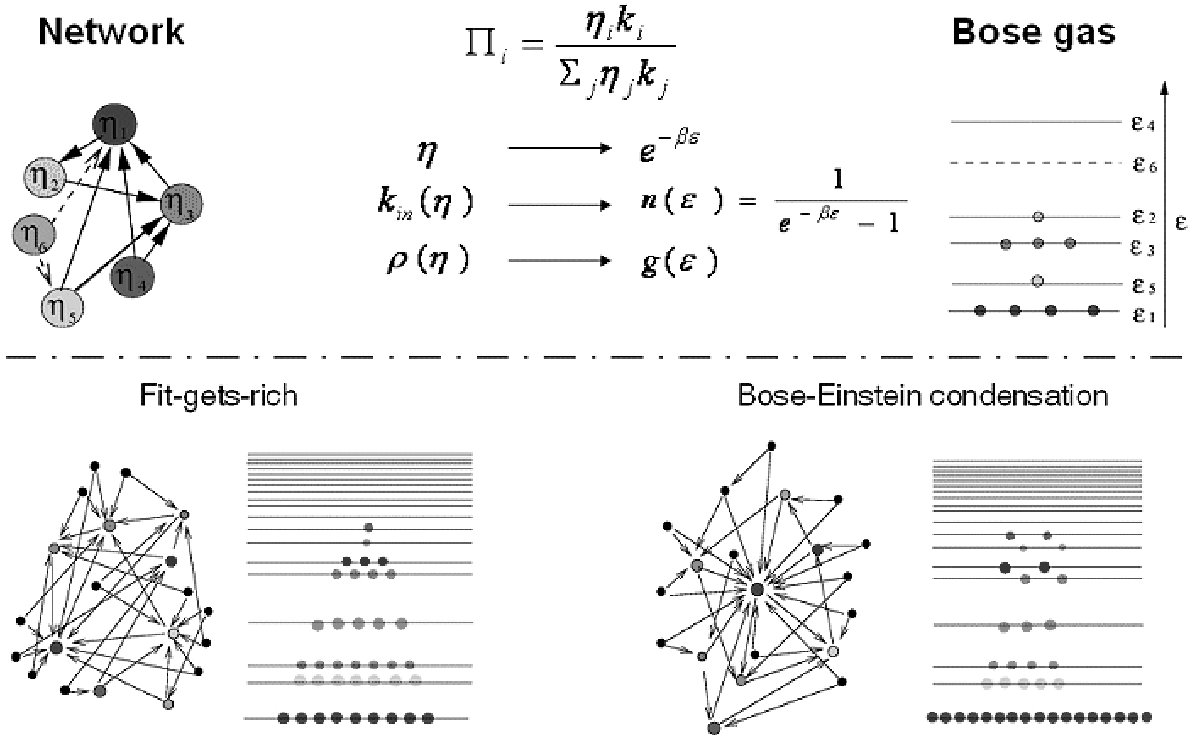


Figure 3.4 Fit-gets-rich phenomenon and Bose-Einstein condensation analogy (after Bianconi & Barabási, 2001)

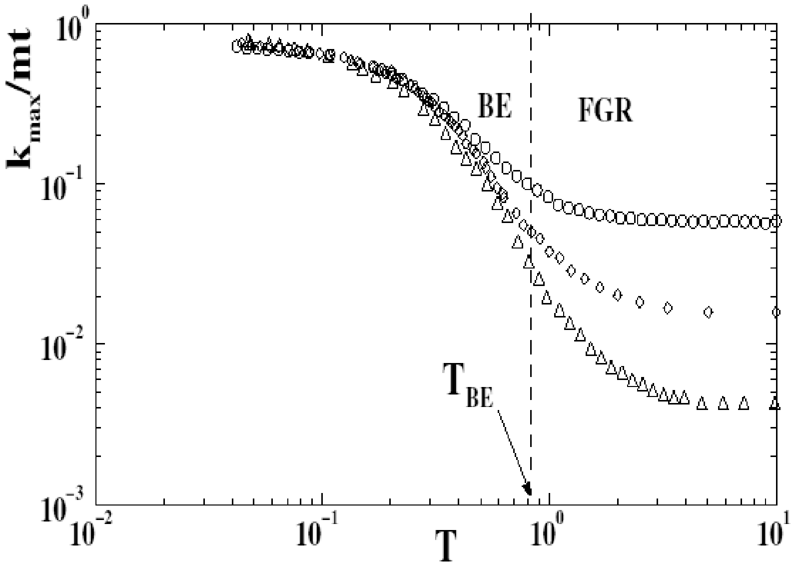


Figure 3.5 Maximum degree in a SF network as function of temperature T (after Bianconi & Barabási, 2001)

The evolution of many complex systems such as the Web, the industry and the economy, follows the dynamic laws that describe the interactions among the parts of the system. In many cases, in spite of their non-equilibrium and irreversibility characteristics, they are ruled by Bose-Einstein statistics and may undergo a condensation. This quantum mechanical analogy is able to explain famous mechanisms known as first-mover advantage, rich-gets-richer, winners-take-all, in terms of distinct thermodynamic phases of a complex network.

3.2 Dynamic processes on complex networks

Complex systems and their associated networks are dynamic entities and, more importantly, are subject to a number of different dynamic processes in which, as it has been discovered in the last years, their topology plays a crucial role. In the following sections some of these processes will be discussed.

3.2.1 Robustness and fragility

One of the most important qualities of a complex system is its reliability. For a network, this depends on its topological characteristics such as connectivity, nodal degrees or average distance. A high degree of tolerance to errors is a feature existing in many different systems. For example, relatively simple biological organisms are able to live and proliferate in spite of difficult environmental conditions in which they may survive, even in presence of external potentially destructive shocks. Complex communication networks (the Internet, for example) exhibit high robustness. Single elements may have malfunctions or may stop working, but the whole network does not lose its capacity to transfer messages or to allow communication between two nodes.

The stability of these systems must be attributed, at least partially, to their topology, independently from the characteristics of single nodes or links. In other words, the robustness of a network depends on the type of connectivity it has.

The stability can be studied by simulating the removal of a certain number of nodes or links. An error tolerant or robust network is a network that still contains a giant connected component after the elimination of a fraction of elements.

Experimental results (Albert et al., 2000; Crucitti et al., 2003) show that there is a close relationship between topology and robustness. In particular, scale-free networks are stronger than ER networks in the case of a random removal of vertices, but they are more vulnerable to targeted attacks to their hubs.

In a simulation, the robustness of a graph is measured by randomly removing a certain number of nodes. A targeted attack is simulated by deleting a certain number of vertices chosen among those with the highest degrees.

Let us consider a random connected graph with N nodes and let us remove a fraction f of them (a node removal also implies the deletion of all the edges attached to that node). There exists a threshold value $f_c(N)$ such that if $f < f_c(N)$ the resulting subgraph stays connected. If $f > f_c(N)$ the graph is disconnected (see also section 2.4). Actually, this is an inverse percolation problem, with the difference that the critical threshold depends on the number of vertices N.

The quantities measured are:

- average path length l between two randomly chosen nodes;
- relative size S of the largest connected component (giant cluster, if it exists);
- average size $\bar{s}$ of the connected components (excluding the giant cluster).

The difference in behaviour between a random network (exponential or ER) and an SF network is remarkable (Figure 3.6).

In an ER network, there is no significant difference between the two deletion methods. The dependency on f (the fraction of nodes deleted) is the same in both cases. A SF network (having a degree distribution with exponent $\gamma \leq 3$) yields, instead, to completely different results. For a targeted attack, the SF network has a behaviour which is qualitatively similar to the one of an ER network: there is a critical f_c beyond which the network is destroyed. On the other hand, when the node deletion is random, the SF network has a substantial lack of critical threshold ($f_c \rightarrow 1$).

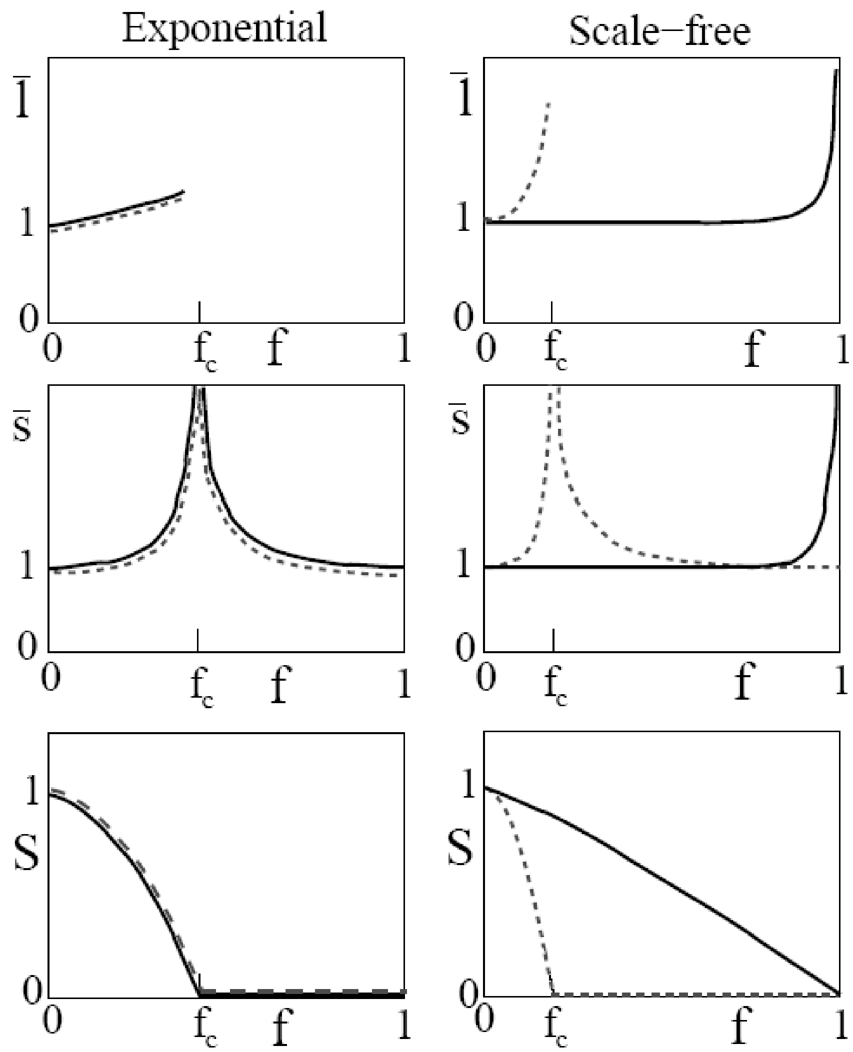


Figure 3.6 Simulation of the effects of malfunctions (random removals, solid line) and targeted attacks (dotted line) for random (exponential and ER) networks and scale-free networks (after Dorogovtsev & Mendes, 2002)

The same behaviour can be observed (Figure 3.7) also for local and global efficiency (Crucitti et al., 2003). It is worth remembering here that there is a close correspondence between the average distance, the clustering coefficient and the two efficiencies: more precisely $L \rightarrow E_{glob}$ and $C \rightarrow E_{loc}$ (see section 2.5.1).

Scale-free networks are robust (have high resilience) to random removals of elements while they show high fragility to attacks to the components with the highest connections.

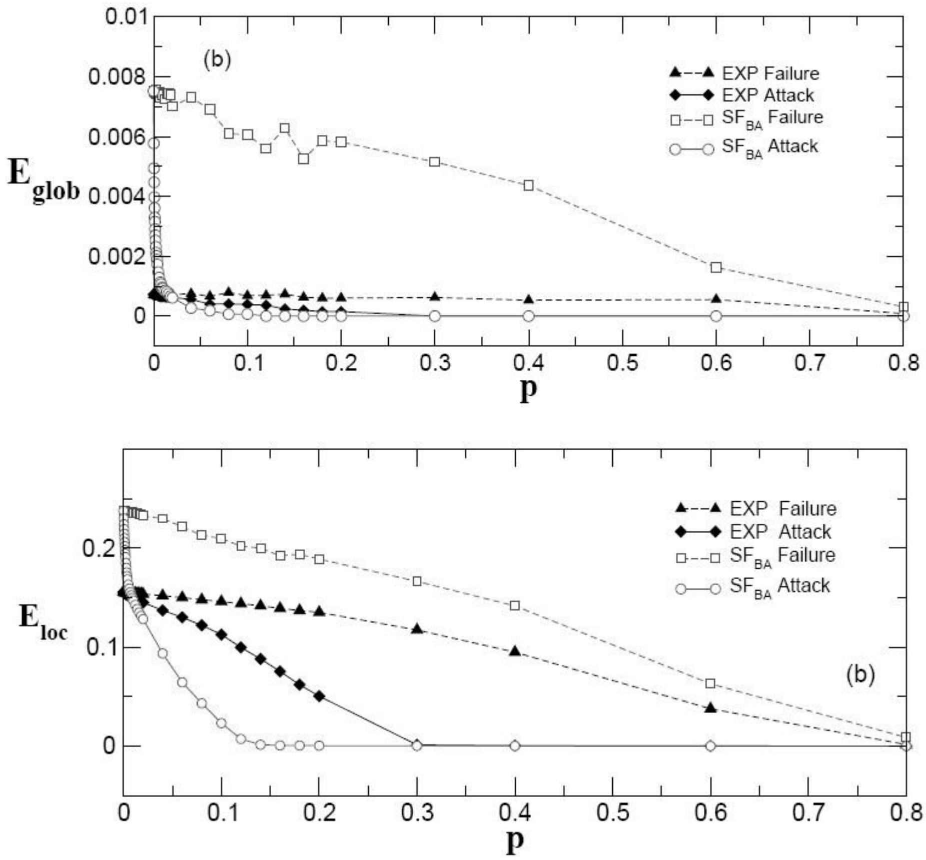


Figure 3.7 Local and global efficiency as a function of random and targeted removal of a fraction *p* of nodes in different networks (after Crucitti et al., 2003)

3.2.2 Diffusion in complex networks

Diffusion processes are of great importance. Diseases, viruses, cascading failures, but also information, messages and ideas are subject to phenomena that can be modelled in similar ways. Recently, then, it has been understood that the shape of the system serving as medium for the diffusion may have a crucial influence.

The process is, in general, bound to the percolation properties of the network studied. One of the most used diffusion models, derived from clinical epidemiology, it the so called SIS (susceptible-infected-susceptible) model, valid to describe a vast class of diffusive phenomena (Anderson & May, 1992; Diekmann & Heesterbeek, 2000).

The model starts by supposing that a node of a network is susceptible to an infection and that it is infected with rate ν when it has at least an infected neighbour. Infected elements are cured with rate δ and turn susceptible again.

The effective infection rate $\lambda \equiv \nu/\delta$ is the main parameter of the model.

Let us take into account undirected equilibrium networks, i.e. those with a definite and fixed number of nodes and edges (static and stable networks).

In exponential homogeneous networks (with a Poissonian degree distribution such as the ER and SW networks), the situation is similar to that of a disease that spreads out: it exists as a non-null threshold $\lambda_c \sim \bar{k}$ ($\bar{k}$ = mean degree) below which the phenomenon extinguishes very rapidly (exponentially). In other words, when a node is contaminated, the average density of infected nodes (prevalence) $\rho(t)$ goes to 0 very quickly (Figure 3.8).

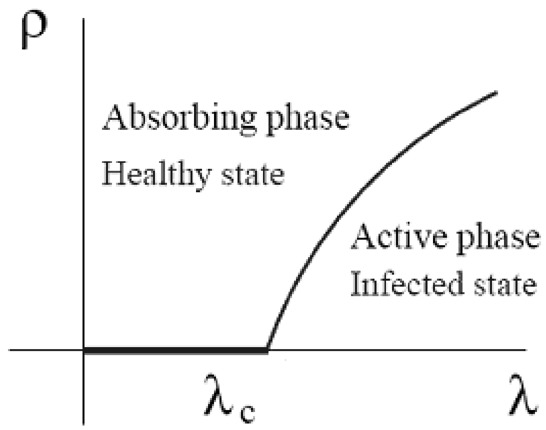


Figure 3.8 Infection model and critical threshold (after Pastor-Satorras & Vespignani, 2003)

When $\lambda > \lambda_c$ the infection spreads and becomes endemic:

$$\rho(t \to \infty) \equiv \rho \propto (\lambda - \lambda_c)$$

In homogeneous networks the threshold is:

$$\lambda_c = \frac{\bar{k}}{k^2}$$

If this value is small, it coincides with the critical percolation value for randomly damaged networks: $\lambda_c \cong p_c$. If κ_1 is the largest eigenvalue of the adjacency matrix, the critical threshold is also $\lambda_c = 1/\kappa_1$.

If an infinite network has a power-law degree distribution (SF network) with exponent $\gamma \leq 3$, the infection is endemic for every $\lambda > 0$ (Pastor-Satorras & Vespignani, 2001, 2002, 2003). Hence, there is no critical threshold. Such networks, so resilient to random damages, are also incredibly sensitive to epidemic diffusion. Both phenomena, in fact, depend on the shape of the degree distribution tail.

The limiting effects of a finite number (even if large) of elements bring us to the condition depicted in Figure 3.9.

If $\gamma = 3$, in an *infinite* network, the prevalence of the endemic state is (B = constant):

$$\rho \sim e^{(-B/\lambda)}$$

if $\gamma > 3$, but close enough to 3:

$$\rho \sim (\lambda - \lambda_c)^{1/(\gamma-3)}$$

Only when $\gamma > 3$ do we go back to a situation in which $\rho \propto (\lambda - \lambda_c)$.

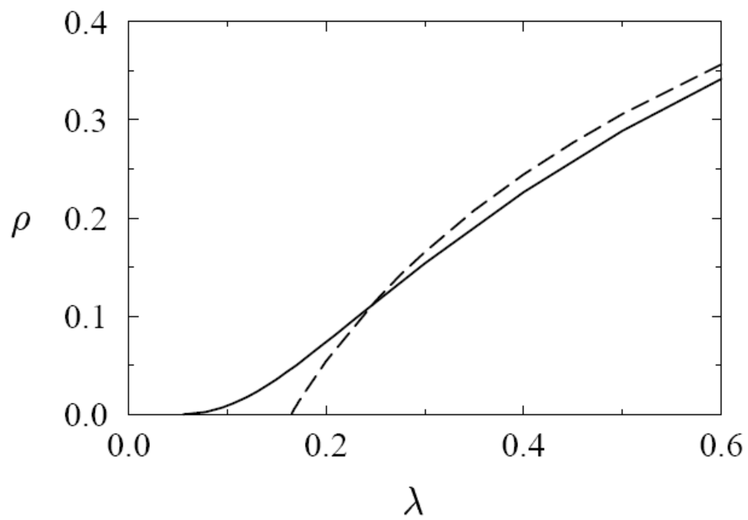


Figure 3.9 Diffusion (infection) in SF networks (solid line) compared to homogeneous random networks (ER, dotted line) (after Pastor-Satorras & Vespignani, 2003)

In *finite* networks, the size effect leads to the presence of an effective critical threshold which depends on the network size (number of nodes N): $\lambda_c(N) \simeq p_c(N)$. For $\gamma \leq 3$ the threshold is large, and it is observable only for values close to 2.

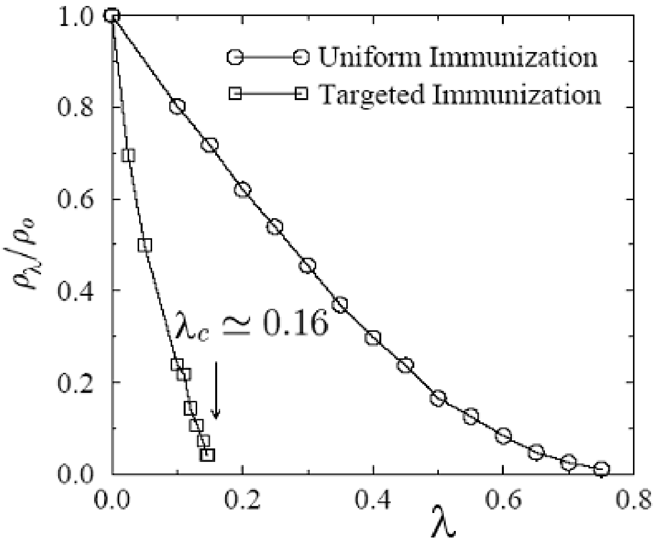


Figure 3.10 Immunisation techniques for SF networks at a fixed spreading rate (after Pastor-Satorras & Vespignani, 2002)

Therefore, real SF networks are subject to diffusion and 'virus' prevalence (for example those affecting Internet computers). On the other hand, their robustness renders ineffective all disinfection attempts based on 'random' actions, i.e. attempting to 'vaccinate' large numbers of nodes taking them more or less at random. The only effective step would be the identification of the most connected vertices and their disinfection to interrupt the diffusion (Figure 3.10).

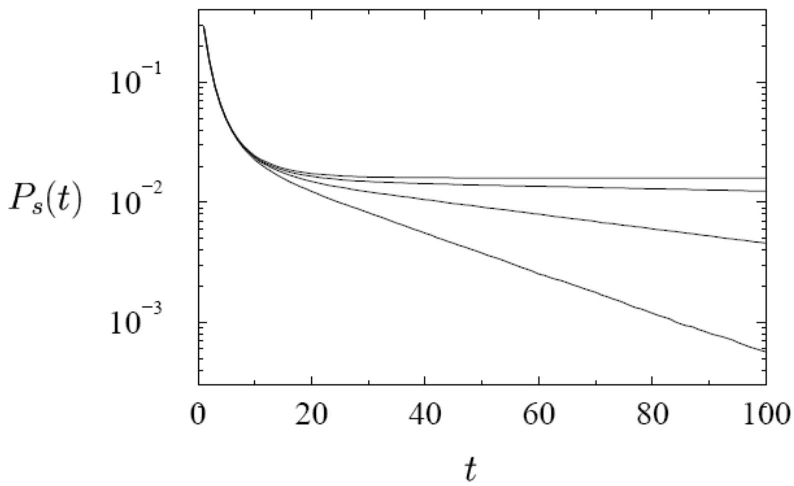


Figure 3.11 Survival probability of an infection for λ=0.065 and N from ~10^3 to ~10^6 (top to bottom) (after Pastor-Satorras & Vespignani, 2003)

If an infection has started in a SF network, its survival probability depends on the network size (Figure 3.11).

3.2.3 Self-organised criticality

A power law seems to be a peculiar attribute for complex systems 'far from equilibrium' and with non-linear characteristics. In general, we think of far from equilibrium systems as those evolving towards a critical state or phase only if the systems parameters are varied from the external.

The Self-Organised Criticality (SOC) theory, instead, postulates that dynamic systems evolve spontaneously to a critical state for a large range of system characteristic parameters. The classical example is the one of a sand pile on which new grains are dropped. As long as the slope is modest, the addition of new grains does not have macroscopic effects - only a few of them will change position. When the pile grows, the slope will reach a critical value. At this point, new grains will produce avalanches of different sizes. After that, the pile self-organises with a new slope.

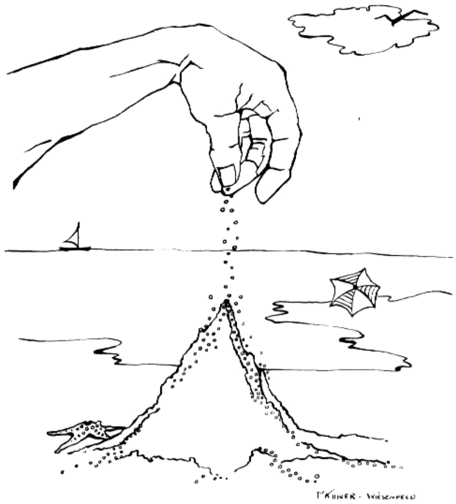


Figure 3.12 The sandpile giving the name to self-organised criticality model (drawing by E. Wiesenfeld in Bak, 1996: 2)

The system goes to a critical state spontaneously, without any external intervention (besides the dropping of grains), or without the necessity (typical of other models, phase change of water, for example) of an external regulation of system's parameters. More importantly, this critical state shows great robustness to initial state perturbations. This robustness renders the SOC a plausible model for the instability of many natural systems.

The *sandpile* model, due to Bak, Tang and Wiesenfeld (1987; 1988) plays an important role in statistical mechanics and its self-organisation property seems to be the basis for the presence of fractals and background noise (the *1/f* noise) in many natural systems (such as the flicker noise in electronic circuits)

Consider a square regular lattice with side L (it will have $L \times L$ nodes) to which a rewiring process is applied (the network will become small-world), then:

1. each node i is assigned a threshold z_i ($\leq k_i$), for example: $z_i = k_i^{1-\eta}$ $(0 \leq \eta < 1)$. Let $\lceil z_i \rceil$ be the smallest integer $\geq z_i$ ($\lceil z_i \rceil \leq k_i$);
2. for every time interval a 'sand grain' is added to a node i chosen at random. The node *height* (integer) h_i increases by 1;
3. if h_i reaches or exceeds z_i the node becomes instable and the $\lceil z_i \rceil$ grains fall (spill) on $\lceil z_i \rceil$ neighbour nodes chosen randomly among the k_i, hence: $h_i \rightarrow h_i - \lceil z_i \rceil$ and $h_j = h_j + 1$ for all other j nodes;
4. if this causes the instability of any adjacent nodes the process is repeated until there are no more instable nodes. This process is called an avalanche;
5. steps 2÷4 are repeated.

Now we would like to know the:

- area of the avalanche: the number of distinct nodes involved;
- size of the avalanche: the number of 'spill' events in a specific avalanche; and
- duration of the avalanche.

Numerical simulations show that a system in stationary state, reached without external modifications of any parameter, has scale invariance characteristics in the distribution of avalanche sizes *n(s)* and their durations *n(t)*:

$$n(s) \sim s^{-\tau} \text{ e } n(t) \sim t^{-\delta}$$

More precisely, the size distribution shows an exponential cut-off (*sc* is the characteristic size):

$$n(s) \sim s^{-\tau} e^{(-s/sc)}$$

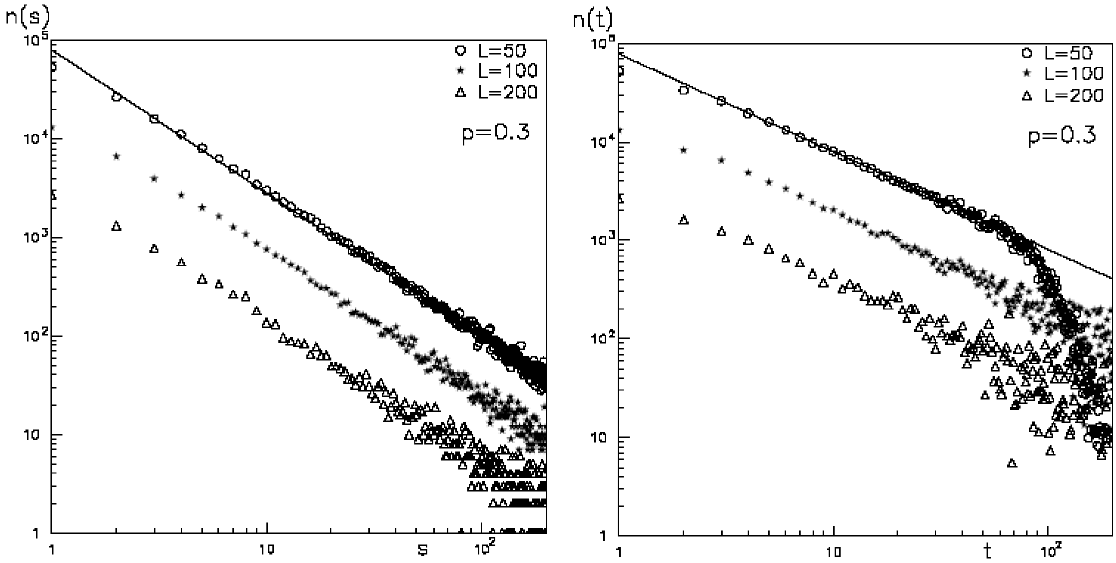


Figure 3.13 Size and lifetime distributions for avalanches in small-world networks different sizes (L) and rewiring probability p=0.3 (after de Arcangelis & Herrmann, 2002)

Size and lifetime distributions of the avalanches follow a power law (Figure 3.13).

The origin of great avalanches that start from an even small perturbation is a phenomenon which occurs in many areas of the economy and society. Some of them, such as stock exchange crashes, electric blackouts, but also diffusion of ideas, innovations, fashion and obsessions, may also be rare, but their size and effects are quite big.

An example can be the diffusion of innovations (Guardiola et al., 2002; López-Pintado, 2004; Rogers et al., 2005). A diffusion cycle starts when an individual (a node in a social network) becomes spontaneously active by adopting a new product or a new technology. If the innovators are strongly connected to a community of vulnerable individuals, the percolation process fires up, otherwise the innovation extinguishes after having 'contaminated' only a small number of elements.

In the study of these events, we are interested in finding the probability that an avalanche is initiated by a single (or a very small number) element and the expected size.

As seen, the condition which produces a large avalanche is the overload of a node (or of a link between two nodes) propagating along the network. The capacity of this node is the maximum load it is able to bear. In man made networks this capacity can be strongly influenced by cost considerations.

Suppose that the capacity C_i of node i is proportional to its initial load L_i:

$$C_i = (1 + \alpha)\, L_i \text{ for } i = 1, 2, ..., N$$

where the constant $\alpha \geq 0$ is a tolerance parameter and N the initial number of nodes in the network. If the network is active, it is in a free flux state as long as $\alpha \geq 0$. The removal of nodes changes the paths distribution and the load of a node may change. If this increases beyond the node capacity, the node fails and its load is distributed among the others. As a result, some of them may be loaded beyond their critical capacity and an avalanche starts. The subsequent behaviour is determined by the topology of the network.

Let us consider, for example, a scale-free network, with a degree distribution:

$$P(k) \sim k^{-\gamma}$$

The damage can be quantified in terms of the relative size G of the largest connected component:

$$G = N'/N$$

where N and N' are the number of nodes of the largest component before and after the avalanche.

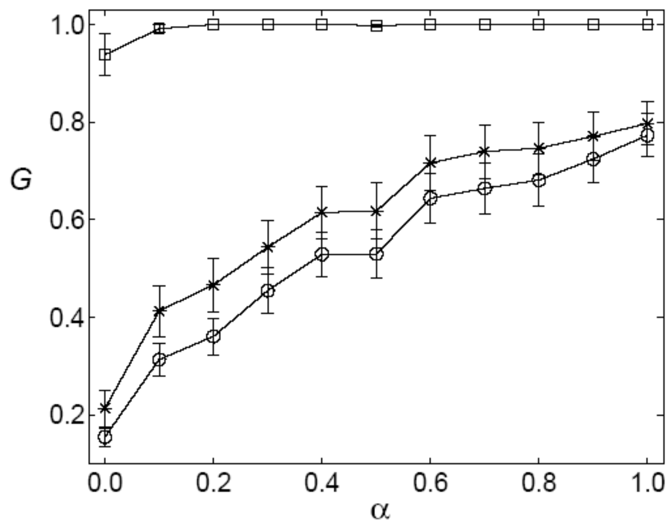


Figure 3.14 Largest component G as function of α in a SF network (squares: random node removal, asterisks: removal of highly connected nodes, circles: removal of high load nodes) (after Motter & Lai, 2002)

Figure 3.14 shows the effect on *G* of a node removal done randomly or in a targeted way (nodes with higher load or connectivity) in scale-free network with $\gamma = 3$.

G remains close to *1* for random overloads, it decreases significantly for targeted removals. For example, for $\alpha = 1$ (capacity double of the one at which the system operates in normal conditions) *G* is reduced to 20%. The relative damage is higher for small α values

Figure 3.15 gives the results for a random homogeneous network (ER). In this case there are no avalanches for $\alpha \leq 0.05$ which acts as critical value for the process.

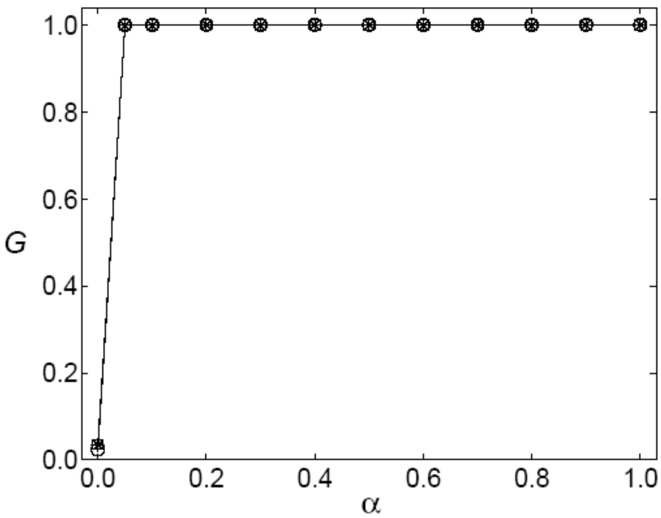


Figure 3.15 G (α) for an ER network (after Motter & Lai, 2002)

The relation between connectivity and resilience suggests the hypothesis that a global avalanche can be initiated only for a window of possible values (Figure 3.16). The conditions can be summarised by identifying the following characteristics (Crucitti et al., 2004a; Motter & Lai, 2002; Watts, 2002):

- *node threshold*: if it is too high, no node will ever become active and the system will stay stable, independently from the type of connectivity;
- *network sparseness*: if the network is sparse (low number of connections), the individual thresholds are sparse as well. Only if the density increases does the percolating cluster reach a sensible size. Obviously, if the network is sparse, highly connected hubs have a disproportionate efficacy in favouring the diffusion;

- *network density*: an upper limit suggests that even densely connected networks can be refractory to the phenomenon. For example, in a social network, if any individual cares of many others, nobody will start the process.

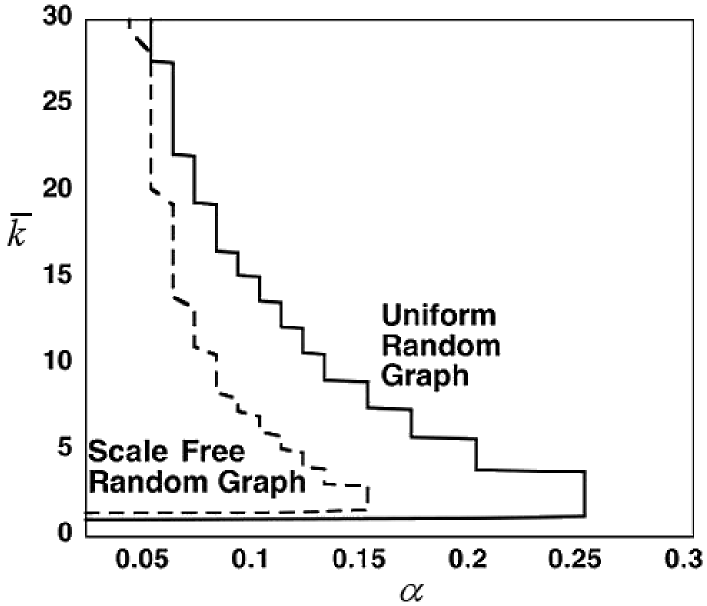


Figure 3.16 The window of possibilities for avalanches as function of the average network degree and the tolerance parameter. The area included in the two curves is the one of the values for which an avalanche is possible in SF and ER networks (after Watts, 2002)

In conclusion, a system will be subject to global avalanches only if a balance between local stability and global connectivity exist. In other words, the phenomenon depends only on the structural characteristics of the network and the relationship with the external environment only determines whether a triggering event exists.

3.2.4 Network optimisation

Watts and Strogatz (1998) have shown how empirically derived small-world networks have intermediate characteristics between regular lattices and completely random networks. The presence of such networks in the real world leads to the idea that self-organisation mechanisms may guide an evolution towards networks with SW characteristics as a balance in the optimisation of the average separation between nodes and in the total cost of connection. A competitive optimisation principle must, therefore, be in charge of the formation of a SW network (Mathias & Gopal, 2001).

On the other hand, the existence of hubs (and therefore of scale invariant phenomena) is a structural component of a SW network and increases the degree of clustering.

It can be seen that, by optimising a network, it is possible to obtain values $C \geq C_{regular}$ and $L \leq L_{random}$.

The density of a network is defined as:

$$\rho = \frac{\bar{k}}{N-1}$$

In real networks, typical values are (see for example Albert & Barabási, 2002; Dorogovtsev & Mendes, 2002): $\rho \in \{10^{-5}, 10^{-1}\}$.

We can define an *energy* function:

$$E(\lambda) = \lambda d + (1-\lambda)\rho$$

where d is the distance between two nodes and λ a parameter that controls the linear combination of d and ρ.

The normalised number of links ρ is (a_{ij} is an element of the adjacency matrix):

$$\rho = \frac{1}{\binom{n}{2}} \sum_{i<j} a_{ij}$$

the normalised distance d is defined as: $d = D/D_{max}$ where D (minimum average distance between vertices) is:

$$D = \frac{1}{\binom{n}{2}} \sum_{i<j} D_{ij}$$

$D_{max} = (N+1)/3$ for a connected network.

Minimising $E(\lambda)$ means minimising distance and number of nodes and imposing that the network be connected (Ferrer i Cancho & Solé, 2003).

The probabilistic algorithm used for the optimisation is known as *simulated annealing*. Originally proposed by Kirkpatrick (1983; 1984), it is named after the annealing of a metal which is first heated to very high temperature and then slowly and gradually cooled as far as a stable, crystalline structure is reached. At high temperatures, the atoms of the metal are in a greatly disordered state, therefore the total energy of the system is elevated. To bring the atoms to a crystalline arrangement, the temperature must be lowered very slowly; otherwise small imperfections in the lattice may be generated with a consequent metastability. At high temperature all the energy states are probabilistically possible, while at low temperatures the system is, for sure, in states with minimum energy. The heating is also used, in physical processes, to settle the atoms of a crystalline lattice after having performed operations that may have modified the geometric symmetry of the crystal.

The method uses the algorithm first proposed by Metropolis (1953). It consists in providing each element of the system a certain (random) amount of energy able to move it, with a certain probability, from its partial equilibrium state to a better one. It can be seen as a way to overcome a local maximum of the *cost* or *energy* function to reach a lower minimum. The initial energy is gradually reduced during the process.

Let ΔE be the energy difference between a configuration and the previous one, if the value is negative (new state is better) the new configuration is accepted. Otherwise the probability that the new state is accepted is distributed as a certain function, for example as (Boltzmann distribution):

$$f(\Delta E; T) = e^{-\Delta E/T}$$

T, for analogy with thermodynamics, is called temperature. Its initial value T_0 is chosen high enough so that the highest number of states can be accepted. Usually, the choice is for a value T_0 much higher than the standard deviation of the cost distribution (supposed normal) as in:

$$T_0 = -3\sigma_c / lnP$$

where P is the probability with which the state should be accepted. At step n, the value of T is determined such that:

$$T_n = \alpha\, T_{n-1}$$

with $0.8 < \alpha < 0.95$. This way the temperature goes down slowly with the subsequent iterations.

A 'well ordered' network, in a minimum energy state (ideally null) will also have a very low entropy (ideally the minimum possible).

If p_k is the frequency of nodes with degree k ($\Sigma p_k = 1$), given λ, the Shannon entropy H can be defined as:

$$H = -\sum_{k=1}^{N-1} p_k \log p_k$$

If $\lambda=0$ the network is random (ER), if $\lambda=1$it is a clique ($\rho(\lambda)=1$).

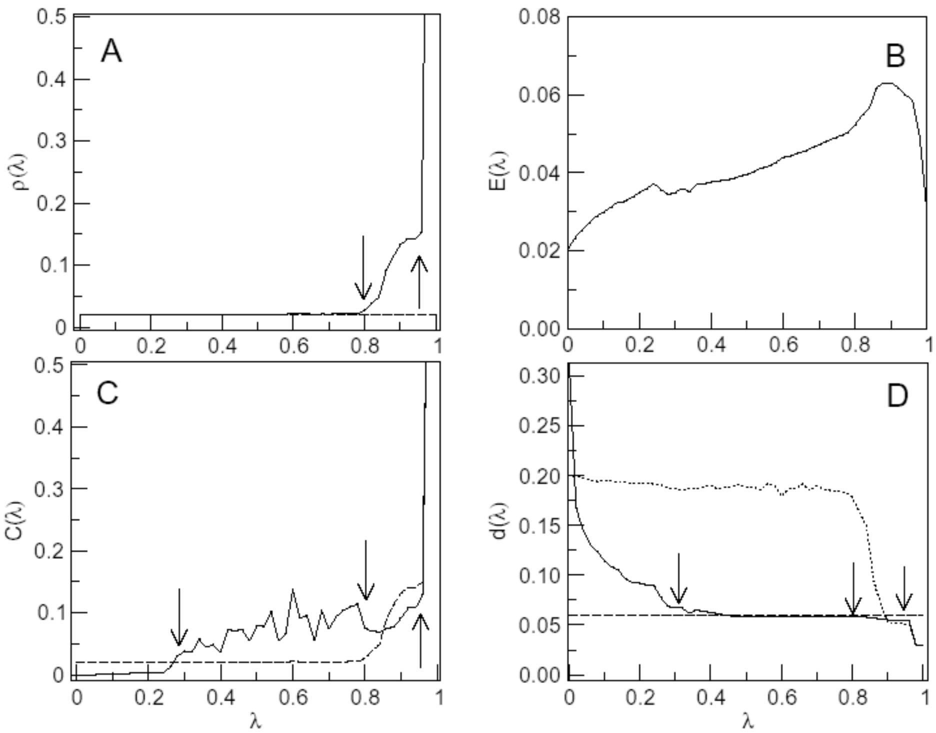


Figure 3.17 Density (A), energy (B), clustering coefficient (C) and distance (D) as function of λ. (C shows the clustering coefficient for an ER network, D the normalised distance for a star

network d=6/(N+1) and for a poissonian network d=logN/log⟨k⟩) (after Ferrer i Cancho & Solé, 2003)

By looking at Figure 3.17, it can be noted that three sharp transitions exist (marked with arrows) for $\lambda^*_1 \approx 0.25$, $\lambda^*_2 \approx 0.80$ and $\lambda^*_3 \approx 0.95$ (see also Figure 3.18):

- exponential networks with $P(k) \sim e^{-k/\xi}$;
- scale-free with exponential cut-off with $P(k) \sim k^{-\gamma} e^{-k/\xi}$ ($\gamma \approx 3$; $\xi \approx 20$ and $N=100$);
- star networks ($\lambda^*_1 < \lambda < \lambda^*_2$) with a central node connecting all others: $p_k = \frac{N-1}{N}\delta_{k,1} + \frac{1}{N}\delta_{k,N-1}$. A star network has, rather obviously, the lowest distance among vertices all over the graph, with the least number of links.

Scale-free networks are in the neighbourhood of $\lambda^*_1 \approx 0.25$; the cumulative degree distribution exponent is $\gamma \approx 2.0$ (the same calculated by Barabási and Albert). Preferential connections are a necessary condition.

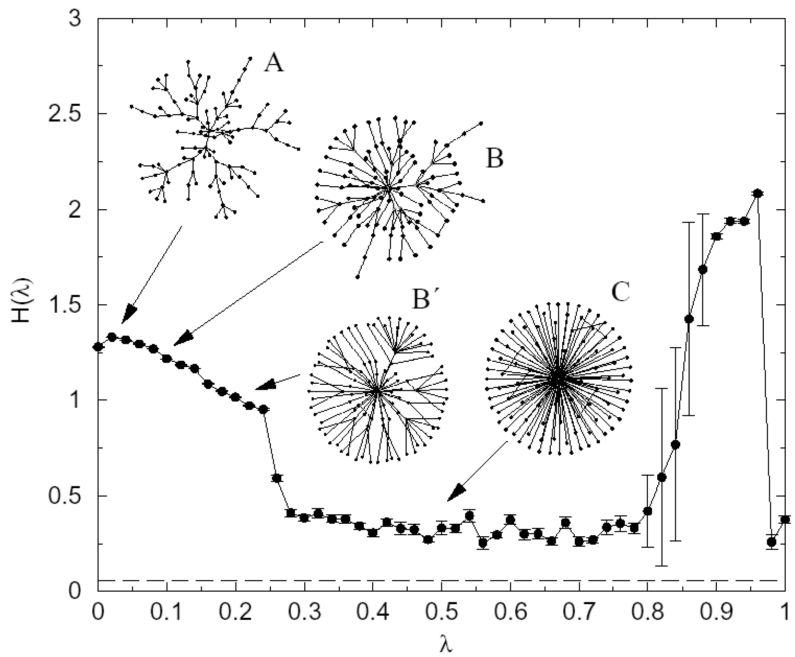


Figure 3.18 Entropy H as function of λ. Different network topologies are shown (after Ferrer i Cancho & Solé, 2003)

The edges can be weighted, for example, by assigning a cost. In this case, the total cost function is (l = link length, a = cost):

$$L(N) = \sum_{i>j} a_{ij} l_{ij}$$

If the link weight corresponds to its physical length, the distribution decays exponentially: $D(l) \sim \exp(-l/l_0)$.

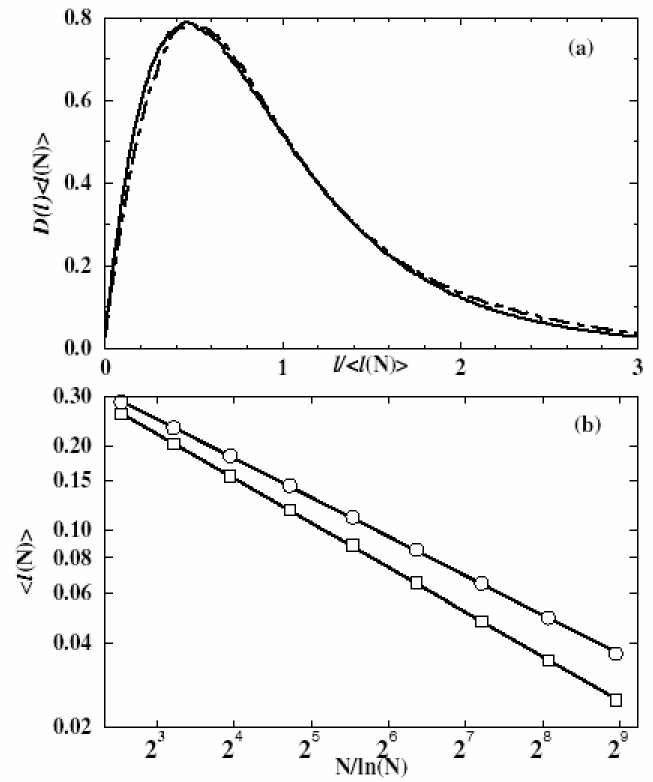


Figure 3.19 Link length distribution for optimised networks (after Manna & Kabakçıoğlu, 2003)

The distributions of link lengths after an optimisation process are depicted in Figure 3.19. $D(l)$ is normalised for average length $\langle l(N)\rangle$ which varies according to $\langle N/logN\rangle^{\mu}$.

In conclusion, networks with optimised 'costs' have the characteristics of many real networks: they are small-world, scale-free and exhibit high clustering levels.

3.2.5 Communities and community discovery

Complex networks can have modular structures. We have examined, in the past sections, possible mechanisms to build up networks so that the main characteristics are in agreement with those found in real cases.

In analysing a network, on the other hand, one of the tasks that can be accomplished is the detection of modular structures that commonly take the name (derived from the social sciences field) of *communities*. The search is important, for example, to reveal social structure through their communication patterns, or to study functional connections in metabolic and protein networks, or for the optimisation of technological systems (see for example Arenas et al., 2004; Danon et al., 2005; Girvan & Newman, 2002; Newman, 2004).

From a technical point of view, the problem is one of dividing the vertices of a network into a number of dense subsets (sometimes they are needed to be of roughly equal size), while minimising the number of connections between vertices belonging to different groups. The *partitioning* of a graph is, typically, a highly computationally intensive problem for which several algorithms have been proposed. The situation, here, is complicated by the fact that, often, there is no prior knowledge of the number of communities to be discovered. Moreover, communities can have a hierarchical structure and be, in turn, further divided into subcommunities.

There is no clear and common definition of a community (Danon et al., 2005), but a general characterisation can be of subsets of nodes which are more densely linked, when compared to the rest of the network. To measure this effect, a widely accepted quantity is the modularity index proposed by Newman and Girvan (2004).

It is defined as follows:

$$Q = \sum_i (e_{ii} - a_i)^2$$

where e_{ii} is the fraction of edges in the network between any two vertices in the subgroup i, and a_i the total fraction of edges with one vertex in the group; edges 'internal' to each group have weight 1 while inter-group links are weighted 1/2. In other words, Q is the fraction of all edges that lie within communities minus the expected value of the same quantity in a graph in which the vertices have the same degrees but edges are placed at random.

In a pure random network (ER), as shown by Guimerà (2003), Q can have values much higher than expected. This suggests that these networks, that should not have a modular structure, actually seem to have one due to fluctuations.

The community identification methods proposed can be basically divided into two large classes. The first one (Arenas et al., 2004; Danon et al., 2005; Girvan & Newman, 2002; Newman, 2004; Newman & Girvan, 2004; Radicchi et al., 2004) is based on an iterative division (with edge removal) of the network in a number of subgraphs. At each iteration some centrality measure (degree centrality, betweenness, shortest path) is calculated and the configuration accepted if the new value satisfies some optimisation criteria.

The second group of methods consists of building up the network, starting by considering all the vertices as 'independent' and clustering them hierarchically (Jain & Dubes, 1988; Scott, 2000). Here, too, the process is iterated looking for the optimisation of the network modularity or clustering characteristics.

A graph can be partitioned also by using spectral methods. As seen previously (section 2.3.4) the second smallest eigenvalue of the Laplacian matrix λ_2 indicates the connectedness of a graph and the multiplicity of the 0 eigenvalue equal to the number of connected components (Fiedler, 1973). A spectral decomposition starts by considering the eigenvector u_2 associated with the second Laplacian eigenvalue. This is treated as a one-dimensional drawing of the graph G (each vertex is assigned the value of its corresponding entry in u_2). The median of the u_2 elements gives a separation of the graph into two clusters (Pothen et al., 1990). The process can be repeated for each cluster.[8] An optimisation criterion helps finding the right decomposition (Donetti & Muñoz, 2004).

The different identification methods have different characteristics in terms of computation time needed and of capacity to effectively find meaningful structures. Moreover, peculiar network topologies (trees, disassortative networks, sparse graphs) can affect their overall efficiency and suitability to achieve their goal (Danon et al., 2005).

[8] The spectrum of the Laplacian of an unconnected graph is the union of the spectra of the disconnected components

Modularity, local clustering and scale-free topology coexist in many real systems. One possible way to explain these circumstances is to suppose that a number of clusters combine to form larger clusters, that combine again and so on. Iteratively, a hierarchical structure is created (Ravasz & Barabási, 2003).

3.3 The structure of the WWW

The World Wide Web (WWW) is, without any doubt, the most important and significant phenomenon of the last few years. Its impact on almost all the forms of our social, economical and personal life is enormous. The WWW remains basically uncontrolled: any organisation, institution or individual can create a website, of any size, with relatively little effort. This growth, without any clear rule or centralised control, produces a huge, self-organising, complex system.[9]

Furthermore, for its nature, the WWW is easily measurable in terms of nodes and connections, and it is no surprise that from the study of the Web topology originated the first seminal works on complex network structures (Barabási & Albert, 1999; Faloutsos et al., 1999; Watts & Strogatz, 1998).

The WWW can be described in terms of a large directed graph whose vertices are the html documents and whose edges are the hyperlinks connecting one document to another; its topology determines its connectivity.

The study of the Web as a graph is not only intriguing *per-se*, but it is also of great importance in providing 'practical' answers to significant problems. First of all the one we may denote as *visibility*, that can be easily translated into the problem of finding functional algorithms for crawling (Deo & Gupta, 2001), searching and community discovery (Gibson et al., 1998; Newman & Girvan, 2004). These problems translate directly into an economic problem, for the definite value acquired by the links in easing the tracking down of a website. The connections (hyperlinks) can be seen as a pseudo-monetary unit (Walker, 2002).

---

[9] This section is a partial edited version form: Baggio, R. (2006a).

A *copying mechanism* has been proposed by Kumar et al. (2000a) to explain the degree distribution for the Web. Following them, a new page on a certain topic *copies* the links from those found on existing pages. At time t a new node connects to the others with a constant number of directed links and it is randomly connected by old nodes.

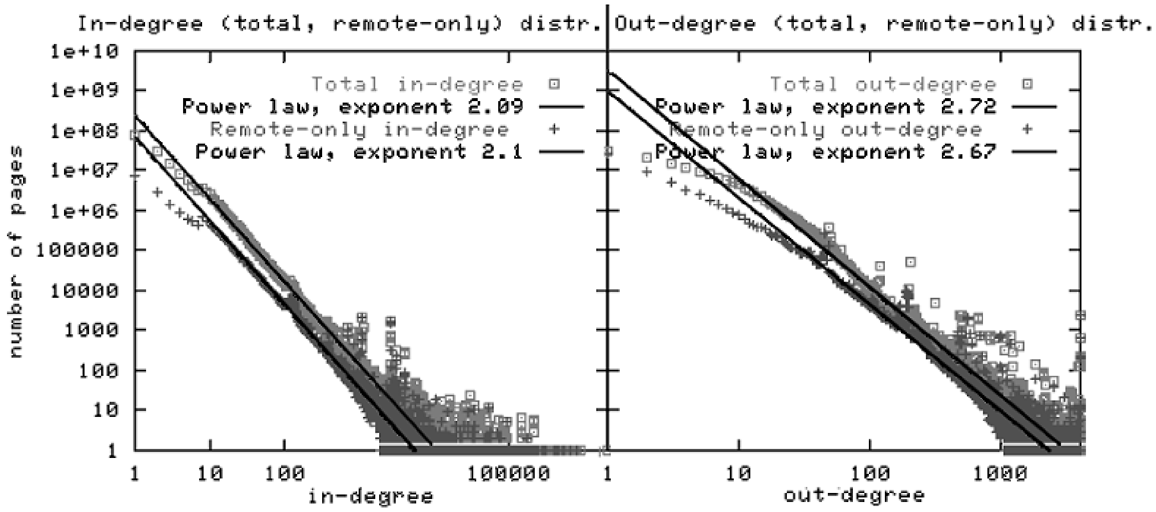


Figure 3.20 In-degree and out-degree distributions for the copying mechanism (after Kumar et al., 2000b)

Arcs outgoing from a node are distributed in this way: destination of link j is taken at random with probability p; the node is a destination (has incoming links) with probability $1-p$. The chance that a web page with degree k receives a link is proportional to $(1-p)k$. This indicates that the copy mechanism works with a linear preferential attachment and gives rise to a power-law degree distribution. Figure 3.20 shows the results of degree distribution measurements.

For the incoming links the distribution is:

$$P(k_{in}) = k^{-2(2-p)/(1-p)}$$

the exponent varies between 2 and ∞.

Based on these considerations, a general model for the structure of the web space has been proposed. The current model is mainly due to the research by Broder et al. (2000), Dill et al. (2002), and Flake et al. (2002).

The bow-tie model (as it is commonly called) sees the Web as a self-organising, self-similar structure, basically divided into three main components:

- a core of strongly connected nodes (SCC), accounting for almost 28% of the web pages in the sample studied, all joined with bidirectional links;
- a set of pages (IN) mainly connected in a monodirectional way to the SCC. The pages in this component (almost 21% of the sample) have outgoing links that reach the SCC, but they are virtually unreachable by other parts of the web; and
- a similar sized (21%) set of pages (OUT) reachable by the pages contained in SCC, but whose links are mainly inward bound; i.e. there is always a path from SCC to OUT pages, there is no direct connection from OUT to SCC or IN.

The picture is completed by two more sets: TENDRILS composed by pages (again 21%) providing paths from IN pages or to OUT pages without passing through the SCC elements.

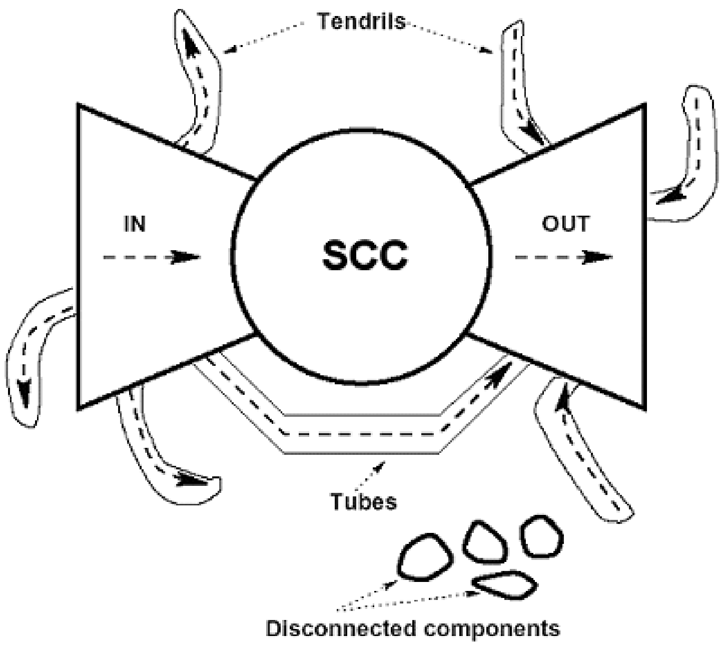


Figure 3.21 The bow-tie, a general model for the Web space (after Broder et al., 2000)

The TENDRILS contain pages that cannot reach the SCC, and cannot be reached from the SCC. It is possible for a TENDRIL hanging off from IN to be hooked into a TENDRIL leading into OUT, forming a TUBE – a passage from a portion of IN to a portion of OUT without touching SCC. To complete the model, there are pages forming a sort of islands disconnected from the other main components (DCC). The general picture, a bow-tie like graph, is shown in Figure 3.21.

The supposed high global connectivity of the Web appears to be not so high. Only in 24% of the cases it is possible to find a path between any two nodes chosen at random (Broder et al., 2000). The average distance between nodes is ~16, and it goes down to 6.8 if we consider the undirected connections in SCC (Table 3.1).

Table 3.1 Average distance between web pages

	incoming links	**outgoing links**	**bidirectional links**
average distance	16.1	16.1	6.83

The degree distribution, on average, follows a power law $P(k) \sim k^{-\gamma}$. The values of the exponent γ are: $\gamma = 2.1$ for the in-degree distribution (incoming links to a page) and $\gamma = 2.72$ for the out-degree distribution (links outgoing from a page).

In social network analysis, out-degree indicates expansiveness and in-degree indicates popularity. The difference (although small) in the measured distributions may be interpreted as a relatively higher tendency to receive links because of their popularity rather than opening towards the external world of web pages (a steeper power law means that the number of websites with *out-degree* = k_{out} decreases much faster).

These values, or values very close to these, have been confirmed by several different empirical studies, that confirmed also the value for the web diameter $D \sim 7$, and the average degree of a node in the Web $\langle k \rangle = 7$ (Pastor-Satorras & Vespignani, 2004).

3.4 Conclusions

The topology of SW or SF networks is directly related to the peculiar characteristics of their behaviour. Besides the numerous empirical investigations whose measurements are shown in Table 2.1, the most important results obtained in this field include the following processes (section 3.2):

- *rich-get-richer effects*: the accumulation of some quantity (links, but also money or knowledge) directly coming from the preferential attachment mechanisms;

- *robustness*: stability of the system to random removal (or failure) of randomly chosen elements, therefore a higher capacity to resist cascading failures or avalanches of breakdowns; but also
- *fragility*: high sensitivity to targeted attacks to the most connected hubs;
- *congestion factors*: a random graph, tends to become more and more dense (in links) with increasing network size, indicating that these networks will become extremely 'congested' in this limit. For SF networks, however, this congestion factor does not depend on the network size; arbitrarily large networks can be considered without increasing their congestion level;
- *low internal friction*: extent and speed of *disease* transmission (viruses, but also messages, fads, beliefs, knowledge etc.) are greatly improved in a SF or SW network with respect to a random ER graph; in some cases it is shown that there are no critical thresholds at all for these phenomena.

Moreover, there exists the possibility of optimising a network topology with respect to some specified characteristic (e.g. efficiency or clustering).

Despite its almost uncontrolled growth, the WWW exhibits a clear structure which is very far from the random network one would have expected. This reinforces the idea that self-organisation mechanisms govern the evolution and the growth of a wide class of networks.

4 Theoretical background: industrial economics and tourism destinations

Defining tourism and tourists is a challenging task (Cooper et al., 2005; Jafari, 2000; Leiper, 1979; Pearce, 1989; UNWTO, 1999), but the words used by Gunn and Var to describe it are enlightening (2002: 4):

> ... tourism in its full breadth is much more complicated and for many reasons. First of all, tourism itself is an abstraction. It doesn't exist, at least not in the same sense as a residence. Tourism is not even a discipline, such as chemistry or geography. Tourism is a field made up of many physical, program, and action parts. It is only the pieces of tourism and their aggregation that can be planned. Tourism is not under the control of one owner, it has no CEO. It is controlled by a multitude of owners, …

Millions of tourists move every year from their homes and play a part in this *non-existent abstraction*. Their expenditure induces the generation of millions of jobs worldwide and contributes significantly to the GNP of many countries. According to the World Travel and Tourism Council (WTTC, 2007), the world's tourism industry is expected to contribute 3.6% to gross domestic product (GDP) in 2007 (US$1,851.2 billion), with a forecast growth of 4.3% per annum in real terms for the next 10 years. The sector has provided 231.2 million jobs in 2007, 8.3% of total employment, with an expected increase to 86.6 million jobs (2.8% of total employment) in 2017. However we define it, a tourist is a person moving to go somewhere, for a certain period of time (see the definition in UNWTO, 1995). Their target is a *destination*.

A destination is a space (a city, a region, a resort) in which the traveller can find the facilities necessary for accommodation, entertainment, and transportation, that are organised by an intermediary, usually a destination management organisation (UNWTO, 2002a). The evolution of transportation technologies and of the social conditions that initiated the phenomenon of mass tourism, combined with the globalisation of world economies has increased the competitive pressure on destination tourism organisations. In these conditions, as happens for other industries, companies within a destination tend to form bound constellations to better face competitive challenges (Castells, 2000; Krugman, 1991b; Porter, 1998; Schucan, 1998; Smeral, 1998).

The approach followed in this thesis considers a tourism destination as a complex socio-economic system. This chapter analyses and discusses what the literature recognises as the most important models which have been used so far to study such a system. In essence, the constellation of specialised companies, organisations and communities gathered into a confined geographical location (even if its boundaries are often poorly defined) can be seen as a form of industrial cluster or district. Thus the analysis of its structure may draw upon the theory of industrial clusters, of their formation mechanisms and their evolution (see for example Hjalager, 1999).

The main models of clusters and networks of companies have been developed by investigating industrial sectors, with limited attention to the service sectors of the economy such as tourism. Tourism destinations deal with different offerings and have peculiar differences with respect to a 'traditional' district. The formation mechanisms, the focus on the service components, the characteristics of the 'products' and the relationship between them and the 'production' system are some of the peculiarities which make a tourism destination different from an industrial cluster. First of all, tourism is essentially a service industry in which the 'product' (a package) is not well defined and is composed of many different elements (Johns, 1999; Sinclair & Stabler, 1997; Wahab & Cooper, 2001). The tourist purchases this package in advance and consumes it at the destination. The diversity of elements which form the package causes the establishment of a wide layer of intermediaries which are an integral component of the same industry (Gollub et al., 2003). Therefore, the models of industrial networks and clusters need modification and adaptation when tourism is the main object of study (Gnoth, 2002, 2006).

The first part of this chapter discusses the main approaches to the study of an industrial cluster which have been developed in the last 15 years. The main components, their interrelationships and their evolution are first described. The advantages, in terms of economic benefits, of a networked cluster model are subsequently briefly illustrated.

These general models constitute a natural introduction to the discussion on the structure and evolution of a tourism destination. The second part of this chapter is dedicated to the main models used to investigate and understand the behaviour of a tourism destination and its components. Attention is given also to the issues related to the management of such a composite system.

The last part of this chapter combines the network approach (discussed in Chapters 2 and 3) with this framework (the tourism destination district) and discusses the contributions a network thinking approach can provide to the understanding of a tourism destination. This intersection between the three traditions: industrial economics, tourism studies and network science is finally identified as the theoretical background for the study described in this thesis.

4.1 Industrial clusters, districts and networks

The idea and the phenomenon of a clustering of economic activities are far from being new. As Harrison states (1992: 470):

> From the market towns of medieval Europe (and much of contemporary Africa) to the industrial parks, wholesaling districts and shopping centres of any large urban area, it is apparent that similar economic activities commonly cluster together in space.

Fairs, markets and local concentrations of merchants and craftsmen can be traced back to the origins of the economic history of the world. In the Middle Ages they provided a strong impetus and were a characterising feature of many densely populated areas, above all the main cities (Hunt & Murray, 1999; Richardson, 2005; UFI, 2005).

Concentration effects in general economic or industrial activities have been studied and measured in detail. Theoretical and empirical research has found that agglomeration effects generally play a crucial role for regional income levels (Brenner & Weigelt, 2001; Krugman, 1991a, 1991b), for the attractiveness toward foreign investments (Barrell & Pain, 1999) and for the competitiveness of the area in which they occur (Norton, 1992). Moreover, economic growth and geographic agglomeration have been found to be *self-reinforcing* (Martin & Ottaviano, 2001); concentration of industries increases with growth and, by reducing the cost of innovation in the region where the economic activities converge, it enhances growth. Even a simple model in which there is no specific connection among industries, and labour is not movable across regions (Baldwin &

Forslid, 2000), leads to the conclusion that concentration favours the overall growth and vice versa.

An idea at the basis of most literature on industry clustering is that firms sited in a geographical area share common ideals, rules and languages so that the social environment they form is consistent. Social, cultural and operational contiguity favours the spread of *tacit* information and knowledge among local actors. This constitutes a competitive advantage for the participants to the cluster (Morrison, 2004; Norton, 1992) because their *tacitness* makes them difficult to access by elements outside the community. Co-location within a concentrated geographical area is a basis for development of other characteristics of a cluster. The most important factor for a functioning agglomeration is the formation of close ties or alliances among the different actors and the degree of cooperation established to improve the competitiveness of the group beyond the incidental (usually external) effects that promote the gathering (Andersson et al., 2004; Mishan, 1971). In the last decade the clustering of industries has been the objective of a very large number of studies, both in advanced and developing countries (Becattini, 1987; Porter, 1998; Pyke & Sengenberger, 1992; Saxenian, 1994; Schmitz, 1992, 1998). The idea, with its implications for economic development, has also received serious consideration by national and international organisations in their role as policy makers (European Commission, 2001; OECD, 1996, 2001; UNCTAD, 1998).

Another theme in the literature of clustering is devoted to the importance of external linkages and how they enhance competitiveness and reduce the risk of negative lock-in (Giuliani et al., 2005; Humphrey & Schmitz, 2001, 2002). This effect, a kind of close self-referentiality, has been seen as one of the main deterrents to the achievement of all the claimed positive advantages and may explain some of the economic failures in developing countries or the decline of well known industrial regions (Hassink & Shin, 2005).

Alfred Marshall introduced the concept of industrial clustering to the economic literature. In his ‘Principles of Economics’ (1920: Book IV, Chapter X) he describes ‘the concentration of specialized industries in particular localities’ that gather with a constructive cooperation. He establishes a relation between the geographical grouping of firms and the benefit from positive externalities associated with their activities. This framework was revived by Italian scholars in the second half of 20th century (Bagnasco, 1977; Becattini, 1979, 1987) to explain the gap in economic performance between the Northeast and Central Italian areas and the Southern regions. The difference was attributed to the clustering of industries and their intrinsic capacity to foster innovation in products and processes while achieving economies of scale. The idea was further developed by French researchers with the concept of *milieu innovateur* (innovative environment). In this meaning, the innovation process is endogenously favoured by a number of factors that include not only economic causes but also social, cultural and environmental elements. These combined factors contribute to the establishment of a unique set of externalities that stimulate innovation and knowledge transfer. Subsequent evolution of this approach has considered the effects of the dynamic connection between the *milieu* and the external environment (Aydalot, 1985, 1986; Quévit & van Doren, 1997).

Two other lines of research contributed to the literature regarding the concentration of businesses in a delimited geographical area. One is the work of Porter on strategy and competitiveness (Porter, 1990) and the other is Krugman's proposal of a new economic geography (Krugman, 1991a, 1991b). Porter's analysis starts from an explanation of the 'competitive advantage of nations' which results from the interaction of four main forces:

- firm rivalry and new entry of competitors;
- power of upstream suppliers and producers of machinery;
- threat of substitutive products;
- factors conditions and demand conditions.

The combination of these factors guarantees a strong competitiveness to a nation and is fundamentally rooted in the structure of its industry. Later, Porter (1998: 78) better specified the role of industrial concentrations and their part in the economic environment:

> Today's economic map of the world is dominated by what I call clusters - critical masses, in one place, of unusual competitive success in particular fields. Clusters are a striking feature of virtually every national, regional, state and even metropolitan economy, especially in more economically advanced nations. Clusters are not unique, they are highly typical, and therein lies a paradox: the enduring competitive advantages in a global economy lie increasingly in local things - knowledge, relationships, motivation.

The functional Porterian concept of cluster is similar to the spatial approach used by Krugman, who writes (1991a: 5):

> …what is the most striking feature of the geography of economic activity? The short answer is concentration. Think of the United States: most of the population of a huge, fertile country lives along parts of two coasts and the Great Lakes; within these belts, population is further concentrated in a relative handful of densely populated urban areas. As I will document in the next lecture, these urban areas in turn are highly specialized, so that production in many industries is remarkably concentrated in space.
>
> This geographic concentration of production is clear evidence of the pervasive influence of some kind of increasing returns.

The agglomeration into clusters is here considered a consequence of the presence of local external economies and determined by increasing returns to scale and imperfect competition. These two works complement each other. From Krugman's point of view, international trade forces cause a concentration of industries in particular places, while the strong interconnections among firms are, according to Porter, a major determinant of the competitiveness of a region or a country.

Many industrial sectors, and tourism above all, show a large prevalence of small and medium enterprises (EUROSTAT, 2006; OECD, 2000b; UNCTAD, 1998). This induces the need for special attention due to the peculiar characteristics of these companies: limited resources and impact in the marketplace, low specialist expertise and narrow competences in *networking* (Gilmore et al., 2001). More contributions to this issue come from a line of research born among the regional scientists and investigators of small and medium enterprises (SMEs). This addition consists, mainly, in the recognition of the importance of the role of public policies and non-market coordination (Scott, 1995, 1998) and the SMEs' tendency towards the integration of public support to the cooperation of the private organisations (Schmitz, 1992). The latter strand argues that an industrial district is successful when it is embedded in an environment that shares risks across a combination of public and private institutions. In this enlarged ensemble, the costs and the development efforts for new technologies, training and education of skilled human resources, expansion towards new markets and the raising of capital are shared by a network of institutions (Humphrey & Schmitz, 2002; Schmitz, 1992, 1998; Schmitz & Nadvi, 1999). This *collective efficiency approach*, while acknowledging the value of Marshallian local external economies, emphasises the fact that these, by themselves, do not give a satisfactory explanation of the growth and the competitiveness of an industrial cluster.

Criticisms of the concept of clusters mainly build on its relatively poor theorisation, the many different methodological approaches and on the lack of rigorous testing (Martin & Sunley, 2002: 7):

> … the mere popularity of a construct is by no means a guarantee of its profundity. Our argument here is that seductive though the concept is, there is much about it that is problematic, in that the rush to employ 'cluster ideas' has run ahead of many fundamental conceptual, theoretical and empirical questions …. Whilst it is not our intention to debunk the cluster idea outright, we do argue for a much more cautious and circumspect use of the notion, especially within a policy context.

Moreover, a too heavy dependency on agglomerations based on strong connections among the participants may risk a tendency to closeness, which results in an increase in internal rigidities and a reduction in the capacity to adapt to external dynamic conditions (Saxenian, 1994). Geographically bound concentrations may become strongly self-referential systems, generating a self-sufficiency syndrome and developing lock-in effects. These can be of a functional nature: firms rely heavily on direct relationships with other organisations and do not feel the necessity to develop functional areas of expertise (marketing, for example). These may give rise to a soundly shared 'view of the world' which may lead to misinterpretations of market or economy dynamics inducing wrong responses (Grabher, 1993; Hassink & Shin, 2005).

4.1.1 Structure and evolution of an industrial cluster

The numerous contributions to the analysis of industrial agglomerations have led, as seen in the preceding section, to different approaches and definitions (Andersson et al., 2004). Nonetheless, analysis of the literature indicates some agreement on the basic characteristics of a cluster and the crucial factors that contribute to their formation as well as attempts to reconcile different interpretations, such as the Italian post-Marshallian school and the American geographical school (Belussi, 2004).

Industrial agglomerations share a number of common characteristics that have been used to describe them (and sometimes to define the concept or the categorisation). Most authors (Andersson et al., 2004; European Commission, 2001; Jacobs & de Man, 1996; OECD, 2001) speak of:

- *spatial (geographical) proximity*: firms gather in geographic proximity pushed by factors such as external economies of scale, social capital and learning processes. The spatial concentration is a distinctive characteristic, even if the concept of spatiality may need a deep revision if contemporary communication technologies are taken into account (see the discussion on virtual clusters further on in this section);
- *multiplicity of actors*: clusters and districts do not only consist of firms, but also involve public authorities, educational institutions and financial organisations. A critical mass is required to achieve the main benefits and to start a dynamic evolution;
- *specialisation*: clusters are focused on a core activity (or specific industrial sector) to which all actors are related. There is a considerable division of work: different elements of the cluster dedicate themselves to specific processes or products;
- *collaboration*: interconnections exist among the cluster elements, the industrial structure and the local institutions. Industrial sectors include the social division of labour and links between customer, suppliers and competitors. Local institutions and culture contribute to create and sustain regular patterns of social interactions;
- *competition*: essentially based on innovation rather than on pure economical bases (prices/costs/wages). A dynamic and equilibrated combination of competitive and collaborative aspects is probably the most striking characteristic of a cluster;
- *innovation*: firms in clusters are continuously involved in processes of technological, commercial and organisational change. These processes require, and are supported by, effective and efficient means of knowledge creation and distribution;
- *institutional support*: state, regional or local organisations with a supportive policy are thought to be crucial factors for the success of many districts;

- *life cycle*: both the system and the clusters and the initiatives and the activities are not temporary short-term phenomena or static ensembles, but are ongoing with long-term perspectives, strategies and plans;
- *cultural homogeneity or social embeddedness*: a socio-cultural identity that promotes trust, reciprocity and social sanction. It stems from the 'atmosphere' considered by Marshall as one of the main features of an industrial district. It has been emphasised mainly by the Italian school (Sforzi, 2000; Sforzi & Lorenzini, 2002). In this vision a district, or a cluster, is a sociological entity besides (or even before) being an economic entity in which the cultural heritage and the historical traditions may play a crucial role in strengthening the whole system structure.

The structural qualities of clusters differ greatly across locations and industries: some have a 'flat' structure, some exhibit a clear hierarchical shape, some form around a central pole, some seem to be characterised by randomly drawn connections (van der Linde, 2003). So far, little has been done to explain the cause of structural features of these agglomerations and how the network of relationships, a basic ingredient of clusters (Biggiero, 1999; Staber, 2001), influence or are influenced by the topology of these systems. More importantly, the origins of an industrial cluster are not known although this lack of understanding may be because it "is such a complex and serendipitous phenomenon that no theoretical approach neither econometric estimation may pretend to fully explain it" (Maggioni, 2004: 31).

As noted above, a common and clear definition of such a cluster does not exist. Broadly, three main dimensions have been identified (Jacobs & de Man, 1996):

- *spatial concentration*: generally refers to a geographical region in which economic activities within specific sectors assemble and a diffuse support (physical or knowledge-based) infrastructure forms the connective tissue;
- *production chain*: full supply chains (sometimes extended to supporting activities) or organisations active in different stages in the production process form the core of the cluster;
- *economic or industrial sector*: single or combined sectors, defined, rather broadly, at a high level of aggregation.

Several attempts have been made to classify the main types of clusters observed. From a series of studies (Giuliani et al., 2005; Guerrieri & Pietrobelli, 2001; Harrison, 1992; Johnston, 2004; Marceau, 1994; Markusen, 1996; Pyke & Sengenberger, 1992) the following general taxonomy can be deduced:

- *horizontal clusters*: between firms (typically small and medium-sized) in an economic sector that both compete and collaborate with each other;
- *vertical clusters*: between large firms and their core suppliers;
- *virtual clusters*: where physical co-location is not important, or is overcome by extensive use of communication technologies;
- *proto-clusters* (at times called emerging clusters or by using the name of the common base, such as technology or peculiar needs): where firms have a common resource base, technology or needs, but in which the system of relationships in production and innovation is still at an embryonic stage; and
- *mixed configurations*: for example it is possible to identify diagonal clusters, in which parts of a vertical cluster are formed by smaller horizontal clusters formed by elements of the supply chain.

Evolution of an industrial cluster

The dynamic history of industrial clusters and districts has been analysed by numerous scholars and is divided into three major phases (Andersson et al., 2004; Maggioni, 2002, 2004; Schmitz & Nadvi, 1999):

- *an initial stage*, in which the cluster is initiated by some (often exogenous) event and sustained by a spontaneous diffusion of information about the suitability of the location;
- *a second stage*, in which the economical drivers of Marshallian agglomeration become crucial in sustaining (endogenously) the growth and the structural setup;
- *a third stage*, in which a successful cluster attains a leadership in the sector of competence, reaches a stable maturity and becomes resilient, i.e. it is able to successfully absorb external and internal shocks (economic and/or technological). If the cluster does not succeed, it declines and disperses, producing unemployment,

migrations or forcing public organisations into heavy interventions to support the social environment.

This typical evolution is well depicted by a classical logistic *s-curve* (Figure 4.1), in which the evolutionary path is given, customarily, as a function of the *industrial mass*: that is, basically the number of firms participating in the cluster.

The process of cluster development is described as having high internal dynamics and is highly path-dependent: the order of ‘entry’ of the different companies matters. The production chains arranged within a cluster undergo a continuous rebalancing through a series of small adjustments. No central regulatory body seems to govern the system, thus being an example of self-organised and decentralised coordination (Belussi, 2004; Belussi et al., 2003; Biggiero, 1999; Brenner, 2000; Brenner & Weigelt, 2001).

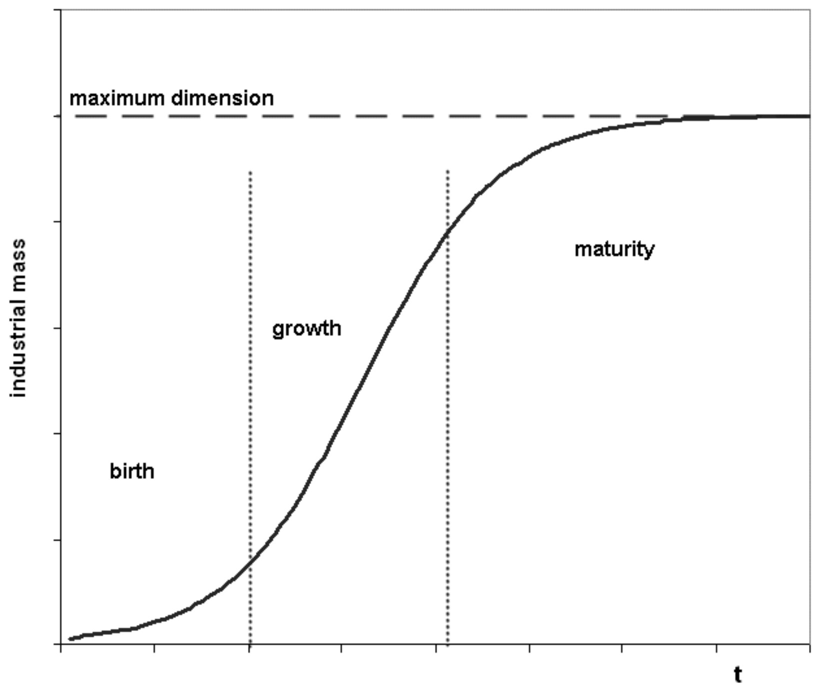


Figure 4.1 The life cycle of an industrial cluster

This recognition of the inherent complexity of these systems has led to a number of attempts to use non-linear techniques and models to explain at least some of the characteristic features of such agglomerations. Agent-based, cellular automata, or neural networks along with game theoretic models have been used to simulate both the emergence and the evolution of an industrial cluster (Biggiero, 1999; Brenner & Weigelt, 2001;

Fioretti, 2005; Giaccaria, 1997; Maggioni, 2002, 2004; Pontes, 2000; Staber, 2001). So far, the results obtained reproduce or reconfirm the empirical findings of the vast catalogue of case studies published. Computational simulation approaches are likely to be improved and enhanced in the future. Another approach using the recent results of 'network science' to investigate the structural and dynamic characteristics of an industrial cluster is discussed further on in this chapter.

The main determinants for the development of a cluster are, as seen above, factors connected with the entrepreneurial spirit, the location atmosphere and supportive policies. However, not all these elements have the same importance or influence. In a review of the published literature (almost 400 papers on more than 700 clusters), van der Linde (2003) identifies competitiveness as the main measure of industrial cluster success and applies Porter's (1990) framework to the analysis. The results indicate the Porterian *factor conditions* (i.e. resources and infrastructures) as the most commonly reported determinant. This is followed by demand conditions, related and supporting industries, the context for strategy and rivalry and, finally, other reasons including chance, isolated individual entrepreneurs, and early mover advantages (Figure 4.2). The same factors influence the establishment of a cluster in a region (Figure 4.3). In this case, however, the importance of the *other factors* is more significant, mainly for what concerns the role of the public institutions.

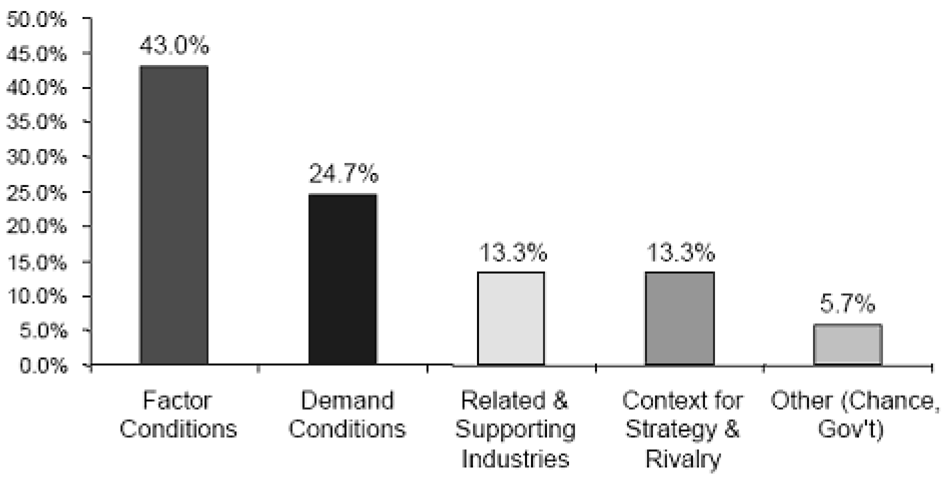


Figure 4.2 Determinants of cluster competitiveness (after van der Linde, 2003)

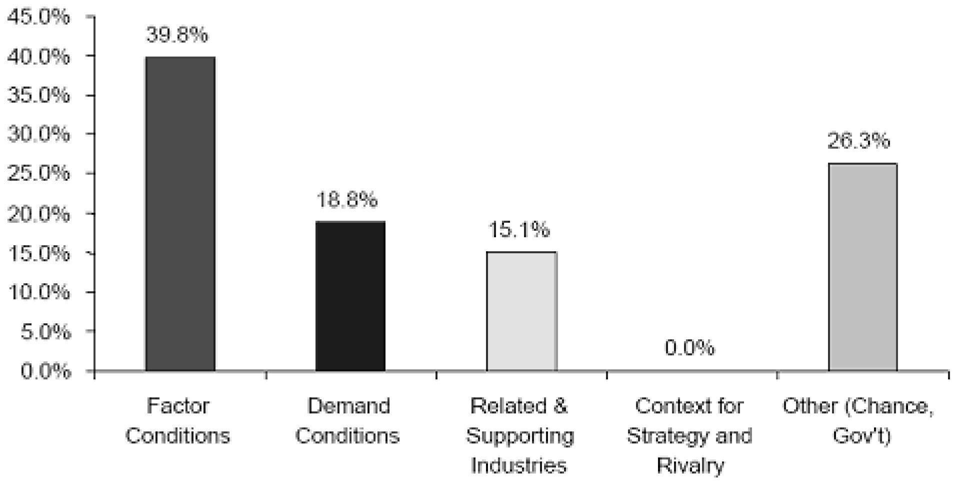


Figure 4.3 Most important reasons for cluster establishment (after van der Linde, 2003)

As discussed above, the spatial dimension has been the basis of the concept of an industrial cluster since its first appearance. The recent developments in information and communication technologies have extended this idea of *spatiality*, translating it into the virtual environment of cyberspace. While traditional causes of clustering are generally recognised to be specialisation and geographic proximity, it has been argued that organisational proximity plays a significant role in substituting geographic nearness as an enabling factor for the formation and development of coordination and knowledge exchange (Romano et al., 2001). Advanced management strategies of supply chains and customer relationships enabled by the contemporary ubiquitous information and communication technologies are the bases for the achievement of added value groupings (Ahuja & Carley, 1999; Rayport & Sviokla, 1995).

The virtual cluster phenomenon is the one of the most significant consequences of the development and diffusion of information and communication technologies in the business environment. A virtual cluster is formed by an e-business community, composed of customers, suppliers, distributors and commerce providers that use shared digital networks for establishing collaborative and competitive environments. Common infrastructures, cooperation agreements, common business practices, mutual trust, sense of community and stability are the main ingredients of a successful virtual cluster (Bönke et al., 2001; Hagel & Armstrong, 1997; Jones, 1997; Kwon et al., 2003; Rheingold, 1993).

The geographic factor, however, still plays a role even in this scenery. Many such groupings, in fact, arise as a technological counterpart of a ‘traditional’ district and then develop over the virtual space. This is particularly true for service enterprises and is a characteristic of the travel and tourism sector, in which the technological dimension is thought to be a crucial determinant for a successful development at all levels: single organisation, groups, areas, regions and countries (Antonioli Corigliano & Baggio, 2003; Archdale, 1995; Buhalis, 2003; Poon, 1993; Werthner & Klein, 1999).

A short summary of industrial clusters

Geographical concentration effects in economic or industrial activities have been extensively studied both in advanced and developing countries. Theoretical and empirical research has found that agglomeration effects generally play a crucial role for regional economies. The sharing of common ideals and practices and constructive cooperation among firms make consistent the socio-economic environment they form. This, coupled with their external linkages, influences their economic competitiveness and contributes to raising the competitive advantage of the country they belong to.

Functional and spatial concentrations of industries are reputed to be mainly due to the presence of local external economies and determined by increasing returns to scale and imperfect competition. To these another strand of research has added the recognition of the importance of the role of public policies and non-market coordination and of the SMEs’ tendency towards the integration of public support into the cooperative forms of the private organisations.

This *collective efficiency approach*, however, does not provide a full satisfactory explanation of the growth and competitiveness of an industrial cluster. Some criticisms have also been directed towards these concepts on the basis of relatively poor theorisation, too many different methodological approaches and lack of rigorous testing.

Despite the different approaches and definitions, the literature shows some agreement on the basic characteristics of a cluster and the crucial factors that contribute to their formation: spatial (geographical) proximity, multiplicity of actors, specialisation, collaboration, competition and innovation. The main determinants for the development of a cluster are the factors connected with the entrepreneurial spirit, the location atmosphere and supportive policies. However, not all these elements have the same importance or influence. The structural qualities of clusters differ greatly across locations and industries, so that no common classification schemes exist.

The dynamic history of industrial clusters and districts has been analysed by numerous scholars and is divided into three major phases:

- an initial stage, commencing the development of the cluster;
- a second growth stage;
- a third maturity stage.

The recognition of the intrinsic complexity of these systems has led to a number of uses of non-linear techniques and models to explain their characteristic features. Emergence and evolution of an industrial cluster have been simulated with good success, and empirical findings have been reproduced and reconfirmed. These computational simulation approaches are likely to be improved and enhanced in the future. Another possible approach is in using the recent results of 'network science' to investigate the structural and dynamic characteristics of an industrial cluster.

Even if the spatial dimension has been at the basis of the concept of an industrial cluster since its first appearance, the developments in information and communication technologies have, as a significant consequence, extended the idea of *spatiality* to the virtual environment of cyberspace.

The geographic factor, however, still plays an important role. Many groupings, in fact, arise as a technological counterpart of a 'traditional' district and then develop over virtual space. This is particularly true for service enterprises and is a characterising feature of the travel and tourism sector.

4.1.2 Cluster economics and benefits

Industrial districts and clusters provide competitive advantages to the firms that take part in them. The most commonly quoted benefits are (Belussi, 2004; Storper, 1995):

- increasing returns determined by the systemic characteristics of the local group in a globalised context;
- reductions in transaction costs;
- innovation diffusion and technological development due to the local interactions;
- reductions in costs by effective learning (by imitation and emulation);
- benefits provided by the localisation of external economies (specialised labour market, local division of labour, specialised suppliers);
- first mover advantages as of initial geographical specialisation;
- advantages connected with product diversification and for being customer driven organisations.

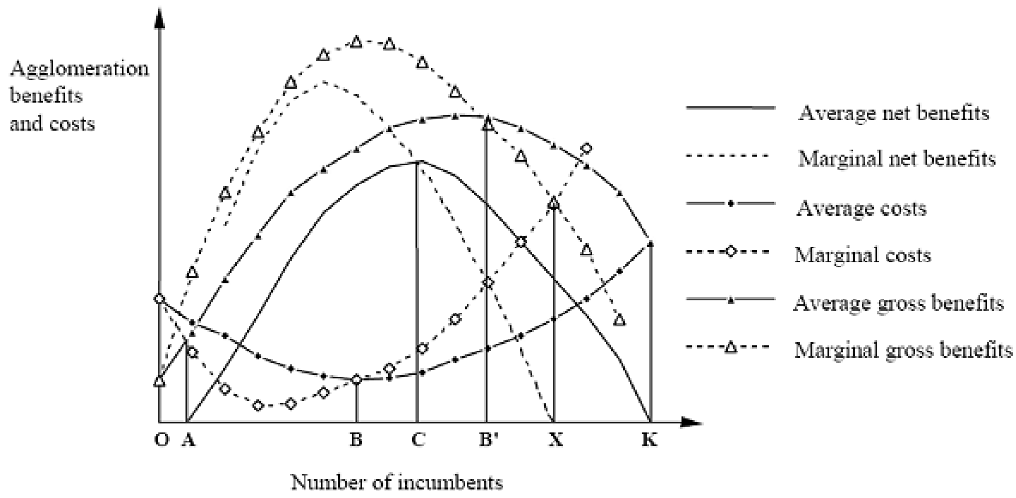


Figure 4.4 Marginal and average costs and benefits for new incumbents in a cluster (after Maggioni, 2004)

Costs and benefits are related to the formation of a critical mass of industries. Some models have attempted to derive formal relations to locate the optimal size of a cluster (Figure 4.4), i.e. the industrial mass which maximises benefits while minimising the entry costs (Maggioni, 2004).

The set of relationships existing among the firms in a cluster or district is one of the main factors influencing the economics of the system. The interest in this topic is mainly due to the seminal work by Katz and Shapiro (1985) and by Arthur (1989). The essential concept introduced by them is the *network externality*, the idea that additional members of the community contribute to the utility of the whole group increasing, the economic advantages for all of them. This is a challenge to the conventional use of decreasing returns and confers a primary role to economies of scale.

From the supplier point of view, the externalities are demand-side economies of scale; as the value of the network increases, the price that consumers are willing to pay rises. This situation may lead to some risks. As has been noted (see for example Sääskilahti, 2005; Shy, 2001), if we consider two competing clusters and there is an incompatibility between the products offered by the two, consumers may become locked-in. This lock-in generates a switching cost, which reduces incentives for price reductions and 'freezes' customers. In order to convince consumers to switch networks, a firm must find some means of compensation for the lost benefits of the rival system. In other words, a perfect compatibility reduces specific strategic values. Consumers do not discriminate between firm-specific networks, but can differentiate between products in terms of brand or quality. Although criticised by some researchers (Liebowitz & Margolis, 1994) the concept has been used innumerable times to describe and forecast the economic behaviour of a system.

Digital virtual clusters experience an amplification of these effects mainly due to the fact that the digitalisation of goods has a very low marginal cost. Therefore, the system undergoes huge economies of scale, so that even in the case of lock-in, the switching costs are very low or non existent.

It must be noted that, from a structural point of view, the traditional model of network externalities assumes that each network actor is connected to everyone else; the links form a complete graph. An added member increases the utility of all of the network elements. That is to say that the value of the network follows the so-called *Metcalfe's Law* (Gilder, 1993; Metcalfe, 1995): the value of a network increases as the square of the number n of the network elements. It is proportional to $n(n-1) \sim n^2$ (assuming the benefit for each link is normalised to 1 unit of *utility*).

This homogeneity is not found in many real world networks (see Chapter 2). In a real social network, typically, some actors have more contacts than others, and some relations are more important than others. The asymmetry of the network topology is usually absent in the network externalities literature. The hypothesis of a complete graph eases the analysis, but it may lead to an incorrect estimation of many characteristic parameters of the network, above all its value. As an example, Sääskilahti (2005) studies size and topology effects on monopoly pricing. The result is that the topological effect dominates the size effect, and the monopolist incorporates the topology in its price. Therefore, asymmetric topologies produce distribution surplus effects due to the monopoly's pricing strategy. Moreover, when networks derive their characteristics from specific physical assets needed to support it (the roads in a transportation network, for example), bottlenecks (a bridge in a transportation network) may exist and the 'owner' of the bottleneck asset is able to exercise monopoly power.

Other important phenomena are greatly affected by the topology of the district network. These may concern general diffusion processes (Balázs et al., 2005; López-Pintado, 2004) such as the diffusion of knowledge (Chen & Hicks, 2004), in the proto-forms of rumours or fads (Grönlund & Holme, 2005; Zanette, 2002), basic information (Watts, 2002; Wu et al., 2004), technology innovations (Guardiola et al., 2002) or policy issues. The latter are of great importance, mainly for public institutions seeking ways to promote or sustain a cluster. For these the mechanisms presiding over the formation of opinions or the building of consensus are critical; so is the optimisation (or the possibilities of optimisation) of the topology towards these mechanisms. In this area too, the topology of the set of actors and ties plays a significant role (Laguna et al., 2005; Sobkowiz, 2003). In examining the life cycle of an industrial district, its robustness or resilience is a crucial characteristic. The capacity to absorb internal or external shocks can determine the survivability of the network and these characteristics are strictly related to the shape of the network (Crucitti et al., 2003, 2004b; Kogut & Walker, 2001; Pastor-Satorras & Vespignani, 2003).

Basically, all of the results discussed in Chapter 2 can be applied to the study of an industrial district, deriving explanatory models for the observed behaviours or indications for action in specific conditions. This has been applied in several situations, and, conversely, the analysis of several networks has led to the construction or the verification of the models.

4.1.3 Clusters and districts

In the preceding discussion, the terms *cluster* and *district* have been used almost interchangeably. However, the two concepts have a fundamental difference. This can be seen in the words of the two most influential scholars in this field.

According to Porter (1990), industrial clusters have a central role in guaranteeing the competitiveness of a country. His analysis is based on the idea that a nation's economy is made of a mix of clusters whose conditions and evolution directly affect the state of the economy. The idea of cluster is therefore expressed as (Porter, 1998: 78):

> … geographic concentrations of interconnected companies and institutions in a particular field. Clusters encompass an array of linked industries and other entities important to competition. They include, for example, suppliers of specialized inputs such as components, machinery, and services, and providers of specialized infrastructure. Clusters also often extend downstream to channels and customers and laterally to manufacturers of complementary products and to companies in industries related by skills, technologies, or common inputs. Finally, many clusters include governmental and other institutions - such as universities, standards-setting agencies, think tanks, vocational training providers, and trade associations - that provide specialized training, education, information, research, and technical support.

The basic feature characterising an industrial cluster is belonging to a common specific sector; the participating firms are connected by horizontal or vertical relationships and concentrated in a geographical area. They may be complemented by some 'external' entities such as the public institutions. The focus, however, stays on the entrepreneurial and business dimension.

On the other hand, the vision of the Italian school interprets a *district* as an overcoming of this specialised spatial concentration idea. As Becattini (1990: 38) writes:

> I define the industrial district as a socio-territorial entity which is characterised by the active presence of both a community of people and a population of firms in one naturally and historically bounded area. In the district, unlike in other environments, such as manufacturing towns, community and firms tend to merge. The fact that the dominant activity is an industrial one differentiates the industrial district from a generic 'economic region'. Self-containment and the progressive division of labour which occurs there bring about an increasing surplus of final products that cannot be sold in the district. There follows an ever-increasing problem of placing

> this surplus on the external - essentially world-wide - market. Such a condition for the survival of the district (the necessity of solving an ever bigger problem of final demand) excludes the possibility of accidental placing of the products of the district on the external market, and requires instead the development of a permanent network of links between the district and its suppliers and clients. Thus, an economic definition of the industrial district which aims at being comprehensive will have to add such a permanent network, and all its interactions with the other elements, to the above-mentioned 'local' conditions.

The contrast with the Porterian vision is significant. In both there is a focus on 'industries', but in Becattini's vision there is a full recognition of the importance of the social environment of the area in which the district works and, perhaps more important in this *globalisation* age, the understanding of the role of the linkages with the 'external' world. Moreover, the firms in a district (often with the help of regional governments and trade associations) intentionally network with each other to solve problems of cycles and over-capacity and to respond to new demands for flexibility. This approach seems to be much closer to what the reality of these agglomerations is and much more suitable as a framework for the study of a tourism destination.

Based on these considerations and the analysis of the literature (see also the list in Belussi, 2004), the following definition can be used for an industrial district:

> A spatially contiguous group of interconnected companies, linked by commonalities and complementarities, involving a localised support infrastructure, with a collective vision for economic and social growth and based on competition and cooperation.

4.2 Tourism destinations

Classical economic theory describes a market sector in terms of demand and supply and of the factors that drive their behaviour. For a tourism system these basic components are identified in the potential or actual visitors and their destination, which is the goal of their travel (Cooper et al., 2005; Gunn & Var, 2002). The growth of the economic outcomes from tourism activities has brought to the attention of researchers and practitioners the ensemble of actors and resources generating these outcomes. The destination, as main target for a tourist, has always been recognised as important, and has been termed *the* *raison d'être for tourism, the reason for travelling* (Cooper et al., 2005: 77). However, this concept has become a central one in tourism only in the last decade.

The development of the general economic, social and technological environment of the world has radically modified the attitude towards 'leaving the usual place of living'. A wider diffusion of information, less expensive travel costs, more disposable free time and higher cultural levels have all contributed to the increase and diversification of the basic motivations for travelling and have greatly enlarged the possibility of choice of the ultimate goal for a trip (Schucan, 1998; Smeral, 1998). This has had significant consequences on the supply side of the tourism market. In order to benefit from the economic outcomes of travelling visitors, organisations and firms have had to increase their efforts to promote and sell their products, basically reinforcing marketing activities, striving to enhance differentiation and focusing on customer service and care. Tourism enterprises are mostly small and medium sized organisations and therefore these external forces have imposed the formation of close collaboration relationships among many of them in order to take advantage of the benefits of such agglomerations and achieve competitiveness (see section 4.1.2), which is in many cases essential for their survival (Smeral, 1998). The process has, thus, led to the formation of different kinds of clusters of independent supply operations, in which interdependent networked structural forms have emerged (Mill & Morrison, 1992; Urry, 2002). The *tourism destination* has, as a consequence, assumed a central role in the field.

This, obviously, raises an issue on the meaning of this concept and on the usefulness and importance for the study of tourism as an economic and social phenomenon. As Framke writes (2002b: 1):

> One of the most used words in the field of tourism is destination – used in marketing, planning and development, and general research. One can find it in tourist guides, brochures and homepages, and, of course, in all kinds of textbooks and readers in tourism. So it is obvious that one should be suspicious. How can one word contain so many of tourism's aspects? Investigating the use of the word shows various ways of using it. The word has no unique content, its meaning depends on ones' purpose, be it describing, communicating or analysing tourism. By taking a geographic perspective one will evidently recognise the differences: 'destinations' can be found in various dimensions and on various regional levels: there exist

> static destinations in connection to one's stay at a certain location; and there exist dynamic destinations, where movement is the motive for a vacation; you can find destinations connected to networks and other relations in the industry, and so on.

And some even contend more strongly that the concept itself could be wrong. For example Leiper (2000a: 366) asserts:

> There is no evidence that any destination ever attracted, in a literal sense, any tourists. […] The main causal factors of tourist flows are not located in destinations but in traveller generating regions, in places where trips begin, where the forces that stimulate tourists' motivations are located and where marker systems directing tourists to nuclear elements of attractions begin.

However, the successful outcomes in terms of economic and social development of many regions in the world do not indicate that the destination is just an abstract idea. Also the theoretical framework of the agglomeration of industries in districts and clusters discussed above, gives us a stronger reason to analyse these 'objects of study'. The many different approaches to the study of a destination have generated a number of different definitions. Obviously, they do not always agree.

The initial idea of tourism destination is closely bound to the geographical characteristics of a destination. The first attempt to define it is attributed to Georgulas (1970, quoted in Framke, 2002a: 95):

> Tourism as an industry occurs at 'destination areas' – areas with different natural and/or man-made features, which attract non local visitors (or tourists) for [a variety of] activities.

The influence of the research on industrial agglomerations, however, led to the realisation that a purely geographical approach was not fully adequate, nor able to represent the richness and the complexity of a tourism destination. None of them have been fully satisfactory mainly because of the difficulties in fully outlining the 'product' provided or in characterising the components of the tourism market or the belonging to 'tourism' of the many diverse economic activities. Therefore, the interrelationships between the tourism phenomenon and exogenous economic, social, and environmental issues have led to numerous attempts to define a destination.

The UNWTO Working Group on Destination Management (2002a) has proposed the following definition:

> A local tourism destination is a physical space in which a visitor spends at least one overnight. It includes tourism products such as support services and attractions, and tourism resources within one day's return travel time. It has physical and administrative boundaries defining its management, and images and perceptions defining its market competitiveness. Local destinations incorporate various stakeholders often including a host community, and can nest and network to form larger destinations.

This definition addresses many of the issues raised above. Many different elements are contained in this definition. Specifically, it contains a characterisation of the basic elements of a tourism system (the companies, the product, the tourist). The effort to overcome the pure geographical characteristic is evident. Apparent is also the attempt at recognising a number of factors which contribute to the standing of destination: not only firms, but also the local social community and the influence of a managing entity. These well conform to the idea of district envisaged by Becattini and his followers (see section 4.1.3).

4.2.1 The tourism destination as an industrial district

Under the pressure of competitive rivalries, as discussed in section 4.1.1, companies tend to cluster and to form agglomerations of interconnected entities in a specific geographic region. Joint efforts and cooperation are increasingly common practices in virtually every industry sector. A tourism destination, in first approximation, may be considered an example of such a cluster.

The mutually dependent attractions, services, transportations, and environmental or cultural resources emphasise the need for collaboration. This is mainly driven by the demand-side. In fact, as Gunn states (1997: 108): “A traveler is more likely to seek the great diversity and volume of services when they are located together. And businesses in such clusters benefit from local as well as travel trade.”

Destination clusters generally establish spontaneously and evolve and change over time, driven by both internal and external factors. They are not isolated entities, but open systems with complex linkages to a number of similar or diverse other systems. The development of new competitive products and services is very often done in cooperation with other ensembles, and the interface between different agglomerations allows the creation of new value (Nordin, 2003: 19):

> As a matter of fact, cluster development often becomes particularly vibrant at the intersection of clusters, where insights, skills, and technologies from different fields merge, and as a result spark new business. Moreover, the diversity of learning stimulates innovation.

The concept of tourism destination implies a systemic approach in tourism studies – an approach in which the main focus is given to the activities and the strategies to foster the development of an area pictured as a system of actors cooperating in order to supply integrated tourist products.

The literature on tourism destinations can be divided into two main strands: the first considers principally the marketing and promotional side of the concept, with the aim of pushing the competitive development of a destination (see for example Buhalis, 2000; Dwyer & Kim, 2003; Ritchie & Crouch, 2003). The second one confronts the scheme from a structural point of view, using the cluster or district lenses to describe the characteristics and functioning of such systems (Antonioli Corigliano, 1999b; Hjalager, 1999; von Friedrichs Grängsjö, 2003).

This is the approach taken by Nordin (2003), which starts from the Porterian idea of response to competitive pressure. The *four forces diamond* model is used to describe a cluster of tourism enterprises, and emphasis is given to the role of innovations. Nordin (2003) stresses the need to develop collaboration and cooperation strategies to gain a sustainable competitive advantage. This Porterian methodology, however, has been subject to a number of critiques that look especially valid when the scheme is transferred to the tourism sector. These criticisms include the lack of specificity and measurability (Martin & Sunley, 2002) and more importantly it can be noted that the model fails to recognise the influence of the public sector and of groups or associations with different interests. This is particularly true in tourism, where government interests (central or local) are highly influential (Stopford & Strange, 1991).

The complex nature of an agglomeration of industries such as tourism businesses and the difficulty of stating a delimitation of a geographical or sectoral nature makes the analysis more difficult. The tourism industry is a very heterogeneous sector, including such diverse segments as accommodation, intermediaries, transportation, marketing and others (Leiper, 1979; Mill & Morrison, 1992). Moreover, it has strong linkages to other related and supporting industries or groups, necessary to satisfy the demands of the customers, so that it can be extremely difficult to draw clear boundaries between clusters. Tourists, besides the destination attractions and services, value several other factors such as safety, security, cleanliness of the environment and traffic. People travel for immaterial cultural reasons that are beyond the logic of supply and demand. The competitive advantage of an area is increasingly related to the whole system of local actors providing a very complex final product (Keller, 1996; Urry, 2002; Vogt et al., 1993).

To achieve competitiveness and development goals, the promotion of industrial clusters is just one of many possible strategies. Even if a region has some successful clusters creating prosperity and employment, it might be worth considering that clusters go through life cycles and face decline. Therefore, strong regional economies should be able to differentiate and count on more sectors to foster their development (Storper, 1997).

On the other hand, the idea of a tourism *district* (see section 4.1.3) can be a more useful framework to interpret the phenomenon. It can be seen as an ideal model for the development of a system which encompasses a number of different enterprises focused on tourism, whose activity is able to generate wealth, occupation and wellbeing for the local community by exploiting and enhancing the local resources. As in the industrial district model, travel and tourism industries, generally of small and medium size, live in a common area and are characterised by linkages of different strength among themselves and with the local community. Moreover, the intervention of public sector institutions is essential in giving policy guidance and support and in managing a large part of the resources that are *public goods* such as culture, environment, lifestyles, and so forth. Hjalager (1999) indicates that a flourishing tourism destination shares the main characteristics of a successful industrial district:

- interdependence of the participating firms;
- cooperative competition, and trust in sustained collaboration;
- a ‘community culture’ with supportive public policies;

- a global market faced by an SME-based economy with main specialisation in one sector;
- an extended vertical interdependence with detectable numerical and functional flexibility;
- tendencies towards the establishment of supportive public and semi-public policies and institutions.

There are differences for which the *district* model needs some adaptation. First of all, the 'product' has peculiar characteristics: it is mainly a service product, whose intangibility, inseparability, heterogeneity and perishability make it rather different from 'usual' industrial goods (Vanhove, 2005). In addition, the acquisition and consumption of the tourism product is typically separated by space and time so that potential visitors are unable to fully assess product attributes prior to consumption (Burns, 1999; Cooper et al., 2005; Mill & Morrison, 1992).

The district interpretation of a tourism destination considers it as formed by two main classes of interacting components (Antonioli Corigliano, 1999a, 1999b, 2000; Capone, 2004; Lazzeretti, 2003; Stamboulis & Skayannis, 2003):

- a large endowment of resources: natural, cultural, artistic, but also artificially built, such as museums, theme parks or sport complexes;
- a network of groups of actors: economic, non economic and institutional, whose prevalent activity is providing tourism related services to visitors and travellers.

The stakeholders of a destination district include, in the tradition of the Becattinian school, not only those whose core activity is 'tourism' as it would be in a Porterian cluster, but also the local social system, the various institutional entities (local or country governments, associations etc.) and other organisations whose activity, although not directly of *touristic* nature, are deemed essential for the good functioning of the system as a tourism destination. Even the geographical delimitation can be somewhat relaxed since, as seen above, virtual groupings with entities external to the specific area can be established, thus overcoming the need for a strict physical proximity.

This enlarged definition takes into account the categorisation of the tourism supply side made by Smith (1988: 184) of *Tier 1* and *Tier 2* businesses:

> Tier 1 commodities are basically 'pure tourism.' Businesses that provide Tier 1 commodities, such as airlines, would cease to exist if there were no travel. Tier 2 commodities, on the other hand, are 'mixed' commodities. These are provided by businesses such as restaurants that serve both travelers and residents. Tier 2 businesses would continue to exist if there were no travel, but they would exist at a substantially reduced level.

And, as Capone notes (2004: 8):

> Regarding the networks of actors, in the Industrial districts literature, there are two main pillars at the base of an industrial district:
>
> - the productive organization; the system of enterprises localized and specialized that work with a flexible division of labour;
> - the social and institutional local environment.

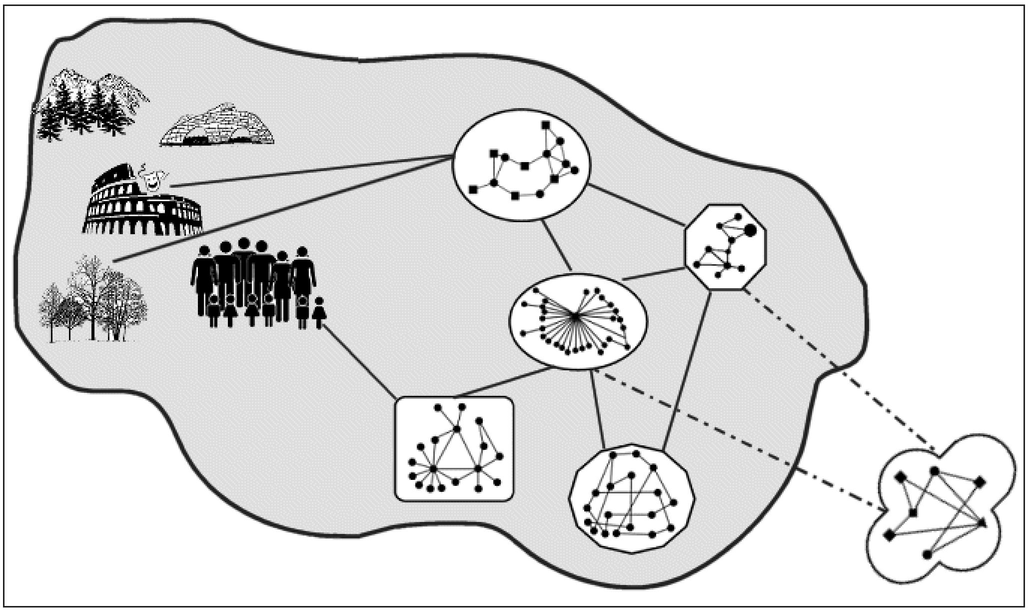


Figure 4.5 A model of a tourism district comprising clusters of diverse industries, associations and public government bodies interlinked and embedded in a 'social' territory endowed with different resources. Virtual links can tie some of the internal clusters to external groups

With this reading, a tourism destination is a complex system formed by a number of interconnected subsystems (see also McKercher, 1999). Each one of these can be a cluster of firms with a particular specialisation. All work together towards the social and economic development of a certain territory which, generally, has definite administrative boundaries, though some may be tied to external actors or groups forming virtual connections. Tourism related activities are the prevalent task of the whole system (Figure 4.5).

4.2.2 The management of a tourism destination

A tourism district, in the meaning given in the previous section, is a complex structure made of different types of organisations which are competing and collaborating in their operations. The dialectic relationships among the components of this system make it a peculiar one. Probably the most important issue, well emphasised by most parts of the literature on this topic, is the tension between different stakeholders' interests. As Buhalis (2000: 98-99) notes:

> Perhaps, the most difficult problem is ensuring the rational use of zero-priced public goods, such as landscapes, mountains, and the sea for the benefit of all stakeholders and at the same time preserving the resources for future generations. Conflicts can easily develop, especially when some (perhaps greedy) stakeholders exploit resources for short-term benefits. A compromise encompassing all these interests is extremely difficult if not impossible, but is the key to long-term success.

It is a common theme in the tourism literature that a destination management problem exists. In other words, the achievement of the benefits generated by tourism in a particular region requires a process in which some entity takes the responsibility of guiding the activities of the different stakeholders.

The literature stresses the idea that the main task of a management entity is to foster the tourism system in an area by looking at the long-term prosperity and wellbeing of the local social community, while maximising the economic profitability of the operators involved and ensuring a sustainable balance between economic benefits and socio-cultural and environmental costs. Obviously, this comprises the demand part of the sector. Therefore one added objective is the overall satisfaction of the tourists and visitors coming to the destination (Buhalis, 2000; Framke, 2002a; Ritchie & Crouch, 2003; Ritchie & Ritchie, 2002).

Management can be achieved by having some regulatory mechanism and some coordinating unit that takes care of harmonising public and private needs and actively works for the promotion and marketing of the destination as a whole: a *destination* *management organisation* (DMO). This will try to achieve the objectives stated above - coordinating a strategic planning approach and involving all the stakeholders of the destination.

For the importance of tourism activities, from both the economic and social point of view, and for the peculiar characteristics of the sector, a DMO is generally a government driven organisation and its main tool is the power of issuing policies, guidelines or rules, although with some important shadings. It has been noted, in fact, that in today's political and economic environment, the role of government in tourism has radically changed, shifting from a traditional public administration model, trying to put into operation policies for the public good, to a different model, closer to a liberistic view of markets stressing efficiency, returns and effective partnerships (Hall, 1999; Hall, 1993).

Efficient and effective destination management aspects can increase the appeal and improve the quality and value of the resources (Ritchie & Crouch, 2003). Moreover, successful management practices can contribute to the generation of satisfaction among both tourists and the local community by adopting shared marketing strategies. These include a careful monitoring of tourist activities and satisfaction levels and of the hosts' reactions to tourists. Considering these as criteria for success, rather than just tending to an increase in the numbers of tourists, allows a sustainable development of the infrastructure of tourism areas (Ryan, 1991).

Managing these processes is particularly challenging for the fragmented nature of the tourism industry and for the possible conflicts that may arise from the different views, values and attitudes held by the diverse components of the district. This, as has been noted, indicates the necessity of recognising common elements and of favouring effective transfer of information among the different stakeholders (Bramwell & Lane, 2000; Font & Ahjem, 1999; Ritchie & Ritchie, 2002). Also crucial is the capability of setting effective and clear means to evaluate the impacts (real or potential) of tourism developments and to formulate a widely agreed mission and strategic vision.

Managing the complex system called a *destination* is quite a challenging task, because it also means finding the way to 'direct' a complex system which, almost by definition, is quite unmanageable. It corresponds, essentially, to finding a control structure able to produce order by inducing certain relations among its elements with the effect of reducing the indistinguishability of the states of the system (see section 2.1). Seeking stable equilibrium relationships is considered detrimental to the development of a system; evolution and growth can only be possible in regions of the phase space at the boundary between order and chaos (Rosenhead, 1998: 3.1):

> Rather than trying to consolidate stable equilibrium, the organisation should aim to position itself in a region of bounded instability, to seek the edge of chaos. The organisation should welcome disorder as a partner, use instability positively. In this way new possible futures for the organisation will emerge, arising out of the (controlled) ferment of ideas which it should try to provoke. Instead of a perfectly planned corporate death, the released creativity leads to an organisation which continuously re-invents itself. Members of an organisation in equilibrium with its environment are locked into stable work patterns and attitudes; far from equilibrium, behaviour can be changed more easily.

For a tourism destination, as well as for other types of organisations, it is possible to adopt the idea (Stacey, 1993, 1996) that systems do not only adapt to their environments, but help to create them. Their success can come from contradiction as well as consistency. When contingency (direct and linear cause and effect relationships) loses its full validity, long-term planning is almost impossible. Success may stem from being part of a self-adapting process, rather than from an explicit 'vision'. Revolutionary as well as incremental changes may form the basis to build organisational success.

Despite these difficulties, it is still possible to manage and understand complex systems, at least at some level. Large scale behaviours might still be foreseeable if it is possible to describe the overall dynamics of the system including the presence of any attractors and their basins (see section 2.1). This can be accomplished by using a number of modelling approaches mainly based on numerical simulation models.

Once the attractors of these complex systems have been identified, it can be possible to determine whether changes in some control parameter can produce sudden shifts in behaviour, or at least establish a probability distribution for their occurrence (Hansell et al., 1997). Simulation methods prove the feasibility of this approach (Smith & Johnson, 2004).

Short-term predictions allow identification of the main evolutionary paths and small 'corrections' to the system behaviour that may be effective in avoiding undesired regimes.

Managing a complex system requires, therefore, an adaptive attitude, rather than a rigid deterministic, authoritarian style. It may require the adoption of strong rules, but it definitely needs the capability to change them dynamically, reacting in short times to all the changes that may occur in the system and in the external environment. The proposal of using *adaptive management* to deal with a system derives from the work of 1970s ecologists (Holling, 1978). It calls for an experimental path to management. The method builds on the idea of exploring alternative possibilities, implementing one or more of them, monitoring the outcomes, testing the predictions and learning which one most effectively allows the achievement of management objectives. The cycle then closes by using the results of the actions to improve knowledge and adjust subsequent management activities. Since then, it has been adopted in different situations, including tourism systems, with encouraging results (Farrell & Twining-Ward, 2004). For example, Agostinho and Teixeira de Castro (2003) analyse a Brazilian experience and bring tangible data showing that an adaptive, self-organising management system produces better performance with respect to more traditional schemes. Bankes (2002) convincingly uses computerised agent base methods to work out different scenarios in order to support the development of policies in complex and uncertain conditions.

4.2.3 Structure and evolution of a tourism destination

Although quite a large number of works have been devoted to the study of a tourism destination, there is no sound theoretical description of its structural characteristics. This is clearly due to the difficulty of defining a tourism district and all the elements that compose it and the same 'idea' of tourism. Most of the analyses on the structure of such systems can be found in the literature dealing with the attractiveness, the competitiveness and the marketing of a tourism destination. They are used as an explanatory framework to stress the importance of certain kinds of activities or processes and to discuss the benefits that may derive from them.

Figure 4.6 and Figure 4.7, for example, show some well known models (Dwyer & Kim, 2003; Heath & Wall, 1992; Ritchie & Crouch, 2003; Viganò, 2004).

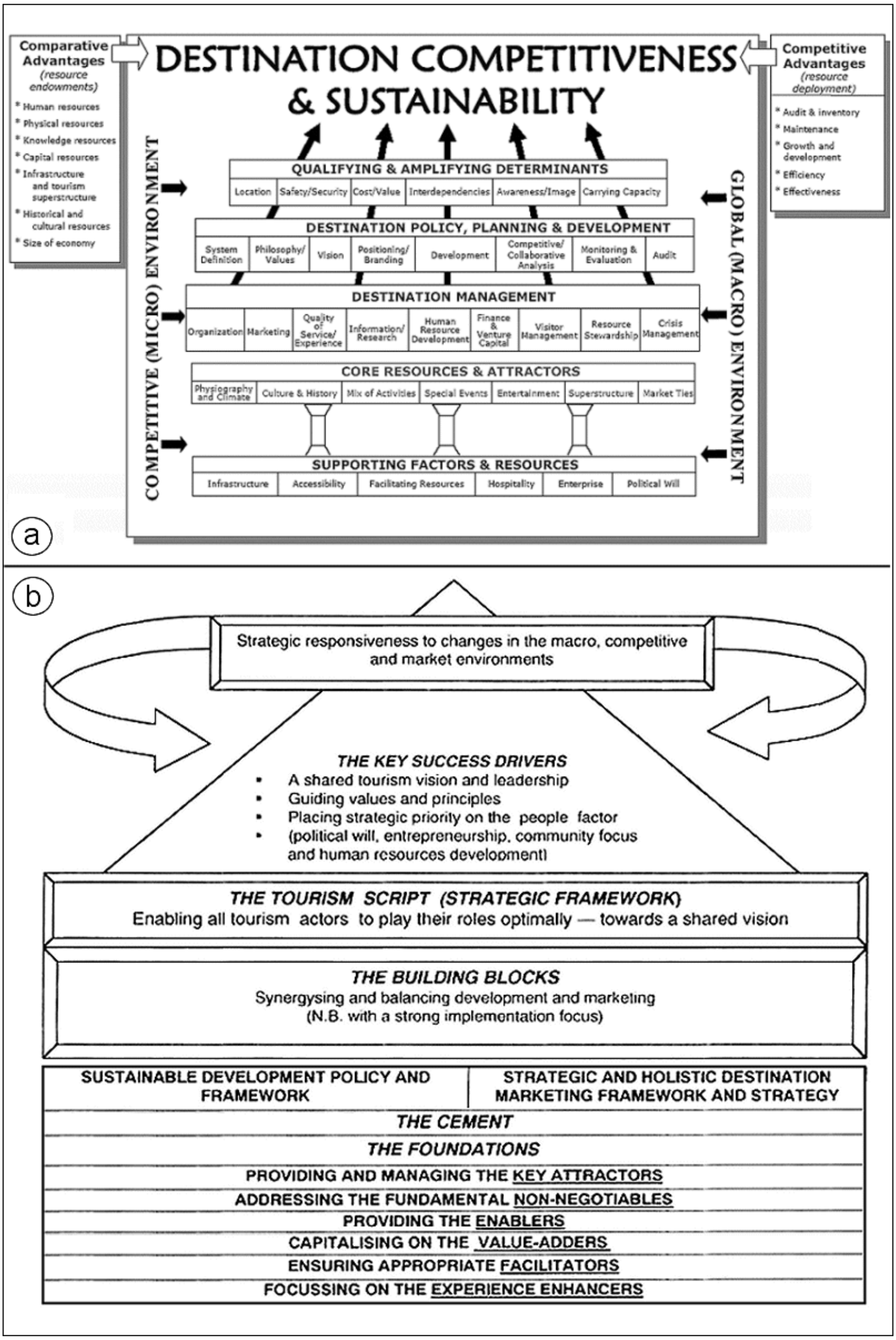


Figure 4.6 Tourism destination models (adapted from a: J. R. B. Ritchie & Crouch, 2003; b: Heath & Wall, 1992)

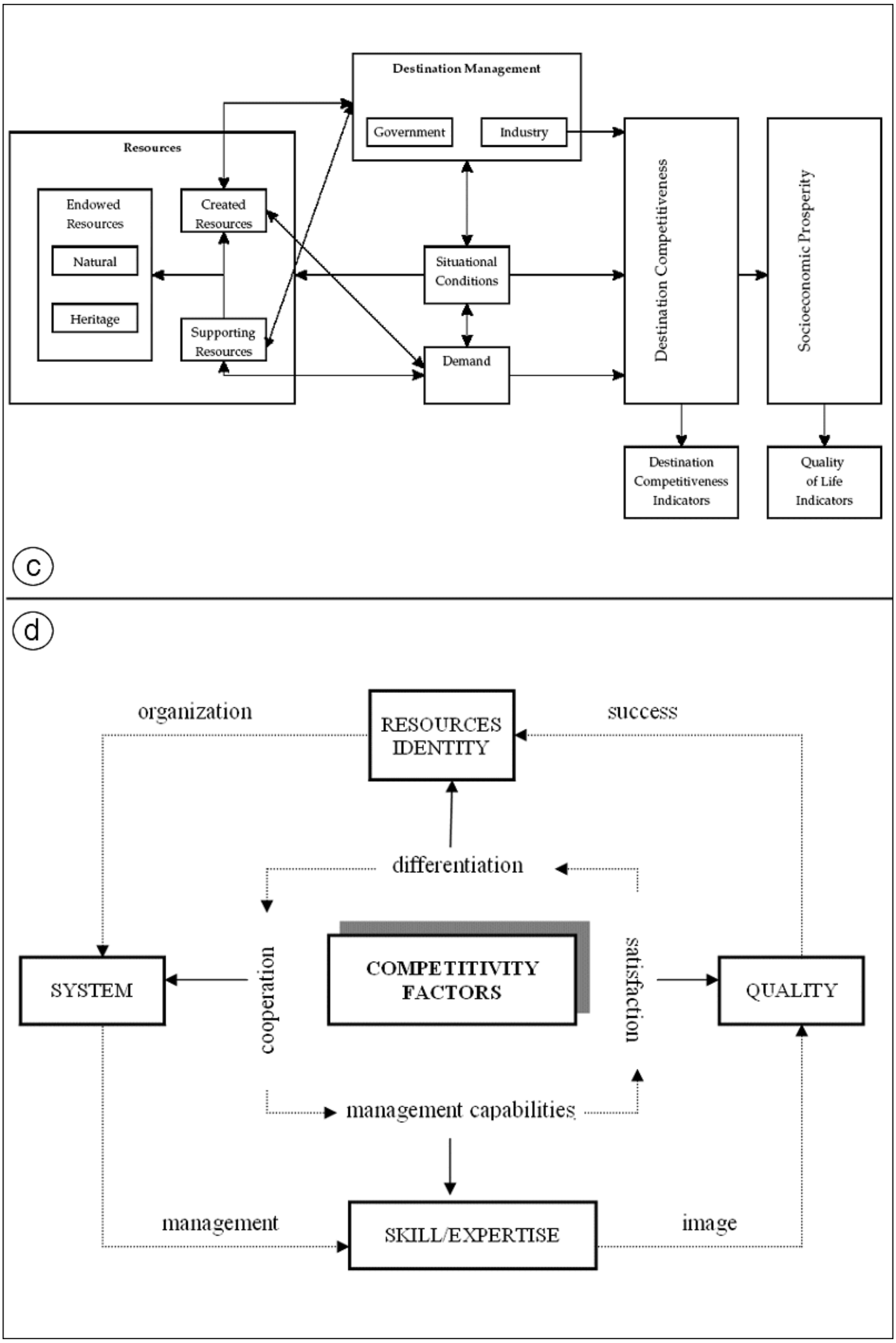


Figure 4.7 Tourism destination models (adapted from c: Dwyer & Kim, 2003; d: Viganò, 2004)

Even with different graphic styles these models look rather similar from a structural point of view. What can be noticed is the presence of a mixture of ‘physical’ (stakeholders, infrastructure, resources) and ‘conceptual’ (processes, activities, behaviours) elements. Even if nothing (or very little) is said about the structure and the type of connections that bind the different elements, one may imagine that these elements are subsystems composed of a number of stakeholders connected by a web of linkages.

Also when new approaches are invoked to better understand the functions and behaviour of a tourism destination, such as in an analysis by McKercher (1999) derived from complexity and chaos theory, the structure (Figure 4.8, see also section 2.1) is described only at a very high level, by showing a classical scheme in which structural considerations are almost absent.

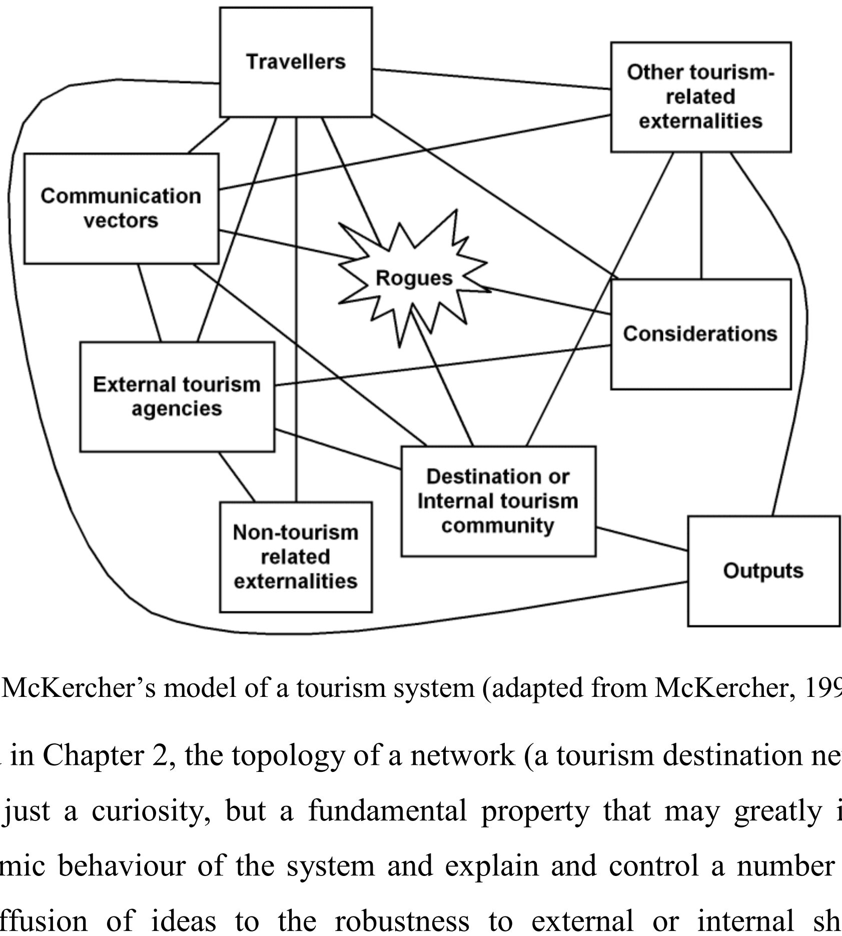


Figure 4.8 McKercher’s model of a tourism system (adapted from McKercher, 1999)

As discussed in Chapter 2, the topology of a network (a tourism destination network, in our case) is not just a curiosity, but a fundamental property that may greatly influence the overall dynamic behaviour of the system and explain and control a number of processes from the diffusion of ideas to the robustness to external or internal shocks, to the optimisation of the relationships among the network components. The networked structure of a tourism destination and its importance has been acknowledged by several authors.

Considering the set of relationships is deemed a remarkably appropriate approach to describe these systems and to give better insights into the whole industry and its coordination and organisational structures (Tremblay, 1998).

The simple existence of a network in a tourism district is not sufficient to generate effective synergies, it is the structure of such networks that is thought to be a crucial determinant (Michael, 2003). The existing theories and research on the relationships between competing and cooperating firms in a tourism destination support a confirmation of this role. In a tourism environment where many and diverse small companies operate, the overall success of the destination is more often found when firms interact more frequently both at a formal and an informal level. Furthermore, efficient information transfer and cooperation in marketing activities or in sharing the knowledge about the 'paths' their tourist segments take through the destination strongly influence a destination's success (Gnoth, 2004).

Multiple ranges of network types exist; they can be categorised according to type of organisation, configuration of interorganisational connections, degrees of formality, or level of intensity of the linkages between members. The success of such networks (in terms of economical and social benefits achieved) depends on a number of different factors: clarity of objectives; organisational structure and leadership; capabilities to manage human, financial and physical resources; and participation of the members. Most of these benefits are difficult to quantify. The evaluation of their qualitative aspects can be very complex, but these benefits are deemed important to fully understand the characteristics and functioning of social groups (Dredge, 2005; Rhodes, 2002). All things considered, the many examples studied confirm a clear relationship between the success of a destination and the structure of the network of its stakeholders. This is valid also for 'virtual' tourism networks, those that include elements not necessarily geographically close, but spread on an international basis and connected by a common vision and an efficient exchange of information and knowledge (Morrison et al., 2004). A series of cases show that:

> the networks generating the greatest range of benefits were those that had embedded a system and a culture to sustain inter-organisational learning and knowledge exchange (Morrison et al., 2004).

The role of the network structure as a mechanism for transferring information among its elements strongly affects coordination and managerial processes. According to Pavlovich (2001: 495), it:

> displays the overall properties of the group; the relational systems within the network determines who get what, through the scope and intensity of individual relationships.

The position and connectivity of the single organisation influence the consequences of its actions, and efficient information transfer fosters the capability of the network to implement self-regulation mechanisms. Elements whose actions are perceived as contrary to the general interest are cast out and put in a peripheral position (Pavlovich, 2001).

The density and intensity of ties in the network are a decisive factor for the overall behaviour of the system. Strong ties that an actor has with others within a densely linked group do not provide as much diversity of knowledge and information as 'weaker' relationships to elements or groups external to the network. The *strong* links keep the network together while the *weak* links stabilise it (Granovetter, 1973, 1983, 1985).

A structural optimisation of the relational connections, obtained as a balance of both weak and strong ties is the optimal solution for achieving strategic leverage in tourism destinations (Pavlovich, 2003b). It may enable richer and more intense information exchanges from which more comprehensive and satisfying supplier and consumer relationships have been detected (Pavlovich, 2003b).

The macro level structural properties also have a strong influence on the single actors in a tourism destination. By studying the structural optimisation of the relationships, their heterogeneity, interconnection and reciprocity, Pavlovich and Kearins (2004) show how the characteristics of *embeddedness* of a destination network favour or hold back an organisation within the system and how this affects reciprocal flows of resources and information for the development of strategic capabilities.

In some cases, different types of network arise, depending on the relative power of the participating firms: 'dominating networks', with a strong main controlling organisation and 'equal partner networks', in which no single party controls the network and close mutual relations are established. The power and position of the entrepreneur and his company at the destination can be a deciding factor in whether the company is prepared to work together with others or whether they are pushed to participate in the joint production of tourism (von Friedrichs Grängsjö, 2003).

Therefore, for these highly fragmented ensembles to be effective functioning systems, the structure of the linkages between organisations is a critical factor to achieve strategic power. The importance of cooperating in networks is especially high for small businesses, the vast majority of companies in the tourism sector. They may consider their own lack of resources as a strong motivating factor to effectively enter a cooperating network (von Friedrichs Grängsjö, 2003).

The tourism destination's structure and dynamics are, therefore, recognised to be strictly connected. What is missing is a quantitative analysis of these factors and their influences, as most of the studies have relied, so far, principally on the qualitative analysis of a number of cases (see for example Dredge, 2006; Grangsjo, 2006; Novelli et al., 2006; Pavlovich, 2003a; Pavlovich & Kearins, 2004; Pforr, 2006).

The evolution of a tourism destination

A tourism destination is not a static system. It evolves over time passing through different evolutionary phases. The analysis of the development of tourism destinations is an important theme in tourism studies. The literature on this subject has been built, basically, around the idea of a tourism area life cycle (TALC), originally proposed by Butler (1980). This model is composed by applying to the development cycle of a tourism destination the theories on the evolution of products (product life cycle model), dating from the 1950s, that were well established in consumer marketing studies by the time that Butler adapted the framework. A new product is launched, achieves acceptance and growth until competitors gain market share (Gardner, 1987). Then, innovation or repositioning is necessary to withstand declines in sales and profits. Butler applies these principles to dynamic, market-driven tourism development and suggests that successful destinations pass through a sequence of growth stages that follow the *s-shaped* logistic curve shown in Figure 4.9 (this model is similar to the one used to describe the general evolution of an industrial district, see section 4.1.1 and Figure 4.1).

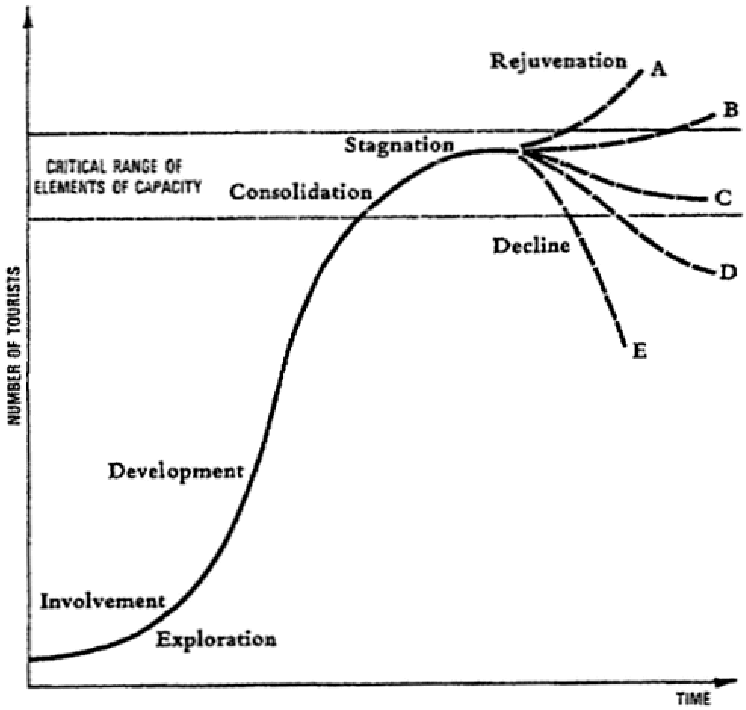


Figure 4.9 Butler's model of a tourism area life cycle (after Butler, 1980)

The model, in the tradition of the diffusion theories, uses the number of tourists as a measurable outcome and identifies seven stages:

- *exploration*: a small number of travellers independently explore a new location, either for personal adventure or to experience new cultures. At the location there is little or no tourist infrastructure;
- *involvement*: as acceptance by the locals increases, the destination becomes more popular. Travel and accommodation facilities are improved, and there local investment in tourism related services and advertising begins;
- *development*: the local community becomes involved in promotion, attracting more visitors, and the area turns into an established tourist destination, with a defined market. Visitors outnumber residents, external investment leads to loss of local control, man-made attractions emerge to replace natural or cultural ones;
- *consolidation*: tourism is an important economic and social activity affecting the traditional economies and lifestyles. Land is given over to resort building, without a simultaneous benefit in increased wealth or jobs. Resentment by the local population may occur. This stage is characterised by slowing tourist growth rates and extensive advertising to overcome seasonality and develop new markets;

- *stagnation*: increased local opposition to tourism and a growing awareness of environmental, social and economic problems results in opposition to further growth. Capacity limits are reached and a destination image unconnected with the environment makes the area no longer fashionable;
- *rejuvenation*: a secondary growth burst is initiated by some kind of renewal. New attractions replace the original facilities. New tourists may be of different socio-economic groups or different age groups than the original;
- *decline*: if nothing is done to react to a stagnation period, a final decline of the destination occurs, with partial or total abandonment of tourism as an economic activity, organised and supported by the community. Different degrees of 'decline' or 'rejuvenation' may give different patterns as outcomes (i.e. the alternative paths A to E in Figure 4.9).

The model has been applied to many areas and in many cases has proved to be quite effective in describing or explaining the tourism history of a district. It has been commented on, improved, developed and criticised numerous times (Butler, 2005a, 2005b). Many authors have used it to test its adequacy in specific situations (Agarwal, 1997, 2002; Baum, 1998; Cooper & Jackson, 1989).

Some of the model's criticisms have been based on expectations (mostly unrealistic) of explanatory and predictive power, or on the concern that it emphasises factors which are specific to individual destinations rather than those common or shared with other places (Prosser, 1995). Some argue that destinations are more complex than products with a specific life cycle and more similar to living ecosystems evolving through time (McKercher, 2005). Nonetheless, Butler's model, even if in some cases it has proved unable to fit the situation, is considered to have an important role in providing useful frameworks for the description of a tourism destination and in emphasising the need for proactive strategic planning processes (Hovinen, 2002).

To overcome at least some of the limitations discussed above, others, such as Lozano et al. (2005), have proposed a more rigorous formalisation by linking the TALC model to economic growth theory; their numerical simulation of an environmental growth model for an economy specialising in tourism leads to the same general pattern of evolution. The most important critique is that the TALC model is founded only on a number of case studies and, essentially, it lacks a standardised approach and a formal base. The absence of uniform measures and accurate quantification make it difficult to use for effective planning purposes (Getz, 1992). As a consequence, conceptual and empirical difficulties have been found in defining the stages, the predictability of the progression, and complications involving multiple products with multiple cycles in a given destination or in geographical subdivisions of the area (Agarwal, 1994, 1997).

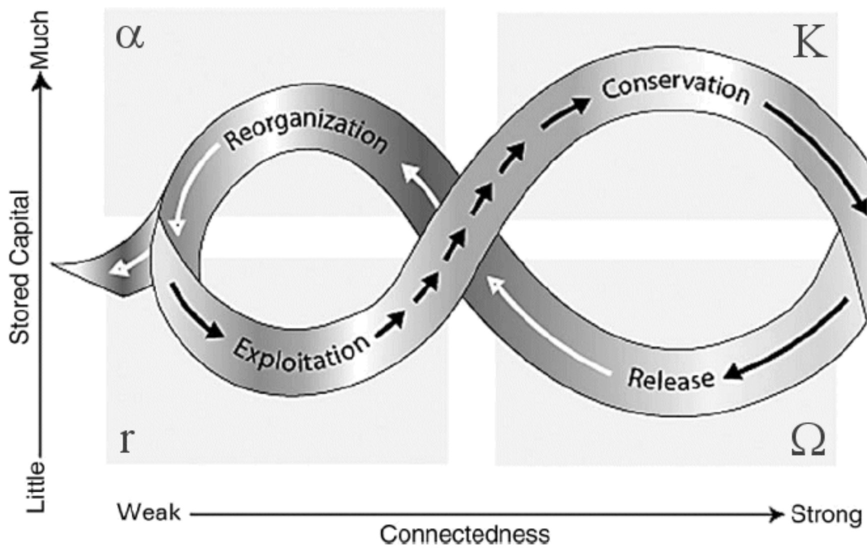


Figure 4.10 Life cycle of an adaptive system (after Holling & Gunderson, 2002)

The recognition of signatures of complexity or chaos in a tourism system has also led to other extensions and interpretations of the TALC model. Russell and Faulkner (2004), for example, combine the two approaches to explain the role of entrepreneurial forces in a tourism destination as elements of turbulence, change, and unpredictability. Farrell and Twining-Ward (2004), recognise in Butler's life cycle a portion of the wider adaptive life cycle introduced by Holling (Holling, 1986; Holling & Gunderson, 2002) to describe complex adaptive ecological systems. In this interpretation, the front part of the cycle (Figure 4.10) closely follows the behaviour of a destination as described by Butler. Having arrived at the *saturating stagnation* phase, the system, left to itself, and without effective management tends to be over-connected, with a too rigid internal control and extremely vulnerable (Farrell & Twining-Ward, 2004). In these conditions most of the internal resources are released, the networking characteristics are relaxed and the system goes into a self-reorganisation.

Adaptive cycles take place at a range of different scales; they interact and can influence one another. The self-reorganisation phase, as for many complex adaptive systems, may lead to the formation of new hierarchical structures. The new level temporarily reduces the complexity of the system which can start a new evolutionary phase.

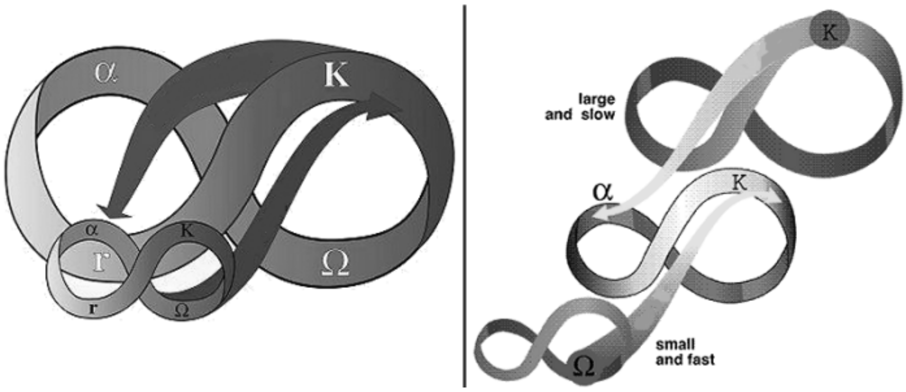


Figure 4.11 Influence of different scales (left) and hierarchical evolution of adaptive life cycles (adapted from Holling & Gunderson, 2002)

This *panarchy* (Holling & Gunderson, 2002) is also proposed as a general framework in which tourism can be analysed (Figure 4.12).

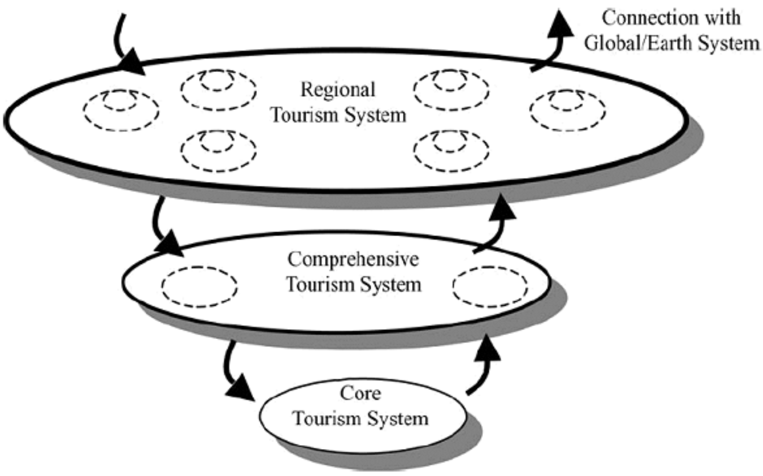


Figure 4.12 The tourism panarchy (after Farrell & Twining-Ward, 2004)

Even if in a purely qualitative way, with no support by quantitative measures, the idea is proposed as a faithful representation of a tourism system with its many diverse and complex relationships and with its different structures, mostly hierarchically nested. Farrell and Twining-Ward state (2004: 279):

> … understanding the destination in terms of the panarchy of systems adds further substance and meaning, and the possibility of discovering how destination character emerges, how it may be kept or even enhanced.

4.2.4 The contribution of network analysis

Most of the results and the models on the structure of an industrial cluster or a tourism destination discussed above are derived by the analysis, mainly qualitative, of single cases. The importance of network structures has, however, pushed at least part of the literature to the usage of the techniques initially tuned by sociologists and, more recently, by so-called 'network science'. Social network analysis has been used several times to explore how networks are formed in the interaction between different players or, taking the network structure as given, to analyse economic phenomena that are constrained by the structure itself.

These studies have given stronger confirmation to the idea that the structural elements of regional networks significantly affect interorganisational collaboration (Martinez-Fernandez, 2001). Moreover, the quantitative analysis has emphasised the fact that not all the elements of a network share the same benefits from participation in a cluster. The performance of firms is associated with individual knowledge contributions and with their position in the knowledge network (Giuliani, 2006). Different structures provide different outcomes, and the role of larger companies may strongly affect the functioning and the success of the whole group or of single actors in the group.

Formal analysis techniques may thus uncover patterns that a casual inspection may not be able to highlight. At the same time, the structure of the interactions with institutional supporting infrastructures helps in understanding the variability of agglomerations that could look apparently similar. Network analysis stresses the fact that what really matters is the existence of ties connecting networks of small firms to larger firms and how they can, in turn, connect them to global partners and suppliers. The issue is no longer whether large or small firms will succeed, but how regional economics link firms of various sizes and competencies together, and with what results (Castilla et al., 2000). On a larger scale, then, inter-regional and international transactions may have a greater importance than the internal connections (Britton, 2003).

Social network analysis also plays an important role in determining the effectiveness of the typology of the linkages. It has been shown, for example, that the claim that regional economic success is a function of strong social embeddedness in local networks is not always confirmed, but looser couplings yield better results in many cases (Staber, 2001).

In addition to its contribution to theory, the understanding of the dynamics underlying industrial clusters can also inform public policy. The analysis techniques show, in fact, that industries cluster to have smoother access to information and resources (Sorenson, 2003). These methods also allow us to determine how, in what terms and with what limitations leading firms in a district act as *knowledge gatekeepers* (Morrison, 2004). It may happen, for example, that public interventions supporting individual leaders, intended as engines of growth for local systems, produce negative effects tending to reinforce internal asymmetries and boost internal conflicts, in particular between large and small size companies.

More recent ideas have also been tested and applied to the structure of industrial networks leading to the conclusion that flexible specialisation in a large-scale industrial district is different from a simple 'small world' phenomenon, but has to consider more complex spatial clusterings (Nakano, 2004).

Most works deal with network processes and principally with diffusion mechanisms (Coulon, 2005), again highlighting the influence of the topology, but also highlighting the importance of the internal performance of the single actors. Although the network analysis methods are quite 'old' and tourism is a network business, little has been done so far to apply these techniques to the study of the tourism sector. Some recent contributions, however, show the usefulness and the effectiveness of this approach.

The networking of single individuals and the diversity of their ties play significant roles in the decision routines and paths in the tourism industry. Furthermore, it has been noticed that there are differences in the use of efficient ties, in terms of strong and weak ties. They change depending on specific tasks, and successful actors are able to optimise their composition (Pesämaa & Skurla, 2003).

This is a reconfirmation of results found in completely different disciplines such as biochemistry (Csermely, 2004). These studies validate the importance of a balanced and heterogeneous composition of links as a factor fostering the success of a tourism network (see Pavlovich, 2004, even if in that case available data were analysed only in a qualitative way with interpretative research methods).

Policy issues are studied by Pforr (2001; 2002; 2006) in a series of papers dealing with the Australian Northern Territories. In these, he shows what usefulness the policy network approach has in describing, analysing, and explaining the dynamics of the tourism policy realm. Quantitative results are used to map the intensity and the density of sets of relations and to measure the relative consideration of the different stakeholders. Pforr also shows that measuring link characteristics leads to effectively examining the role of different actors in the network. He claims that it is possible to see how actors deemed as relevant for strategic planning processes turned out to be more or less insignificant for plan formulation processes.

The formal network approach, besides highlighting the differences in the organisation of tourism activities in different destinations, proves useful in stressing the necessity of collaboration and cooperation, typically lacking in this sector (Bramwell & Lane, 2000). The emphasis on the formation of a value creation system through a balanced set of relationships substantiate this, and offers guidance to policy makers and management organisations (Scott et al., 2008a).

Not only political and strategic tourism issues benefit from network analysis, but also more ‘practical’ problems. A recent paper (Shih, 2006) shows, in fact, how to use these methods to revise the organisation of tourist facilities and services in particular destinations by mapping and measuring the structural characteristics of routes taken by tourists in multi-destination trips.

In conclusion, sound quantitative methods for the analysis of networks are of great importance for the study of a tourism destination, not only as a fascinating intellectual problem, but also as a means to improve abilities and capabilities to understand the functioning mechanisms of a tourism destination in order to manage it effectively and efficiently.

4.3 Conclusions: a theoretical framework for the thesis

Chapters 2, 3 and 4 have presented a number of models and theories that form the broad theoretical basis in which the work discussed in this thesis is conducted.

The object of the present study is a tourism destination. It is, with a systemic approach, seen as an integrated ensemble of different elements: social, economical and technical. Consistent with the industrial district models, it includes a number of different interconnected companies, exercising tourism activities at different levels, and public resources.

This ensemble possesses a structure and works towards the achievement of defined (although sometimes not explicitly stated) economic and social objectives. The relationships among the different components are dynamic non-linear relationships; therefore, from a broader point of view a tourism destination is a *complex adaptive system*.

The general framework of complexity science gives a starting point for the study of a tourism destination. Within this framework, the structure of a system can be described in terms of its components and of the linkages that connect them. Moreover, among the many possible approaches, the study of the topological structure is able to give insights on the functioning of the system both from a static and a dynamic point of view. Therefore, it is possible to model a tourism destination as a complex network and use the ideas, the concepts and the techniques of network science to study its topology and its evolution over time. In fact, the topology of a network, as a predictable property, may greatly affect the overall dynamic behaviour and explain and control a number of processes from the diffusion of ideas to the optimisation of the relationships among the network components. The conceptual framework thus described belongs to the area sketched in Figure 4.13.

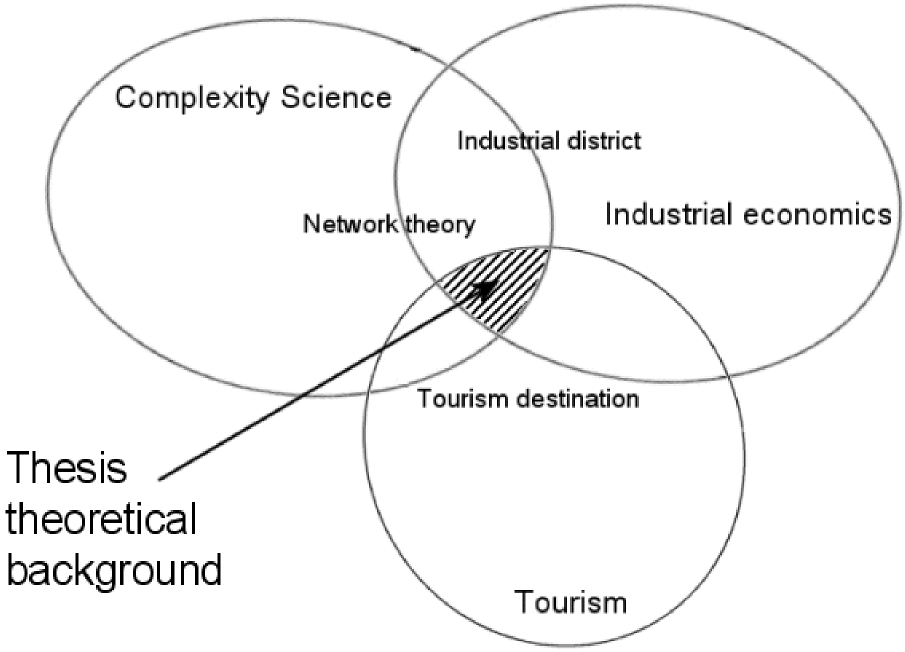


Figure 4.13 Schematic diagram of the theoretical background for the thesis

Apart from the specific research questions (see Chapter 1 and the next chapter), a broader objective of this work is that of testing the validity of this general framework and of setting and refining a methodological path for the analysis of a tourism destination. The implications that can be derived, in terms of capabilities of understanding the general behaviour and the dynamic evolution of a destination, may give tourism organisation managers a strong leverage to improve the flow of information and to target opportunities where this flow may have the most impact on business activities.

Moreover, a better comprehension of the nature and evolution of a tourism district may support public authorities, through efficient policy formation and diffusion, in developing the infrastructure and maintaining environmental quality and sustainability.

5 Methods

The previous chapters have reviewed the literature relevant to this thesis. This chapter gives an account of the methods used in this study, the techniques adopted for the data collection and the issues that may arise from their use. This constitutes an important link to the research question (Creswell, 2003; Punch, 2005) and gives credibility to the results obtained and reported in the next chapter (Chapter 6).

The object of this study is a *tourism destination* (or district) which is considered to be formed by an interconnected ensemble of organisations - a network. The research question this study answers (already stated in the introduction) is:

How do tourism destination networks evolve over time?

The development path of a tourism district is linked to its economic development and to the organisational forms the destination acquires along with those of the bodies that govern it (Pearce, 1989, 1992). The implications for both stakeholders and destination management functions have been highlighted in numerous contributions (see, for example, the extensive review in Lagiewski, 2005). The objective here is to account for this evolutionary path from a structural viewpoint. Structure, as stated in Chapter 1, refers to the topology of the network formed by the stakeholders of a tourism destination and the ties that connect them. The analysis also takes into consideration the possible relations between the dynamic evolution of the system and well-established models of the evolution of a tourism destination and their variations (Butler, 2005a, 2005b).

To answer this question, the first step is to investigate the characteristics of a tourism destination from a network theory point of view. Therefore, the following question must be answered:

RQ1: What is the structure of the tourism destination network?

This can be done by enumerating the different system components and by studying the topology of their relationships (see Chapters 2 and 3).

Then, as discussed previously, it is possible to discriminate which model is better suited to explain the evolution (to date) of a network which has, at a certain time, particular characteristics, answering then the question:

RQ2: How can we build an evolutionary model that reproduces the present structure?

This model and knowledge of the network's structure then allows two interesting problems to be examined:

RQ3: How does the network structure affect dynamic processes?

and

RQ4: How can the network be optimised?

In other words, how can we use analytical or simulation methods to maximise the efficiency or the speed of a dynamic process? If we know the topology of a tourism destination's organisational network and we measure the efficiency of this structure with respect to some process (the diffusion of a policy, or the formation of an opinion on a specific issue), what are the analytical tools we may have to maximise the efficiency of this process? Or, if we consider the cost of information transfer activities, how can we minimise them? This can be particularly important for phenomena such as the diffusion of information, knowledge or policy messages. It is well known, in fact, that a good efficiency in these processes has an important meaning for the overall functioning of the destination system or for its competitiveness (see for example Argote & Ingram, 2000; Porter, 1998).

All this, combined with the knowledge of the main events and the main accidents of the destination history, will put us in a position to answer our final sub-question:

RQ5: How does this model relate to the history of the destination?

This investigation will thus help to provide forecasts for future destination behaviour scenarios and inform the adoption of policy measures.

As stated in the introduction, this thesis sets forth a number of assumptions that bound the scope of the research. These are examined in section 5.6.

This study could be conducted by simulating numerically a destination network (Albert & Barabási, 2002; Dorogovtsev & Mendes, 2002; Newman, 2003b). However, in the tradition of tourism studies, the approach taken has been to consider a real case. A study conducted on a *real* destination provides interesting and useful results: for example, by allowing comparison of the outcomes of the 'experimental simulations' with real situations or, in case a simulation study is needed to analyse possible different configurations, by posing constraints, derived from actual conditions, to the number and type of simulations that could be implemented.

A case study approach also provides a field test of the set of methods proposed, thus providing an answer to the wider problem which underlies this whole thesis: the need for new and, possibly, more rigorous approaches to the study of a tourism system as discussed by Farrell and Twining-Ward (2004: 276):

> It is frequently acknowledged that tourism study is lacking in substantial theory of its own […] and has failed to capitalize on progress made in other disciplines. Consequently, as a field of study it appears isolated and research and teaching appear to have grave shortcomings attributable to its multidisciplinary history, organization, and relations with other fields that should inform the study.

5.1 The research path

The rest of this chapter describes the methods used for the analysis. Most of these come from the research tradition established in statistical physics. The epistemological attitude of a physicist is first described, with a discussion on the usage of 'classical' scientific method for a case such as the one presented in this thesis. The idea of considering a tourism destination as a complex network to which these investigation methods are applied raises the issue, discussed next, of the role of analogy in such studies. Finally, a justification of the possibility of approaching a socio-economic system with physics concepts is examined. After this discussion of the main epistemological approach to the present work, the first topic examined is the one concerning the selection of a suitable case to analyse. The choice of Elba Island, a well-known Italian tourism destination, is justified and the collection of the data needed is discussed. This includes the identification of the actors of the tourism network and the enumeration of the linkages connecting them. The effects of sampling of actors are taken into account, as well as the influence that an incomplete list of network elements may have on the possible outcomes.

Given that data can be collected, a network graph may then be built. This model of the system will then be analysed using the procedures shown in Chapter 2. The objective is to define and measure the main topological characteristics of the network model (RQ1 and RQ2[10]). These measures are compared with those obtained for a synthetic graph formed by the same number of nodes and links, but with a random distribution of the connections. The topology of the network affects the dynamic processes that may occur such as information spreading, robustness with respect to external and/or internal crises, fragility to structural modifications, and cascading failures. Simulation algorithms applicable to this type of system will then be implemented, taking into account available graph theoretic computational methods and the principles of statistical mechanics. An optimisation process, with respect to efficiency in information or knowledge diffusion, allows assessment of the behaviour of the system and the advantages of the optimisation (RQ3 and RQ4).

The definition of an appropriate growth model for a given network can be used to predict its extent and structure at later times. The evolution of the system is inferred by applying the known models (see Chapters 2 and 3) and by choosing those able to render the present structure. The history of the tourism destination, represented, as it is customarily done (see for example Butler, 2005a, 2005b), by the time series of tourist arrivals or overnight stays, is analysed in terms of network evolution. An investigation of the structure of the system at some past period in time, obtained with the analysis of historical data, is used to further confirm the evolutionary model previously identified (RQ2 and RQ5).

The wide diffusion of the Internet may lead us to think of a correlation between the topology of the WWW and the one of the tourism destination. This possible relationship

[10] Reference is to the research questions listed in the introduction to this chapter (Chapter 5)

will be verified. By comparing the 'real' network with its 'virtual' counterpart it is possible to highlight the differences, if any, and to examine the efficiency and effectiveness of the investments that tourism operators make in the field of information and communication technologies. This allows definition of one more question to investigate:

RQ6: How do the real and technological networks compare in a tourism destination?

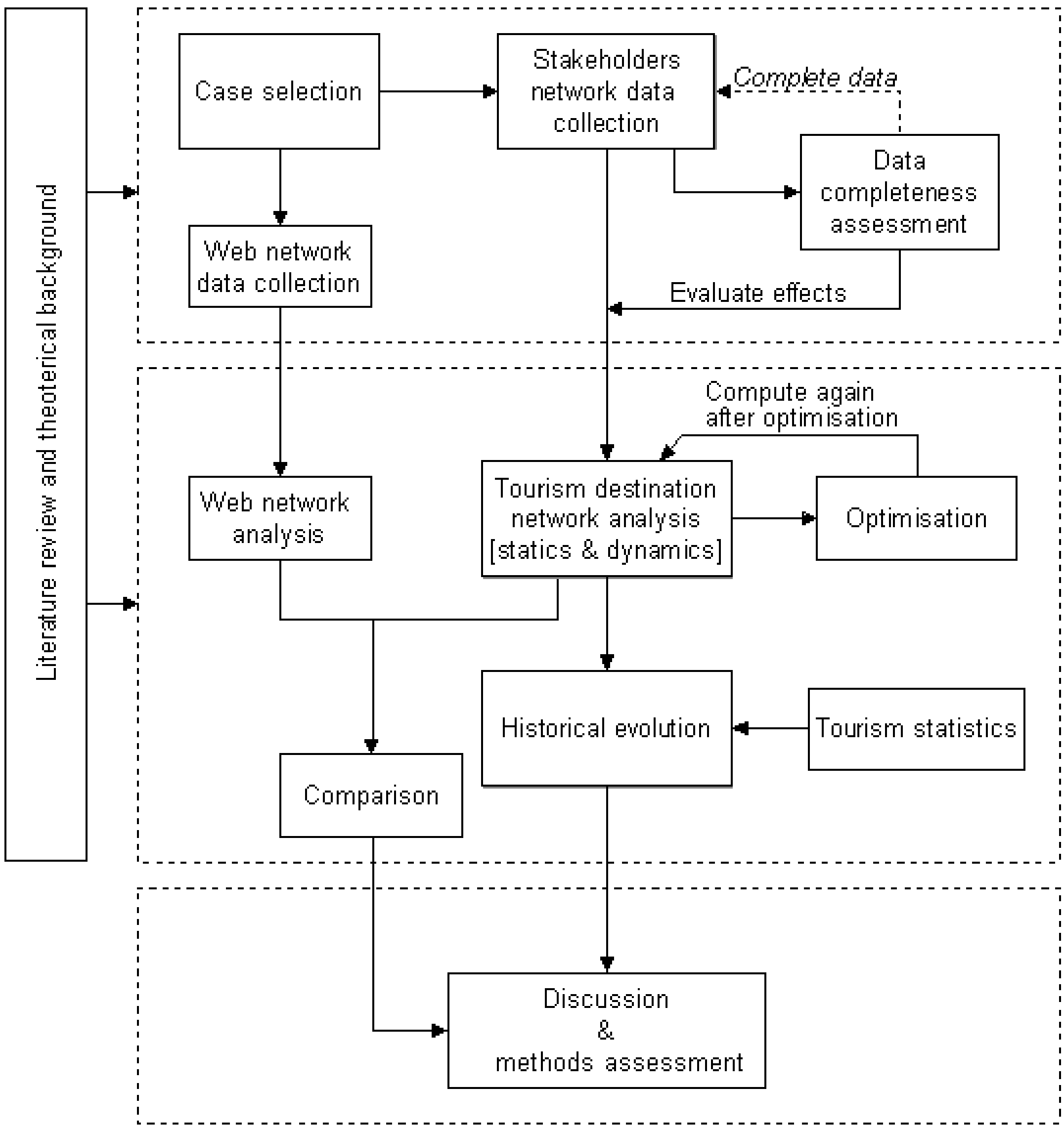


Figure 5.1 A schematic depiction of the methodology used in this work

A schematic depiction of the methodology used in this work is shown in Figure 5.1. Three main blocks can be recognised: the data collection; the network analysis and comparison of the social (tourism stakeholders) and technological (websites) networks; and the final discussion on the methods used.

5.2 The approach to this work: an epistemological digression

This section contains a short discussion of some epistemological issues connected with the present thesis. The section contains a brief account of the position a physicist generally adopts in pursuing research objectives, a discussion of the concept of analogy, central to this work and a justification for the application of physics concepts to a socio-economic system such as a tourism destination.

Starting from an initial conjecture of correspondence between a complex dynamic network and a tourism destination district, this work sets up a research program to build a methodological approach to the object of study. In the tradition of the 'scientific method' employed in an applied field, well represented by Dubin's ideas on theory building (Dubin, 1978; Lynham, 2002; Torraco, 2002), a general framework describing the initial assumptions has been set (see Chapters 2 and 3). These conjectures are verified by collecting suitable data and by using the instruments and procedures described in the rest of this chapter. After that, the original model will be revised and adjusted based on the results obtained.

5.2.1 The epistemology of a physicist

The approach of a researcher to a field of study is undoubtedly influenced by his own convictions and viewpoints. Philosophical and epistemological beliefs have always played a crucial role in the history of science. In many cases they have strongly affected the development of ideas and knowledge.

In pursuing a research idea Weinberg (1994: 167) states:

> Physicists do of course carry around with them a working philosophy. For most of us, it is a rough-and-ready realism, a belief in the objective reality of the ingredients of our scientific theories. But this has been learned through the experience of scientific research and rarely from the teachings of philosophers.

The epistemological positions of researchers are seldom defined fully or coherently. Probably the best explanation of a physicist's attitude is the one given by Albert Einstein (1949: 684) in which he represents the unsystematic and incoherent (at times) approach. In essence, what Weinberg and Einstein (a few among many, including the author of this thesis) maintain is that, even if recognising the importance of an 'epistemological scheme' in which the work is embedded, a researcher avoids considering this as a too strict guideline. Led by the activities he performs for finding an explanation to the phenomena under study, researcher's methods may also look incoherent to a 'systematic epistemologist'. The epistemological interpretation, for a physicist, can be postponed until a clear framework (models and theories) is established instead of adopting it as a guidance throughout his or her work.

Interdisciplinarity is one of the main motivations of this work, which may also be seen as an example of the tendency physicists have to "invade other people's disciplines" as Duncan Watts (2003: 61) says. In recent years, a number of ideas, concepts and techniques typical of physics have been applied to different fields such as biology, economics and sociology. Some of the results of this work are quite interesting and have helped in improving the general understanding of these areas. In particular, in economics, physicists are having an increasingly important role, so much so that a new discipline has been born: *econophysics* (Mantegna & Stanley, 2000).

The basic idea of applying physics concepts in other fields of study is far from new (see for example Mirowski, 1989) and instead may be dated back to the birth of what we today call physics. Examples include Hobbes's *The Leviathan* (1651) which attempts to use the Galilean laws of motion to derive an ideal configuration of the society, or David Hume's *An Enquiry Concerning Human Understanding* (1748) which set out to build up a science of mind and society in the image of Newton's theories of the solar system. Today, however, the scientific community takes a different approach. Instead of mechanically applying physical laws and theorems to the economy, the approach is to make the best possible use of the background knowledge to build specific methods aimed at predicting the performance of a system (a society or an economy, in this case). This is done by deriving this prediction from a set of initial conditions and a series of mathematical models. In such a way we hope to be able to obtain a reasonable account of the system, to determine what choices lead to what consequences and to make reasonable forecasts of its future behaviour.

The awe-inspiring words used by Ettore Majorana in closing his contribution (Majorana, 1942) to this way of reasoning are probably the best testimony. Referring to the possibility of using the then newly developed quantum mechanical statistical methods in the area of social and human sciences, he says (Mantegna, 2005: 140):

> From a scientific point of view nothing prevents one from considering that an equally simple, invisible and unpredictable vital fact could be found at the origin of human events. If this is so, as we believe it is, the statistical laws of social sciences increase their function. Their function is not only of empirically establishing the resultant of a great number of unknown causes, but, above all, it is to provide an immediate and concrete evidence of reality. The interpretation of this evidence requires a special skill, which is an important support of the art of government.

The 'method' is the one set by thinkers such as Bacon (1561-1626), Galileo (1564-1642), Descartes (1596-1650) and Newton (1642-1727). It is the method that, although with many variations, amendments, modifications and interpretations, has allowed generations of scientists to come to our present understanding of many natural phenomena.

Even if questioned, or questionable, from an epistemological point of view, scientists, and physicists in particular, think that their own job is to formulate theories that can be exposed to empirical testing and rejected if they fail the tests (Popper, 1959). These theories fit in a general framework (the paradigms discussed by Kuhn, 1962) that is kept until better alternatives are found. In this framework, a research program is set up (Lakatos, 1974). It consists of a hard core that states goals, issues and enquiries - in essence what the program is about. These are not open to falsification within the system, but are rather shielded by a 'protective belt' of statements that are open to verification. When falsified, these are replaced by variations that are also falsifiable, but which continue to protect the hard core ideas. Thus a research program provides a framework within which research can be undertaken with constant reference to the presumed first principles which are shared by those involved in the program, and without continually defending these first principles.

Hard core paradigms are changed only after the empirical evidence shows that it is impossible to defend them any more. In some cases, however, even in the absence of strong empirical evidence conceptual problems may arise (i.e. problems of consistency, either internal or with dominant traditions in other fields). Then, as Laudan (1977; 1981; 1984) maintains, we accept the 'research tradition' that is capable of solving the most problems, thus changing the dominant paradigms. In his view (Laudan, 1977), the framework in which a scientist works, the *research tradition*, has two basic aspects: it defines the characteristics considered problematic (mainly quantification, but also conceptual aspects) and it defines the method (or the methods) to be used to deal with these problems. Obviously, these definitions have a significant impact on how the research is conducted. Different traditions involve different hypotheses and different measurement models. These, in turn, affect the choice of issues to be studied, the approach to data analysis, the results and the implications drawn from the research. We select (or should select) the problems to investigate based on their consistency level and their possible empirical implementation. Scientific progress occurs when it is possible to increase the range of applicability of existing theories and their capacity of solving problems or if simpler and more elegant solutions can be found to more effectively describe natural phenomena. This follows Ockham's (1284-1347) saying: "Pluralitas non est ponenda sine necessitate" (Entities should not be multiplied unnecessarily).

Whether the basic methodological instruments are heavily quantitative or rely upon qualitative techniques, this framework is the one that characterises any *science* (or discipline willing to be recognised as such), be it natural, social or philosophical.

5.2.2 The use of analogy

A tourism destination is considered here to be a *complex adaptive non-linear system*. It is represented by a network of interconnected elements, and a description of this system is sought by adopting the approach described so far. In this attempt, methods and techniques originally developed for different purposes are employed. What guides this endeavour is, basically, an analogical line of reasoning.

Analogies have been widely used in a number of different disciplines. In physics the method is well known and has a long tradition. One of the best examples is probably the work by James Clerk Maxwell (Turner, 1955). In his writings (for example Maxwell, 1870: 10) he has stated several times the effectiveness of a formal analogy between two systems as a way of increasing and deepening the knowledge of them.

Whenever a new object of study emerges, a physicist tends to use analogical methods. Olson (1943: 1) expresses this idea as follows:

> Analogies are useful when it is desired to compare an unfamiliar system with one that is better known. The relations and actions are more easily visualized, the mathematics more readily applied and the analytical solutions more readily obtained in the familiar system. Analogies make it possible to extend the line of reasoning into unexplored fields.

Metaphors and analogies arise as rhetorical figures. They are usually employed to clarify certain situations. Hundreds of different definitions can be found for these terms and dozens of philosophers have debated the subject, sometimes with long discussions on subtle differences (see for example Geymonat, 1972), so that it may be legitimate to give a personal interpretation of the whole matter.

A metaphor is a very wide term suggesting a likeness between two phenomena or two objects. It is widely employed to exemplify various situations and to ease an intuitive comprehension of diverse concepts. This use for 'aesthetic reasons' does not always bring to practical descriptions or deep understandings (Gentner, 1982; Gentner & Jeziorski, 1993). When a similarity between different phenomena may be established more firmly, it can be suggested that there exists some common underlying law or principle. This may be especially true where such a similarity exists between the functions of elements in different systems or between their structures. The usefulness of this approach depends on whether the consequences that can be drawn can be tested or observed and on the correctness of the theoretical framework in which the analogies are set (Daniel, 1955; Gentner, 1983).

If structural relations can be reproduced in a simplified form in a distinct (and better known) environment, a mathematical model can be assembled. The effectiveness of this procedure has been proved in innumerable cases and in various disciplines (Gentner, 2002; Krieger, 2005; Pask, 2003; Wigner, 1960). From an epistemological point of view, although a concept can be taken with caution to avoid potential abuses (Daniel, 1955), it has been claimed that theories (stated as a set of postulates) showing not even a formal analogy to some already existing system of abstract relations, would provide no means to understand how the theory could be applied to concrete problems (Nagel, 1961).

In an ideal scale of possibilities, the concept of homology comes next. Derived from the biological sciences, it has the meaning of a "similarity, often attributable to common origin" or "likeness in structure between parts of different organisms due to evolutionary differentiation from the same or a corresponding part of a remote ancestor" (Merriam-Webster, 2005). In other words it indicates a strong likeness (but not identicalness) in structure or function between parts of different structures or systems. From this it derives that homology implies both substantial and formal analogy (UIA, 2006).

The difference between analogy and homology, however, out of their original 'biological' meanings (and sometimes even there), does not look very clear (Brigandt, 2004; Fitch, 2000; Handford, 1999). The two terms are often confused and the definitions and the boundaries between the two concepts are not completely clear. Here, a homology between two different systems is considered to require a firmly set and empirically verified correspondence of properties and mechanisms. The 'simple' analogy, limited to a single case, is used in this work is a starting point. Further confirmations and further investigations will be needed to assess the capacity of the research path proposed here to 'solve problems' (in Laudan's sense) and to demonstrate its usefulness and its usability.

Moreover, the use of the analogy has a catalysing function. As has happened many times in the history of science (in physics, for example, with scientists such as Faraday, Coulomb, Helmoltz, Maxwell), it serves as initial help in the development of new disciplinary fields. Known schemes, models and criteria may initially help in finding one's bearings in the vast quantity of data, evidence, phenomena and to start organising all these into an organic set. With this contribution, it is hoped, tourism studies may acquire that status of a 'legitimate' discipline that today is lacking, or is not fully recognised (Davies, 2003; Leiper, 2000b; Tribe, 1997).

5.2.3 Physics and socio-economic systems

One more consideration is in order concerning the epistemological implications of using the laws and methods of physics applied to social systems. This is an area where these is little relevant prior literature. While a variety of papers deal with the epistemological issues of both natural and social sciences, comparing the attitudes and positions of researchers with regard to the similarities and differences of approaches and methodologies (see for example Durlauf, 1997; van Gigch, 2002a, 2002b), the specific problem of applicability of a 'physical' approach to social systems is discussed very few times and mostly only as a secondary topic.

A physicist does not seem to feel the necessity to be epistemologically justified in using the knowledge and tools of physics to investigate even unrelated fields. Justifications and discussions are the job of the epistemologist and usually come very late and certainly not when, as in the case of network science, it is still in a very early stage of development (section 5.2.1). From a sociologist's perspective, the application of physical network theory is either not relevant because the main interest is in studying qualitatively the behaviour of groups of people, or it is refused as non applicable. One of the reasons for this refusal can be that a non-physicist has, sometimes, a mistaken idea of what physics is. Bernstein et al. (2000), for example, consider that sociologists believe the ideas of physics are mainly those of Newtonian mechanics, where single or small sets of *particles* are studied. They have well defined characteristics (mass, velocity, energy) and, more importantly, their equations of motion can be described and investigated. Based on this idea, sociologists consequently object that a 'social actor' is completely different from an undifferentiated particle, its behaviour is totally dependent on its history, its beliefs and its decisions and thus a system of particles is too simplistic a representation. If we consider models such as the ones proposed by Schelling (1971), Axelrod (1997) or Sznajd-Weron (2000) this remark seems justified.

However, in studying a socio-economic system we are mainly interested in its global behaviour and in the possibility of making predictions at this level rather than guessing the conduct of every single element (individual actors). The objective is to understand how regularities may emerge (when they do) out of the apparently erratic behaviour of single individuals (Majorana, 1942). In this perspective, the comparison with empirical data has the primary objective of verifying whether the trends seen in the data are compatible with a ‘reasonable’ microscopic modelling of the individuals and whether they are self-consistent or require additional factors.

In these circumstances, as Castellano et al. (2007) note, only high level characteristics, such as symmetries, critical transitions or conservation laws are relevant. These, as the principles of statistical physics show, do not depend on the microscopic details of the system. Therefore, Castellano et al. (2007: 2) state:

> With this concept of *universality* in mind one can then approach the modelization of social systems, trying to include only the simplest and most important properties of single individuals and looking for qualitative features exhibited by models.

In conclusion, the application of statistical physics laws and methods to the study of a socio-economic system such as a tourism destination can be considered justified, mainly if, as it will be shown in the rest of this thesis, the quantitative techniques rely strongly on a sound and accepted qualitative interpretation of the phenomena described.

5.3 The case study

Social science research (and economics and tourism studies as part of it) makes extensive use of case studies. The quest for new theories, the testing of existing models, ideas and assumptions, are activities performed very often by studying one or more *cases* carefully chosen. From these (or with these), it is possible to overcome the almost general practical and sometimes theoretical impossibility of performing systematic measurements to base theories and models. The discussion on the validity of the case method is still ongoing, and a recent paper states (Gerring, 2004: 341):

> Although much of what we know about the empirical world is drawn from case studies and case studies continue to constitute a large proportion of work generated by the discipline, the case study method is held in low regard or is simply ignored. Even among its defenders there is confusion over the virtues and vices of this ambiguous research design. Practitioners continue to ply their trade but have difficulty articulating what it is that they are doing, methodologically speaking. The case study survives in a curious methodological limbo.

However, the usage of case studies is widely spread in the scientific community. Other disciplines, even those commonly considered ‘hard science’ such as physics, geology or biology, have extensive recourse to it. In all cases in which it is not possible to perform a replicable laboratory experiment, an examination of a natural phenomenon is performed. From a single observation of a ‘case’, models and theories are built or existing ones are confirmed or refuted. Tycho Brahe’s observation of the comet in 1577 or the new star in Cassiopeia in 1572 led him to the conclusion that the Aristotelian view of the immutability of heaven was wrong. The eclipse of Jupiter's moon Io in 1676 induced the Danish astronomer Ole Rømer to uncover the finiteness of the speed of light. A single voyage and the ‘survey’ of the fauna of the Galapagos Islands in 1835 brought Charles Darwin to the idea of evolution. The presence of identical fossil species along the coastal parts of Africa and South America convinced Alfred Wegener to propose the continental drift theory in 1915. The measurements made during the 1919 total eclipse of the sun led Arthur Eddington to the first important confirmation of Einstein’s general theory of relativity. These are just a few well known examples, but the list could be much longer.

A “case study is an intensive study of a single unit for the purpose of understanding a larger class of (similar) units” (Gerring, 2004: 342) and whatever the field of study, the choice of a suitable case starts with a clear theoretical framework (Tellis, 1997a, 1997b). The rationalisation of theoretical topics dealing with the purpose of the study is a major determinant in the choice process. A review and a critical appraisal of the literature allows us to better specify the characteristics of the ‘possible case’ by comparing previous results, their validity and their limitations (Yin, 1994). Good use of theory helps not only to effectively design a case study, but also to guide the generalisation and the extension of the results. Yin (1994) focuses on the role of theories that exist prior to the study. He shows how knowledge of prior research is useful in case studies, and maintains that theory is important in order to build up a cumulative body of knowledge from a case instead of simply answering isolated empirical inquiries.

A sound theoretical framework can thus allow us to:

- select the cases to be studied in the first place, whether following a single case or multiple-case (replication) design;
- specify what is being explored when doing exploratory case studies;
- define complete and appropriate description for descriptive case studies;

- stipulate rival theories for explanatory case studies; and
- generalise the results to other cases.

Case studies can be single or multiple-case designs. Usually, the research is limited to a single-case design if it is not possible to find different examples for replication. Different types of case studies can be defined (Tellis, 1997a, 1997b; Yin, 1994): exploratory, explanatory and descriptive. In exploratory case studies, data collection usually helps in defining the research questions and hypotheses. Explanatory case studies help in causal investigations. Descriptive cases require a theory to be developed before starting the study and are used in the formation and validation of hypotheses of cause-effect relationships. Moreover, a case can be (Stake, 1995; Tellis, 1997b):

- *intrinsic*: when the researcher has an interest in the case;
- *instrumental*: when the case is used to understand more than what is obvious to the observer; and
- *collective*: when a group of cases is studied.

In any event, case study research, even in multiple-case designs, is different from sampling research which relies on considerations derived from measurements performed on probabilistic samples extracted from a defined population.

The unit of analysis is a critical factor for successful investigations. Typically, it consists of a system of action (Loubser, 1976: 244): "the organization of the components of action into a set of relatively stable mechanisms that reduce the complexity of the contingencies impinging on the achievement of the goal states of the actors involved". The idea is to be selective, choosing a limited number of issues considered fundamental to the understanding of the system being examined (Tellis, 1997b).

Generalisation of case studies is a debated issue. Some critics state that results are not widely applicable (see for example Lincoln & Guba, 1985). Yin refutes this idea by making a difference between analytic generalisation, meant as "a template with which to compare the empirical results of the case study" (1994: 31) and statistical generalisation in which "an inference is made about a population (or universe) on the basis of empirical data collected about a sample" (1994: 30). The one pertaining to case study research is an analytical rather than a statistical generalisation, therefore no point can be made in discussing the limitations in ‘sampling’. Single-unit studies provide cases that are likely to be *comparable* to one another (Gerring, 2004) and may allow us extensive testing of the causal implications of a theory. They are useful in providing evidence supporting the theoretical arguments. On the other hand, single-unit investigations may exhibit a *representativeness* problem in the extension and the generalisation of the results. Multiple-unit cases pose the problem of making assumptions about the comparability across the instances chosen. As for hypotheses testing, only a few of them are checked, but the multiplicity of units may give a stronger degree of confidence in confirming the results. However, it is at least possible to think of a ‘weak’ form of generalisation, named ‘naturalistic generalisation’ by Gomm et al. (2000) by which a case study is able to provide working hypotheses that may be appropriate for other (almost similar) cases.

Apart from these considerations, one issue in the choice of a suitable case is the availability of the data needed. The validity of case study research is greatly enhanced when multiple data collection methods are employed and both qualitative and quantitative methods are used (Eisenhardt, 1989; Yin, 1994).

The case chosen for this study, therefore, should follow the requirements discussed above. In addition, it should help to identify in the clearest possible way the boundaries of the destination. The issue here is that, like in many other parts of the world, the unit for which most of the data and information (mainly statistical) are collected is coincident with an administrative unit. From a general point of view, though, a tourism destination (at least as seen by tourists) could include a geographical area which is different from the administrative boundaries.

Moreover, the network built should have a size useful for conducting the analyses required. It is, obviously, preferable to deal with the smallest possible size in order to ease the data collection process and for ‘computational’ reasons; some of the algorithms involved could be very long for very large networks. However, networks with less than some hundreds of vertices are not considered sufficient to show their properties in a significant way (see for example Dorogovtsev & Mendes, 2003; Solé et al., 2000).

These considerations led to the choice of Elba Island as the destination to be investigated (see a description in section 5.3.1 and a discussion about data collection in section 5.4). Elsewhere it has been argued that Elba is a complex adaptive system (Baggio, 2008) which can be well represented in terms of a network. It is also a destination whose structure is similar to the one exhibited by many other destinations (Pechlaner et al., 2003; Ritchie & Crouch, 2003; Tallinucci & Testa, 2006). Being the first study performed with network analysis techniques, this work is exploratory, and the choice of a *typical destination* (Sheehan & Ritchie, 2005) will give more strength to the findings and the conclusions presented in this thesis.

The main reasons for this choice can be summarised as follows:

- it is an Italian summer destination and one of the most significant in terms of tourist flows (both domestic and international). It is a suitable and representative case study which complies with the requirements described in this section;
- its geographical boundaries practically coincide with the administrative unit. Elba is part of the Province of Livorno which includes two tourism boards: *Costa degli Etruschi* (covering the province coasts) and *Arcipelago Toscano* (the islands). In the latter, excluding Elba, only Capraia island has some (limited) tourism establishments, but they are easily identifiable. The destination boundaries, therefore, almost coincide with the administrative boundaries, thus easing the collection of tourism related data (arrivals and overnight stays);
- the 'size' of the destination in terms of number of actors involved (accommodation, services, intermediaries, transport, associations etc.) is of about a thousand elements, more than sufficient for the purpose of this research. If the sample was smaller, some of the statistical analyses used in this work might not have enough power to allow a clear discrimination between different hypotheses; and
- being an island, the destination is 'self-contained' and the boundaries are well delimited as is required for a case study (Yin, 1994). Even if not always true in a general case, the knowledge of this particular destination ensures that boundaries' effects (organisations physically located outside the destination administrative unit but mainly working in the destination itself) are quite limited and can be considered not significant for our purposes.

5.3.1 Elba Island

This section describes briefly the main characteristics of Elba Island with a focus on the economic conditions and the 'touristic importance' of the destination.

Elba is the largest island in the Tuscan archipelago, in the heart of the western Mediterranean Sea (coordinates: 42°46' N; 10°17' E). It is the third largest island of Italy (223.5 km^2). Its shape is elongated (27 km x 18 km) and its major axis is aligned almost exactly along the east-west direction.

At only three nautical miles from the mainland, it is strategically located in the middle of the naval routes reaching Tuscany, Sardinia and Corsica. Elba represents an important environmental resource in Italy. Its geographic position, temperate climate, the variety and beauty of its landscapes, the coast and the sea, make it a well known destination.

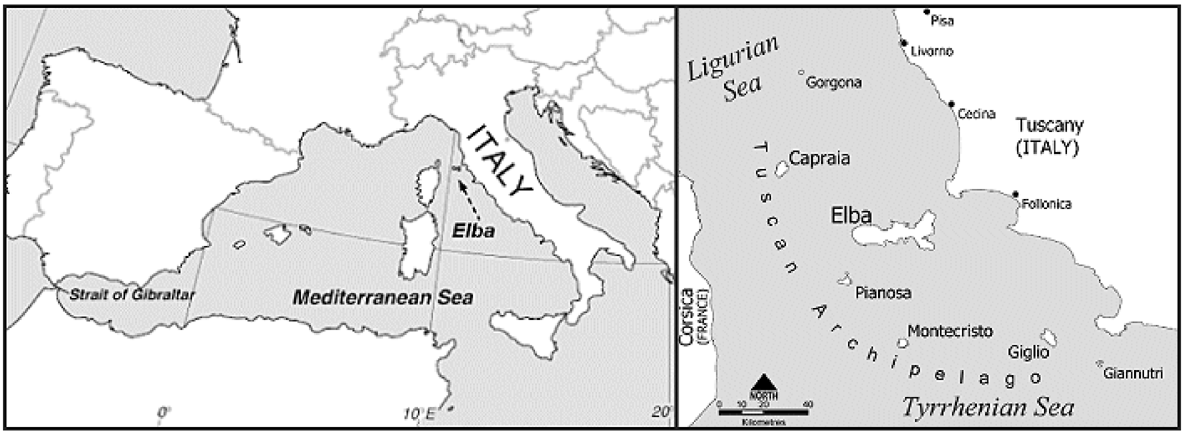


Figure 5.2 The geographical position of Elba

Geography, morphology and climate

The morphologic characteristics of the Tuscan islands come from their volcanic origins. Elba offers several different landscapes including mountains, such as Mount Capanne (1018 m), alternating with valleys and coastal plains.

About one third of the surface (223.5 km^2) is more than 200 m above the sea, while three quarters is more than 50 metres high. The coastline is 147 km long, with over 70 beaches (17,126 metres of beach; 7,908 sand, 3,510 shingle and 5,708 mixed). Almost two thirds of Elba is high and rocky. Small beaches can be found in correspondence to the major rivers.

The soil of the islands of the Tuscan archipelago is rich in minerals, and there is a natural museum for curious and interested people and for mineral scientists. The minerals have determined many economic and historic events of the archipelago, especially for Elba and Giglio. Their presence is so important that the archipelago belongs to UNESCO’s list of geologic sites.

The Tuscan islands have a peculiar ecosystem. There are interesting animals and very rare endemic plants. The location is of inestimable scientific value, but the ecosystem is fragile and delicate. The islands form also the most important corridor of the whole Mediterranean basin for providing shelter to many bird species on their migration paths from Europe to Africa. For these reasons the archipelago is a protected area. Founded in 1989, this is the largest marine park in Europe: 600 km^2 of protected water and 180 km^2 of protected land.

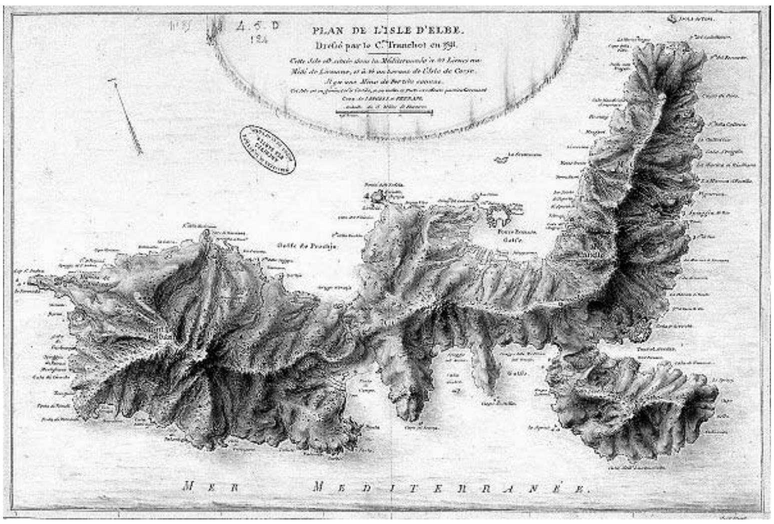


Figure 5.3 Elba in a map (dated 1791) of the Napoleonic period (after Godlewska, 2002)

The climate of Elba is typical of most marine areas, providing mild weather even throughout the winter period. Rainy days are reasonably scarce, particularly throughout the summer season.

Elba’s economy

Mining, active since the Greek exploitation started in the fifth century BC, had practically ended at the beginning of 1990s. Agriculture and farming are still present on the island, but have rapidly declined in the last 50 years. In 1971, an area of about 44.41 km^2 was agriculturally productive (almost 20% of the total); 10 years later, in 1981, the area had decreased to 23.85 km^2 and in 1991 only 16.50 km^2 were left to agriculture production. Today, no more than about 9% of Elba inhabitants are still employed in this sector.

During the World War II, the island was bombarded and occupied by German troops. Mining, the main activity that had supported the local economy, was practically destroyed and, after the war, the only alternative for the population was emigration.

The biggest economic efforts made by the Italian government go back to the period just after the end of World War II, with the launch of construction as a substitute for the closing of the iron and steel factory in Portoferraio. This lasted up until the last years of the 1950s, when Elba discovered tourism.

The population of Elba, around 30,000, has remained more or less constant in the last few years. One fifth of the population live in the major cities. From an administrative point of view, Elba belongs to the province of Livorno (Tuscany region) and comprises eight municipalities: Capoliveri, Marciana, Marciana Marina, Marina di Campo, Portoferraio, Porto Azzurro, Rio nell'Elba and Rio Marina.

Since the 1960s, tourism has grown to become the main (almost exclusive) economic activity on the island. Today, 200 hotels and 400 other accommodation establishments (camping, residences, vacation houses, etc.) provide more than 12,000 rooms with 36,000 bed places. A significant number (29%) of Elba people work directly in tourism related businesses. The tourism infrastructure has remained rather constant over the last years (Table 5.1 shows the number of hotels and bed places for the last five years).

The average tourist arrivals, with a steady increase in the last 50 years, have reached 500,000 with about 3 million overnight stays (

Table 5.2 and Figure 5.4). Elba (and the nearby small islands) is one of the top 20 Italian tourist destinations.

Table 5.1 Elba hotels and bed places in the last years (source: Elba Tourism Board)

Year	Hotels	Bed places
1970	153	7,319
1980	195	10,497
1999	210	16,255

2000	209	14,067
2001	212	14,022
2002	212	14,032
2003	211	16,082
2004	210	16,009
2005	208	15,924
2006	207	16,066

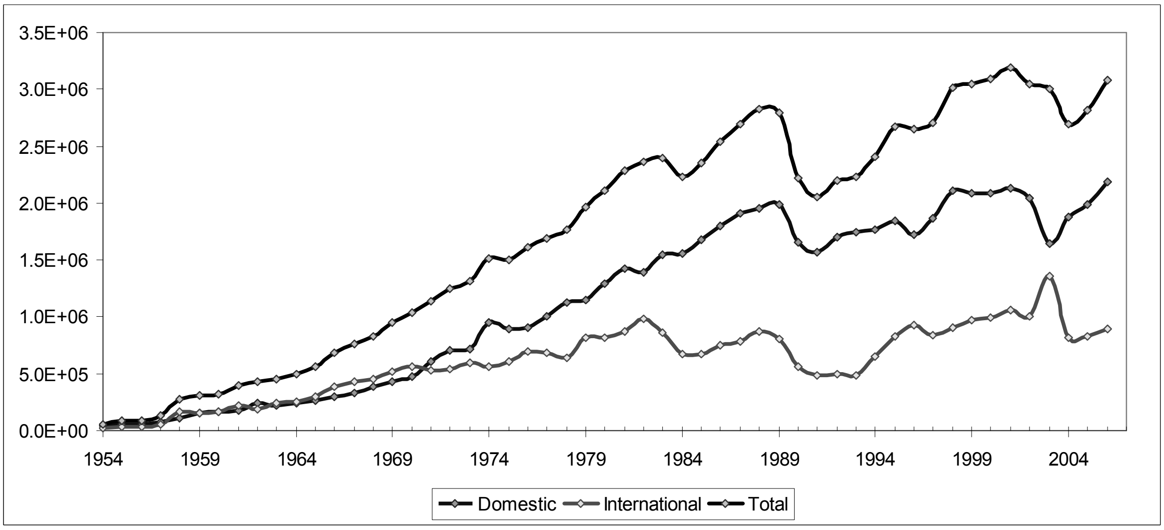


Figure 5.4 Tourism in Elba: overnight stays for the period 1954-2006 (source: Elba Tourism Board)

Table 5.2 Elba tourism statistics since 2000 (source: Elba Tourism Board)

	Arrivals			Overnight stays		
Year	**Domestic**	**International**	**Total**	**Domestic**	**International**	**Total**
2000	348,624	141,044	489,668	2,091,484	995,638	3,087,122
2001	349,469	149,255	498,724	2,132,504	1,059,729	3,192,233
2002	329,262	143,680	472,942	2,046,338	1,002,668	3,049,006
2003	242,288	150,803	393,091	1,650,343	1,352,772	3,003,115
2004	333,909	117,127	451,036	1,880,281	818,178	2,698,459
2005	334,993	112,790	447,783	1,992,083	824,585	2,816,668
2006	350,672	118,287	468,959	2,180,825	899,681	3,080,506
Average	*327,031*	*133,284*	*460,315*	*1,996,265*	*993,322*	*2,989,587*

As might be expected for such a destination, tourism shows a strong seasonality. The tourists' presence begins in April and May and declines in October, with a major concentration in July and August (Figure 5.5). Italians account for 70% of the tourism demand for Elba. Germany, Austria and the northern European countries are the most important foreign markets.

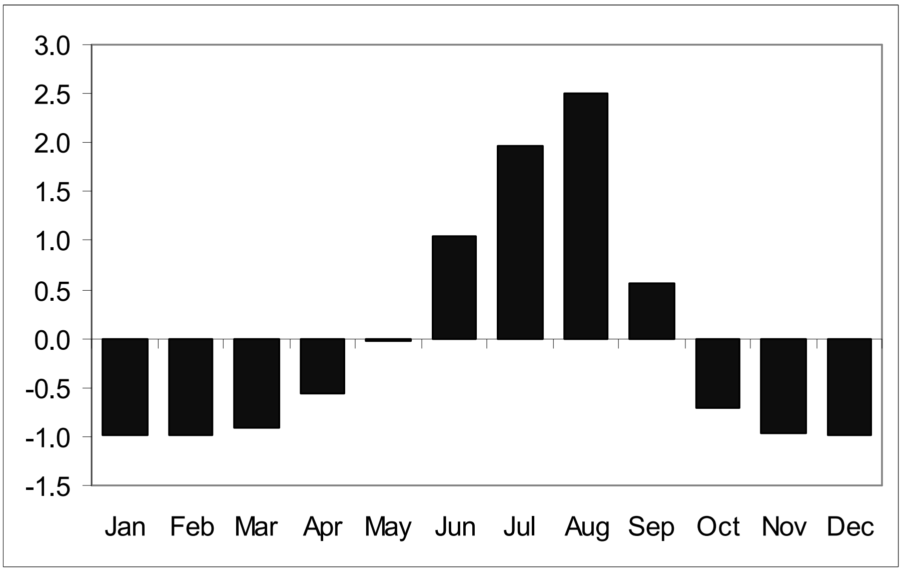


Figure 5.5 Overnight stays seasonality: 1996-2006 average (reference is year average) (source: Elba Tourism Board)

In conclusion, the geographical and environmental characteristics of Elba make it a pleasant tourism destination and its economic history has acknowledged this natural bent by making the island one of the main destinations in the Mediterranean. This 'case', therefore, looks especially interesting for the aims of this thesis and because of its resemblance to similar locations, facilitating some kind of generalisation of the results obtained here. Moreover, as will be shown in the next sections, the size, in terms of number of stakeholders, is more than sufficient to provide meaningful statistical analysis of their web of connections.

5.4 Data collection

The analytical framework developed in the previous chapters provides the basis for understanding how a tourism district can be represented and analysed.

The requirements of the theoretical framework are important in the selection of a method for data collection. They determine what data are relevant and need to be collected for adequate explanations (Yin, 1994). The unit of analysis considered here is a tourism destination composed of a collection of organisations (public or private) and their common relations. In other words, the destination is seen as a network whose actors are the single organisations and whose links are the connections established among them.

The investigation, therefore, starts with the specification and identification of the elements (stakeholders and their relations) of the district network and with the collection of the data needed to perform the different analyses. Social network analysis literature has dealt quite extensively with the problem of enumerating the ties among a network's actors and with the reliability of the different methods proposed to collect the necessary data (Marsden, 1990). The problem has been well known since the seminal work by Ove Frank on statistical inferences in graphs (Frank, 1971).

The analysis of such networks relies on data collection procedures that may be difficult to execute or lead to incomplete or unreliable outcomes. The collection is typically based on surveys, with a number of different techniques that aim at highlighting the connections among the different actors, and statistical sampling procedures are applied. Sampling global network properties, however, is challenging since errors grow exponentially. Informants may be implicated in relationships with numerous other actors with different 'intensity' so that their subjective judgement may induce strong bias in reporting the situation and their accuracy may be questionable (Bernard et al., 1984). In a complex network, even minor errors can have striking effects on the properties as a whole. Moreover, after the extensive studies conducted recently (see Chapter 2), it is now clear that the topology of a network may greatly affect these measurements.

In order to draw up feasible and replicable methods for the collection of useful data, it has been deemed important to proceed in two steps: the first one is the design of a range of survey techniques and the second is the assessment of the completeness of the sample gathered and the evaluation of the effects on the quantities to be derived. This latter issue is discussed in the next section.

5.4.1 Sampling issues

Classical statistical methods allow a wide range of possibilities to estimate, infer, generalise and model some characteristics of a set of objects (a population) by studying the same characteristic in a subset of elements: a sample. A number of randomly chosen items are investigated and, if the distribution of the population is known or can be supposed, we assume that the sample possesses the same distribution. A wide variety of sampling schemes (simple random, stratified, probability proportional to size, systematic, cluster, multistage) are also available, all well described and studied in terms of validity, reliability and applicability to different situations (Cochran, 1977; Gentle et al., 2004; Langley, 1971; Shao, 1999; Sheskin, 2000).

Until the early 2000s, the only network model available was the one described by Erdős and Rényi (section 2.4.1), in which the connections among the vertices are distributed at random and follow a Poisson-Normal distribution. In this situation, all known sampling techniques hold well and even a large network can be approached by using 'standard' statistical methods. The problem of adapting known sampling schemes to these systems has been approached by social scientists, and the main issues discussed have been the identification and enumeration of the social network members, the delimitation of such a 'fluid' system, and the reliability of the response to surveys which try to highlight connections among the system components or the accuracy of informants reporting on social group characteristics (Bernard et al., 1984; Doreian & Woodard, 1992; Erickson & Nosanchuk, 1983; Frank, 1979; Frank & Snijders, 1994; Grannis, 2004; Holling, 1978; Killworth & Bernard, 1976; Laumann et al., 1983; Leydesdorff, 1991; Marsden, 1990; Rothenberg, 1995; Scott, 1988; Tichy et al., 1979a; Wasserman & Faust, 1994).

The boundary specification problem is one of the most important issues (Laumann et al., 1978; Laumann et al., 1983). It consists of the definition of the rules for the choice of actors and the ties of a network. Ignoring part of the elements, or lacking part of the data, can have a direct influence on the estimation of the main network parameters. Simulation studies show that clustering and assortativity coefficients are overestimated when both actors and relations are omitted, and underestimated when ties are missing or incorrectly reported in surveying the network components (Kossinets, 2006). On the other hand, when the links are correctly identified, some centrality measures (in-degree and simple eigenvector, see Chapter 2) are found to be relatively insensitive to the sampled number of vertices in the network, so that even 50% of missing actors still allows a reliable estimate to be formed (Costenbader & Valente, 2003). Important for this research is that, even in the presence of some data collection bias, significant evaluations of important network parameters can be performed.

In sampling social networks, a widely used method is so-called *snowball sampling* (Frank, 1979; Frank & Snijders, 1994). The technique allows the building up of a sample by starting from an initial set of network actors and asking them to list other actors with whom ties exist. The new 'members' are then queried to determine their connections and so forth. The statistical properties of snowball samples have been studied and the method is thought to provide good samples of populations where a large number of steps are performed. However, since sample elements are not selected independently from the environment in which the investigation is conducted, a 'snowball' can be strongly biased (Rothenberg, 1995; Sudman & Kalton, 1986). Several ways to overcome this problem have been proposed. One consists of creating a *fixed list* containing an inventory (or a random sample) of network members and asking the respondents to identify those with whom a relationship exists. The results, nevertheless, are not fully satisfactory and the method is shown to perform more poorly, especially, if the network has some growth mechanism (Doreian & Woodard, 1992).

A better proposal is the respondent-driven sampling method in which a basic snowball sample is weighted to compensate for non-random selection patterns (Heckathorn, 1997). Weights are determined by assuming that the different snowball steps select vertices with probability proportional to their degree and by using Markov chain theory. It has been shown that this technique allows one to make asymptotically unbiased estimates of network size and clustering across groups and to make estimates about hidden parts of a social network (Salganik & Heckathorn, 2004).

One more crucial issue, stressed by many authors (Bernard et al., 1984; Killworth & Bernard, 1976), is the accuracy of the reports by the informants contributing to the sampling schemes, mainly for what pertains to 'historical' information on the evolution of the social network. The main answer to this problem resides in the necessity to cross-test all the data gathered and to use and compare different collection methods. Moreover, it is noted that the time elapsed between an interview and the time at which information is sought is strongly related to the reliability of the data (Marsden, 1990; Wasserman & Faust, 1994).

However, the most recent results on network topologies (see Chapter 2) have radically modified this view and added a number of new issues in the sampling of a complex network. All the considerations made so far assume that the network has a random Erdös-Rényi (ER) structure (section 2.4.1) while, as seen in Chapter 2, most real networks exhibit a structured non-homogeneous topology. The effects of an incomplete sampling of network components (both nodes and links) on the structural characteristics needs, therefore, to be assessed in a different way. This can be done by using the results provided with regard to the robustness or the fragility of such networks (section 3.2.1).

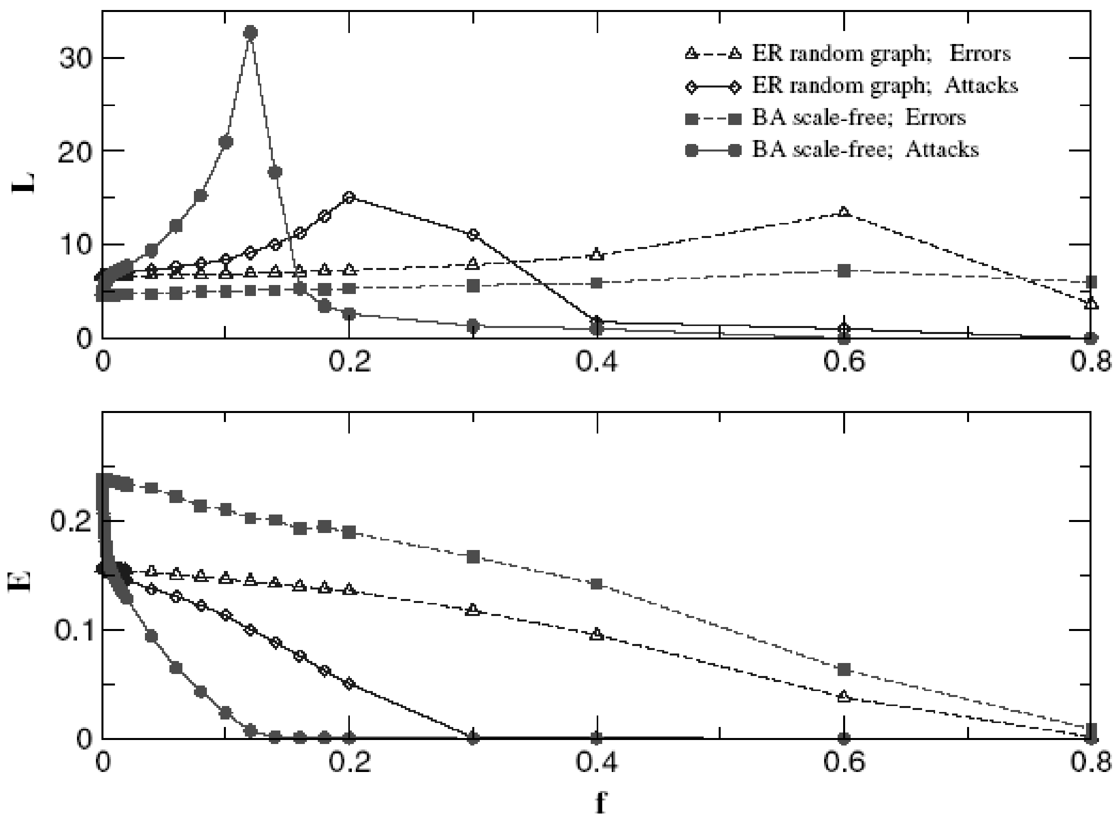


Figure 5.6 Effects of random removals (errors) and targeted attacks (attacks) for random (ER) and scale-free (BA) networks (f is the fraction removed) on average path length (L) and efficiency (E) of the system (after Boccaletti et al., 2006)

Consider, for example, a scale-free (SF) network and compare its behaviour with respect to a random removal of vertices or to a targeted removal of the most connected ones. As seen in section 3.2.1, the results of this procedure lead to the situation depicted in Figure 5.6. The overall connectivity of the network is quite robust under random node removal, while it is easily lost when coordinated attacks affect just a few highly-connected nodes. It is clear, then, that an incomplete network may show properties that have different characteristics from those possessed by the complete system.

The relationship between the network parameters measured on a partial sample, and the complete one, can be inferred by simulating a sampling process and by calculating the difference effects. With this approach, a number of recent studies have provided insights into these issues (Han et al., 2005; Lee et al., 2006; Rafiei & Curial, 2005; Smith et al., 2003; Stumpf et al., 2005a; Stumpf & Wiuf, 2005; Stumpf et al., 2005b).

Table 5.3 Effects of sampling methods on SF network properties. As the sampling fraction decreases (⇩), the quantity may increase (⇧), decrease (⇩), stay the same (⇔), or behave according to the specific network (⇳) (adapted from Lee et al., 2006)

	Degree exponent	**Average path length**	**Clustering coefficient**	**Betweenness exponent**	**Assortativity**
Node ⇩	⇧	⇧ (⇩)	⇳	⇧	⇔
Link ⇩	⇧	⇧ (⇩)	⇩	⇧	⇔
Snowball ⇩	⇩	⇩	⇳	⇩	⇩

Table 5.3 shows the behaviour of main network parameters in the case of an SF network for different sampling methods. It must be noted that the average path length may have a stronger dependency on the average degree, decreasing with the increase in degree. This effect may overcome the one shown in the table for peculiar cases.

If the network has an ER (or, generally, exponential) topology, the dependency is different (as discussed above). For example, a random sample preserves the shape of the degree distribution while modifying the average degree value (see Figure 5.7).

These results provide general criteria for sampling methods when some specific parameter is investigated and allow a better evaluation of the analysis results. By estimating the sampled fraction, therefore, it is possible, at least, to determine bounds for the main network parameters.

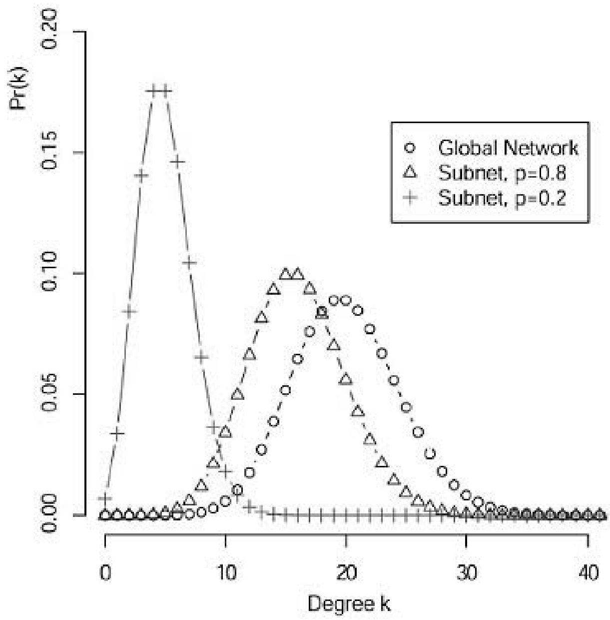


Figure 5.7 Effects of the sampling fraction (p) on the degree distribution of an exponential (ER) network (after Stumpf & Wiuf, 2005)

The main implications of the preceding discussion for this work is that an assessment of the completeness of the data gathered is needed in order to correctly evaluate its influence on the estimate of the main topological network parameters (see section 5.4.2).

5.4.2 The empirical data

Tourism activities are difficult to define precisely and the same, if not increased, difficulty is found in defining what economic activities should be included in the idea of 'tourism destination' (see the discussion in Chapter 2 and section 4.2). For our purposes, and with the general interpretation of a tourism district given in section 4.2.3, the elements considered as part of our unit of analysis were those organisations (companies, associations, groups etc.) belonging to the areas specified in the following list (Table 5.4).

The list draws mainly on similar inventories found in the principal recommendations of international organisations for the collection of tourism statistical data (EUROSTAT, 2000, 2002; OECD, 2000a, 2004; UNWTO, 1995, 2000). This way, a better replicability of the study methodology was ensured.

Table 5.4 Components of a tourism destination

Category	Subgroups
Accommodation	Hotels and similar establishments
	Other accommodation establishments
Food-and beverage-serving	Restaurants and similar
	Pubs, bars and other
Transport	Road and water transport
	Transportation means rentals
	Maintenance and repair services
Tourism intermediaries	Tour operators, travel agencies
	Information and tourist guide services
	Real estate services
Cultural resources	Museum and other cultural services
	Performing arts
Recreation and entertainment	Sports and recreational sport services
	Night clubs, discos
	Other amusement and recreational services
Tourism associations	Industry associations
	Consortia
Public tourism bodies	
Other tourism services	

Table 5.5 Elba destination network actors collected for the study

Category	No.
Accommodation	408
Food & beverage	341
Transport	5
Tourism intermediaries	46
Cultural resources	28
Recreation and entertainment	170
Tourism associations & consortia	17
Public tourism bodies	2
Other tourism services	11
Total	*1028*

Table 5.5 were identified and included in the study. The sources for this data are the lists published by the Elba Tourist Board and by the register of companies for the province of Livorno (the administrative area to which Elba belongs).

The official estimate of tourism suppliers by the Elba Tourism Board (in year 2005) was 204 hotels; 251 other accommodation establishments; and 535 restaurants, bars etc.

The sample considered is lacking some units in the 'other accommodations' and in the 'food and beverage' categories (typically some of the smallest private houses or restaurants and bars). This difference is mainly due to difficulties in the updating of official lists by the local tourist board. Besides this difference, which may well be regarded as non significant, the sample used can be considered complete. The total number (1028 items) fully complies with the requirements for a significant application of the network analysis methods proposed in this thesis (see section 5.4).

The second component of a network is the set of linkages among the system elements.

The main focus of this work is on the general topology of a tourism district network. Therefore, for the present study, all the links were considered bidirectional and unweighted. In other words, no importance was given to their strength or intensity. For a network of the type involved in this research (a social network formed by organisations belonging to a tourism destination), relational linkages can be:

- transaction relations;
- communication relations;
- information exchange relations;
- boundary relations;
- instrumental relations;
- power relations;
- social relations;
- kinship relations.

These relations were identified in the following way. First of all, publicly available information, documents or records were consulted. These consisted of the promotional materials published by the local tourist board, the tourist operators' websites, the official company records of the Province of Livorno (for the board of directors compositions and companies' ownerships), and the member lists of the different associations. A link between two actors was considered to exist if any of the following occurred:

- membership in associations, groups or consortia;
- belonging to the same group (e.g. owned by the same individual or group of people);
- co-presence of members in the board of directors;
- commercial agreements;
- cooperation or collaboration agreements;
- joint advertising on brochures or catalogues, or hyperlinks on organisation's website.

All data were collated into a database table. The elements coded were: stakeholder's name, type of business (as categorised in Table 5.4 and Table 5.5), geographical location (municipality of the company's headquarters), size of the company or organisation (small, medium, large). Each name was assigned an alphanumeric code representing these characteristics. The links were recorded into a different table. Each record consisted of two entries (fields) containing the codes of the connected stakeholders. This table was then used as input for the network analysis software packages listed in section 5.5.

One more set of data was collected regarding the technological network. The destination stakeholders' inventory (section 5.4.2) was examined and the list of existing websites was composed. Only websites with an *independent* URL were considered (i.e. pages belonging to larger portals, whose Internet address looks like: www.someportal.tld/operator_name were disregarded). Links among the Elban websites (defined as websites belonging to operators based in Elba) and from these to others outside Elba were counted by using a simple crawler and complementing the data obtained with a visual inspection of the websites. The visual inspection was needed to ensure the absence of missing connections and to overcome the limitations of the crawler in recognising the different forms of hyperlinking. The final dataset comprised 468 elements (websites) with 495 links. This size, although not large, can be considered sufficient to show the statistical properties of the graphs in a meaningful way (Angeloudis & Fisk, 2006; Dunne et al., 2002). Also in this case all the data were hand coded into a database table following the same layout used previously.

A series of in-depth, semi-structured interviews to the principal experts in the field were used to confirm the quality of the data collected and to highlight other possible interconnections. The participants were chosen among the directors of the Elba tourism board, the presidents of the main operators associations, members of the board of the main consortia and professional tourism consultants operating at Elba. In total, 13 people were consulted.

The semi-structured form of interview was chosen for its characteristics with respect to the objectives of this part of the research (Lofland et al., 2006). Apart from a few direct questions on the existence and type of interorganisational connections (see the interview protocol in Appendix 2), the aim of this series of interviews was to gather information on the general situation of tourism at the destination and on its development history. Moreover, the expertise and direct knowledge of the informants was used to qualitatively assess the completeness of the set of connections collected with the techniques discussed above.

Given these objectives and the methodology chosen, the number of interviewees can be considered sufficient to reach a good result. In qualitative investigations such as this, in fact, the literature suggests ‘sample sizes’ of this order of magnitude (Creswell, 1998; Dukes, 1984; Kvale, 1996). Even if the experience and knowledge of Elba tourism was not uniform across the people involved, the overall level was quite high due to their personal histories and their present roles. These differences were taken into account during the interviews by making them flexible as to what to ask and how, and which answers to follow up and which not to (Kvale, 1996).

During the course of each conversation the responsiveness was very good, all the participants were collaborative and, as will be seen in the exposition of the outcomes and the discussion (Chapters 7 and 8), their contribution was effective in pointing out several issues which proved extremely useful for the qualitative interpretation of the results of the quantitative analysis presented in this thesis. All the interviews were documented, and a second run (performed by e-mail) helped to clarify the thoughts and facts discussed by some of the interviewees.

Finally, once all these elements (actors and ties) were collected, an assessment was performed to estimate the completeness of the sample and to derive possible influences on the main network parameters (see section 5.4.1). The assessment strategy for this work was, therefore, based on triangulation (Yin, 1994). Triangulation has been described (Denzin, 1989; Flick, 2004) as a method that allows the researcher to validate the results of the investigation by using data drawn from different sources, by using different researchers in the same study, by approaching data using different theories and perspectives, or by using different methods. The recognised advantage is to limit personal and methodological biases while enhancing the generalisability of the study (Decrop, 1999; Olsen, 2004)

5.5 Network analysis

The data collected allowed a graph of the tourism district network to be drawn. After that, the main characteristics were deduced with the analysis starting with the categorisation of the static properties of the system and then proceeding with the study of its dynamics and evolution. All the values computed for the case under investigation were compared with those of a synthetic graph, which has the same order (number of nodes) and size (number of links) and a random distribution of the links. This procedure allows a better understanding of the significance and of the 'physical' interpretation of the quantities involved.

The calculations were performed by using several software packages widely employed in network analysis, and with the help of programs specifically developed for the purpose of the present study. The list of tools used includes:

- PAJEK: a software for large network analysis and visualisation (Batagelj & Mrvar, 1998);
- UCINET: a general program for the analysis of social network data as well as matrix analysis and algebra and multivariate statistics (Borgatti et al., 1992);
- SPSS: for statistical data analysis and charting (SPSS, 2004);

- MATLAB: a high-level technical computing language and interactive environment for algorithm development, data and analysis, and numeric computation (MATLAB, 2004);
- NETLOGO: a toolkit developed to implement agent-based simulation models (Wilensky, 1999).

5.5.1 Static characteristics

The basic measures computed for the Elba tourism network are those described in Chapter 2. They were used in this context to assess the general structure of the network and the general characteristics of the tourism district.

The quantities computed are (see Chapter 2 for definitions and formulas):

- density of links;
- degree distribution;
- distribution of paths and average path length;
- diameter;
- clustering coefficient;
- local and global efficiency;
- betweenness centrality; and
- assortativity.

Moreover, the entropy of the network can be calculated. Several definitions and computation methods have been proposed for this quantity. According to Demetrius and Manke (2005), for a Boolean graph (unweighted and undirected), whose dominant eigenvalue is λ (the largest eigenvalue of the adjacency matrix A), it is possible to compute the Kolmogorov-Sinai (KS) entropy as:

$$H = \log \lambda.$$

H is also called *topological entropy*, since it is an invariant quantity, depending only on the characteristics of the network and takes into account both the degree distribution and the average path length (see section 2.3.2). *H* is shown to be positively correlated with the robustness characteristics of a network and can be taken as an index for this feature (Demetrius & Manke, 2005).

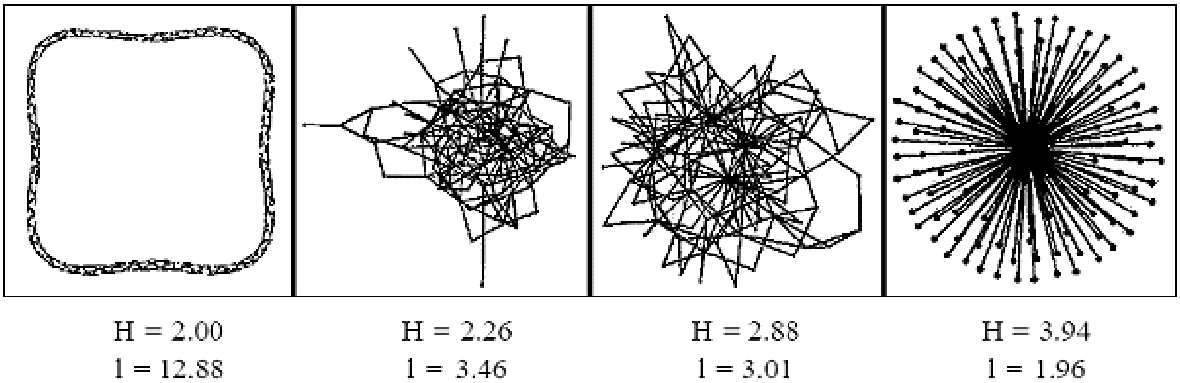


Figure 5.8 Topological entropy (H) and average path length (l) for different network topologies. The networks have same number of nodes (N = 100) and edges (M = 200) (after Demetrius & Manke, 2005)

A social network such as a tourism district is, intuitively, not a uniform network, but it may exhibit some degree of clustering. Two possible measures for this feature can be used: the clustering coefficient and the modularity index. All the proposals for these measurements (see section 2.3.2 and 3.2.5), however, take into account only the general structure of the network, ignoring 'semantic' characteristics. In our case, we were interested in assessing the degree of cohesion of defined sets of actors, whether grouped by geography (actors in the same area of the destination) or by type of business (accommodation, intermediaries etc.), as representative of the degree of collaboration among them.

For this purpose the modularity index was calculated. Following what discussed in Chapter 3, the modularity index can be written as:

$$M = \sum_{s=1}^{N} \left[\frac{l_s}{L} - \left(\frac{d_s}{2L} \right)^2 \right]$$

where: N is the number of groups of actors, L is the total number of edges in the network, l_s is the number of links between vertices in group s and d_s the sum of the degrees of the nodes in group s.

The groups considered were derived from the geographical division (stakeholders located in the eight Elba municipalities) and by business type. The results thus obtained were compared with those of the algorithm proposed by Clauset, Newman and Moore (Clauset et al., 2004) which improved the Newman-Girvan method described in section 3.2.5. Moreover, a random network of the same size (same number of nodes and density, with a random distribution of the links) was used as *null model*.

The spectrum of a graph is strictly connected with its topological characteristics. The spectral analysis of the adjacency matrix and of the Laplacian matrix of the tourism destination network can be a useful, and in many cases computationally more efficient, method to derive its main parameters (see Chapter 2). Moreover, some features such as cohesive clustering, long paths and bottlenecks, and randomness of the network are highlighted by the study of the network spectrum.

A spectral analysis of the tourism district network, therefore, was performed and all the main considerations discussed in Chapter 2 were applied to the specific case under investigation.

5.5.2 Dynamic processes

The structural properties of the tourism district system, as of any interconnected system, have a great impact on its dynamic characteristics and on the behaviour of dynamic processes that may occur (see section 3.2).

In this scenario, the statistical and spectral methods described above (Chapter 3) were used to assess the robustness of the Elban network. The basic hypothesis is that a relationship between the network connectivity features and its resilience exists, where resilience means the capability to adapt, by resisting or changing, to potential hazards in order to maintain an acceptable level of functioning. In particular, the topological entropy measure (see section 5.5.1), taken as an index, can be compared to the same quantity for other similar economic and social systems as stated in the literature. More precisely, a simulation on the network data was performed. By iteratively removing network elements (randomly or targeted) the connectivity is measured to find the presence of possible critical parameters that may have affected the general network robustness properties. The simulation applied the model and the procedure described by Albert et al. (2000) and Latora and Marchiori (2001). We started with the whole network and computed the local E_{loc} and the global E_{glob}

efficiencies. At each step a random number of links (3%) was removed and the parameters calculated again. The simulation and the parameter calculations were programmed with the Matlab development environment (MATLAB, 2004). The results are compared with a similar simulation performed on a random network. The comparison allowed us to infer the effects of the actual topology on the capability of the network to stay connected (robustness).

In a tourism destination system, issues regarding the diffusion of information or knowledge are crucial matters for its efficiency and effectiveness (see section 2.5). The method used to assess the network properties was derived from the one described in section 3.2.2 and stems from the tradition of mathematical models used in epidemiology to study the diffusion of human diseases (Bailey, 1975; Diekmann & Heesterbeek, 2000; Hethcote, 2000). There is clear analogy here between the transmission of disease and the transmission of information or knowledge through a network (Bettencourt et al., 2006; Iribarren & Moro, 2007). The main difference is that, conventionally, epidemiologocal models have assumed *perfect mixing*: i.e., all individuals are equally able to infect all others, and have taken into account a random distribution of the contacts between individuals that are responsible for the infection (diseases spread through some kind of contact between the population elements). However (see section 3.2), recent advances in the study of complex networks have forced a reconsideration of diffusion models in order to take into account the effects of non-homogeneous network topologies (Kuperman & Abramson, 2001; Pastor-Satorras & Vespignani, 2001).

In the simulation conducted here a simple SI epidemiological model was used. In this, the single individual can assume only two states, susceptible (S) to the infection, or infected (I). Despite its simplicity, this class of models has been shown to be quite effective and to be a good approximation of more refined and complex models (Barthélemy et al., 2005; Xu et al., 2007). In addition, it is suitable for describing the knowledge transfer process. In fact, we may well reasonably assume that once knowledge has been transferred to a new host, it will retain the knowledge received, therefore it will remain infected. This is an essential pre-requisite to innovation as unless the knowledge is transferred and used by enterprises at the destination, innovation will not occur.

The algorithm used for the simulation was the following:

- the network is loaded;
- one randomly chosen stakeholder starts the spread by *infecting* a proportion k_i of its immediate neighbours. In tourism, this stakeholder is often a government-funded tourist board or economic development agency;
- at each time step the infected elements transfer the knowledge to a proportion k_i of their immediate neighbours; and
- the process ends when all the network nodes have been infected.

As a parameter for the model, the capacity of the single stakeholder to transfer knowledge was used. It can be expressed as a probability $k_{i,}$ whose value controls the number of neighbours which are informed by a single stakeholder. This accounts for an important difference between information and knowledge flows and the spread of viruses. While viruses tend to be indiscriminate, infecting any susceptible individual, knowledge is selective and is passed by its host only to a limited set of the individuals with which it has relations (Huberman & Adamic, 2004). Moreover, particular actors can have difficulties in acquiring and retaining all the knowledge available to them (a feature usually called absorptive capacity, see for example Cohen & Levinthal, 1990; Priestley & Samaddar, 2007) due to their internal functioning or because of the associated costs. In tourism, this issue of absorptive capacity is critical, particularly given the dominance of SMEs in the sector.

We can assume that the capacity of transferring knowledge is different for the different ‘sizes’ of companies involved. Therefore, the network nodes were divided into three categories: large, medium and small. In our case we had the following proportions: large = 7%, medium = 16%, small = 77%. The values for the proportion of neighbours informed used in the simulations run were (arbitrarily) set as: $k_{large} = 1$, $k_{medium} = 0.8$, and $k_{small} = 0.6$. Since the structural characteristics of the network, and particularly the cohesion among stakeholders, can be a factor influencing the knowledge transfer process, the experiment was also performed with a modified version of the original network. This was obtained by rewiring the connections while leaving unchanged the original connectivity (i.e. the number of immediate neighbours of each stakeholder and the overall density of linkages), in order to obtain a higher clustering coefficient and efficiency. The model described was implemented by using the NetLogo (Wilensky, 1999) software environment and is a derivation of some of the distribution library models (Rumor Mill as modified by F. Stonedahl: http://www.cs.northwestern.edu/~fjs750/netlogo/).

The efficiency of the network, both global (E_{glob}) and local (E_{loc}), was used as an overall index to assess the capability of the system to ensure smooth and effective flows of information. This efficiency can be optimised. The optimisation was achieved by using a simulated annealing algorithm and a Metropolis algorithm (see section 3.2.4). The function to be optimised was the local efficiency (E_{loc}) of the network. The following procedure was used:

- the actual network, with addition of a fraction p of the original number of links, is used as the starting configuration;
- the efficiency for the initial configuration C_0, $E(C_0)$, is calculated and an optimal final efficiency E_{opt} is set;
- a 'temperature' value T, a temperature schedule ΔT (intervals of reduction) and a minimum threshold T_{min} are set;
- a random rewiring of the network is performed, obtaining a new configuration;
- the efficiency of the new configuration is calculated with its variation with respect to the previous value;
- the new configuration is accepted if the difference in efficiency ΔE is positive (the new efficiency is higher); if the new efficiency is lower it is accepted with probability $p = e^{-\Delta E/T}$;
- if T_{min} or E_{opt} have been reached, the configuration obtained is the final one, otherwise the temperature T is decreased by ΔT and the whole process is repeated.

The algorithm is schematically depicted in Figure 5.9.

The structural parameters of the optimised configuration were finally calculated and tested again with respect to robustness or diffusion processes.

```
INITIALIZE (configuration, temperature, temperature
            schedule, criterion)
compute function E for configuration C0: E (C0)
LOOP
  generate new configuration: C1
  compute ENEW (C1)
  evaluate: ΔE = ENEW (C1) - EOLD (C0)
  IF < 0 THEN
    accept new configuration: C0 := C1
  ELSE
    accept with probability p = e^(-ΔE/T)
  update configuration
  EXIT for criterion satisfied
  update temperature
END LOOP
```

Figure 5.9 Schematisation of the Metropolis algorithm used for network optimisation

One more way to empirically test an information diffusion process is to perform an experiment by simulating the spread of a piece of information and comparing the time needed to notify the whole network (or a reasonable part of it). A comparison between the original system and the optimised one gives a quantitative measure of the effectiveness of the optimisation process. This can be done by using the following algorithm:

- at time t_0, an initial number of nodes is informed (in an active status);
- at time t, an informed node will inform a fraction f of its first neighbours, skipping the one from which it received the information at time $t-1$ and those that have already been informed (at time $t+1$ a node informed at time t will start informing their neighbours, the nodes that were already active will continue informing);
- the process is repeated until a fraction F of nodes of the whole network is informed;
- the possible disconnected nodes are treated as having very weak ties within the social group; therefore, at the end of the process, in a final step, a fraction G of the disconnected nodes is activated and connected with some randomly chosen nodes. This may help in better simulating a real situation and taking into account possible problems in the complete enumeration of the links.

The total number of steps represents the period of time needed to notify the whole network (or the fraction $F+G$ wanted).

5.5.3 System evolution

The actual topology of a network can be explained with different growth models (see section 3.1). The approach followed for identifying the candidate models is to simulate a number of different evolution mechanisms in order to find those that better match the resulting network characteristics in terms of connectivity patterns: degree distribution, average path lengths and clustering coefficients.

The results deduced in this way provide the first part of the answer to the main research question: how does this tourism network evolve?

The second part of the answer is found by considering the possible relationship between the network development and the evolution of the destination.

As discussed in section 4.2.3, the development of a tourism destination is usually studied by considering some measurable outcome, such as tourist arrivals (TA) or overnight stays (OS). The first one, TA, is a mostly demand oriented measure, while the second, OS, takes into account some destination (supply side) structural parameters as well (see for example Ferro Luzzi & Flückiger, 2003; Garín-Muñoz & Amaral, 2000). This consideration (coupled with a wider availability of the basic data) led to the choice of overnight stays as the indicator.

Traditionally, one of the well known models of tourism destination evolution (in the vast majority the model due to Butler, 1980, 2005a, 2005b) is used to explain the destination behaviour as a whole. In many cases, amendments, modifications or extensions of the model are provided to better match the theoretical framework and the empirical data (see section 4.2.3).

In the present study a different approach was taken and the destination history was evaluated in terms of structural (topological) changes. The method consisted of analysing the behaviour of the overnight stays time series for Elba. As can be seen in Figure 5.10, the steady increase of the series has three 'breaks' (a, b and c). A period of peaking overnights is somehow interrupted, then starts again.

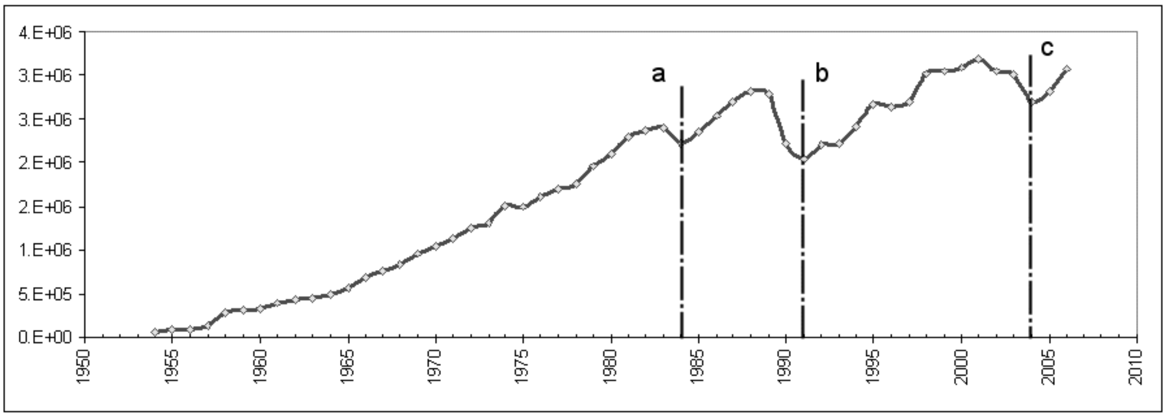


Figure 5.10 Overnight stays at Elba. Possible structural breaks in the time series are indicated by vertical lines and marked a, b and c (data source: Elban Tourism Board)

The tourism destination network, in these periods, can be investigated to check whether there is any correlation between the series behaviour and the topological characteristics of the network. This can also give more elements to confirm the evolutionary model chosen above. Availability (and reliability) of historical data and informants' accuracy typically decreases for more distant periods (see Bernard et al., 1984; Killworth & Bernard, 1976) leading us to consider a timeframe in the neighbourhood of 1990-1992.

A complete quantitative assessment of the historical development of the Elba network would require a precise knowledge of the configurations at different times in the past. This knowledge is not available, at least not at the level required for exact calculations, therefore the layout of the network was estimated by considering the available data and the interviews conducted.

The network was rebuilt with the information collected and its characteristics deduced following the methods presented in section 5.5.1. After that, a new simulation determines the dynamic evolution. The results were compared with those obtained as per section 5.5.2 and a final assessment of the dynamic behaviour of the destination network was performed.

In order to gain a better understanding of the possible relations between structural evolution and the outcomes in terms of tourist stays, the general environmental situation is described. Political, social and global conditions for that time period are addressed.

5.5.4 The technological network

The web space is, nowadays, an important virtual counterpart of an economic and social system. The Internet age has allowed the development of new ways for producing and distributing tourism services. Web-based approaches and technologies are helping tourism suppliers and agencies reduce service costs and attract customers (Buhalis, 2000, 2003). A website looks to be a major tool (and, probably, it will be the only one in the future) to conduct business in the tourism field (see the surveys by Marcussen, 2006; PhoCusWright, 2005). As for many other destinations, the Web has become, in recent years, an important means of promotion and commercialisation for the whole community of Elban tourism operators. The wide diffusion of these technological tools for both the demand side (tourists) and the supply side (operators) may allow us to use the website network as a further element to assess destination characteristics.

The analysis of the technological networks started by considering them as nodes of a complex network, and the data needed was collected as described in section 5.4.2. The topology of this network was studied to assess its agreement with the known models for the Web and to compare the derived values with those published in the literature for other similar examples (see section 3.3).

The results, mainly in terms of connectivity, were then compared with those of the ‘real’ destination network. This allowed us to determine similarities and differences between the ‘real’ and ‘virtual’ behaviour of the destination stakeholders. Moreover, these results are used to assess the capacity of Elban tourism operators to exploit modern technologies in conducting their business.

5.6 Research boundaries

The scope of this study was potentially large, and therefore a number of boundaries were needed to make it both focused and manageable. These are given below.

The literature in the fields concerning this study is expanding very rapidly and knowledge has developed conspicuously. There are no doubts that this will continue in the timeframe necessary for the completion of this work. Every attempt has been made to review the most important literature, but there are, inevitably, works that might have been missed. However, the theoretical framework derived in the previous chapters can be considered fairly stable and complete as the most recent reviews in this field (network science) show (Boccaletti et al., 2006; Castellano et al., 2007; da Fontoura Costa et al., 2007). It can be reasonably assumed that the risk of having neglected some specific result which could radically alter the framework described is small.

Like any methodology, the approach followed here has its limitations. Collection of data and estimation of its completeness, important for the elaborations performed, are delicate issues. Nonetheless, the procedures adopted and the considerations made on the influence of missing data on the network properties give a reasonable confidence that the results derived hold. A single case was examined, therefore the results are specific to that situation. The general methods discussed, however, have a much broader applicability.

Being interested in the topological characteristics of a tourism destination network, and in its dynamic evolution, we disregarded the functioning of the single elements (the network actors), as well as the influence this may have had on the properties studied here. The analysis of these effects needs a further series of investigations and a further development of the methodology that go beyond the objectives of this work.

With this in mind, it is worthwhile to recall here the main assumptions made in conducting this research:

- since only topological characteristics of the networks are studied, no reference to possible specific functions of the single network elements is made. All the elements (actors) are considered equivalent;
- all connections are supposed undirected (symmetrical) and unweighted unless explicitly stated in specific parts or for specific analyses;
- the network nodes are organisations, companies and associations, and the links studied are considered to be connecting these elements, not parts of them (single persons or smaller parts);

- the network and the system are studied at a macro level. Therefore all characteristics derived have the same scale.

Some minor limitations or more restrictive hypotheses, needed for specific computations, have been presented when required.

Most initial studies of real world complex networks use these assumptions (Albert & Barabási, 2002; Boccaletti et al., 2006; Newman, 2003b). In this case their main function is to reduce the weight of the computations. This way it is possible to better direct efforts into interpreting the results and their implications for the management of a tourism destination.

The author considers that the acknowledgment of these boundaries and limitations adds focus to this research program and is able to provide an effective possibility of replication of the methodology proposed.

5.7 Ethical issues

The empirical collection of data for this work involves also human participants, therefore the ethical concerns related with this type of research must be addressed (Veal, 2006; Vessuri, 2002). The University of Queensland guidelines for the ethical conduct of research have been considered: Policy No. 4.20.1 - Research Ethics Policy, and Policy No. 4.20.2 - Procedures for the Conduct of Research (UQ HUPP, 2002). In the present case, the most critical issues are the informed consent and the confidentiality of the data obtained.

According to the guidelines, the following have been adopted:

- all the participants are informed of the reasons and the context of the research;
- participants are made aware that their contribution is purely voluntary and they can withdraw at any moment;
- participants are informed of their privacy rights and of the protection measures taken to guarantee confidentiality. This is done in agreement with the strict regulations imposed by Italian legislation (Laws No. 196/2003 and No. 675/1996);
- information sheets are provided and a written consent form is signed by the participants;

- participants are given the option to request a summary of the final results.

The ethical clearance for the investigation was requested and obtained by the Ethics Officer of the School of Tourism of The University of Queensland (Ethical Clearance Number TALM 25, 21st September, 2006, see Appendix 2).

5.8 Conclusions: the methodological approach

This chapter has described the methodological path followed for the study of the network of a tourism destination. The process can be summarised as follows:

- a case is chosen. Its function, in the present work, is to be a field test for the development of the methods used;
- data are collected. This implies the definition of the destination actors and the linkages among them. The sample collected is evaluated and, if not complete, the implications of the incompleteness for the subsequent measurements are assessed;
- a network graph is built and analysed;
- the static topological features of the graph are calculated;
- dynamic aspects of the topologies are characterised. Main diffusion processes are simulated;
- the network is optimised with respect to efficiency parameters in information or knowledge transfer and the results are compared with those obtained for the non-optimised network;
- evolution models for the destination network are assessed and compared with the 'traditional' measures of tourism destination development.

All the results derived are discussed for the influences they may have in the management of the destination and in fostering its development. Moreover, the methods and techniques used are evaluated, within the boundaries set, with respect to their usability and usefulness as management and knowledge tools.

Table 5.6 summarises the main activities performed and the methods used in answering the research questions.

Table 5.6 Summary of the methods used

Research question	Activity performed	Methods used
RQ1: What is the structure of the tourism destination network?	Static topological analysis	Calculation of main network metrics (section 5.5.1)
RQ2: How can we build an evolutionary model that reproduces the present structure?	Comparison of actual topology with existing network evolution models	Calculation of network metrics and degree distributions at different times (section 5.5.2)
RQ3: How does the network structure affect dynamic processes?	Analysis of network robustness and of information diffusion	Edge deletion and recalculation of network efficiency; SI epidemiological infection model (section 5.5.2)
RQ4: How can the network be optimised?	Efficiency optimisation via rewiring	Simulated annealing dependent on network efficiency (section 5.5.2)
RQ5: How does this model relate to the history of the destination?	Reconstruction of destination network at a past time and analysis of the evolution to present time	Calculation of network metrics and degree distributions at different times (section 5.5.2)
RQ6: How do the real and technological networks compare in a tourism destination?	Comparison of 'real' and 'virtual' networks	Calculation of metrics and degree distributions of both networks, assessment of the topological similarity (section 5.5.4)

6 Results

This chapter contains the results of the analysis performed on the Elba destination following the methods described previously (Chapter 5). The main objective of this thesis was to collect empirical evidence to answer the research questions stated at the beginning of this thesis. This chapter provides also an assessment of the validity and viability of the methodological approach used. A full discussion of the results presented here and of their interpretation is given in the next chapter (Chapter 7).

Following the analysis path described in Chapter 5, Figure 5.1, the first section of this chapter reports the main metrics calculated for the stakeholder network data collected. These quantities allow the structural properties of Elba's destination network to be described. Next, the interviews conducted in the destination (using the method discussed in section 5.4.2) are used to provide an estimate of the completeness of the data collected.

In section 6.3 the virtual network formed by the websites belonging to the tourism operators is discussed. The main metrics for this virtual network are calculated by using the same techniques and algorithms used for the destination stakeholder network and a comparison is made of the two resulting topologies and an assessment of their similarity.

Section 6.4 describes the results of the dynamic study. The robustness characteristics of the Elban tourism system are analysed by using a simulation method discussed in section 5.5.2. In this method the efficiency parameters of the network are calculated while removing different sets of edges from the network. Next, an information diffusion model is employed to simulate the behaviour of the network in two different conditions. In the first case a non-uniform distribution of the single actors' capacities to transmit the information received is assumed. In a second simulation the diffusion is performed by using an optimised distribution of the network's linkages in which the clustering (and efficiency) characteristics are increased (while the actors' capacities are held equal).

Finally in section 6.5, the historical evolution of the tourism destination is synthesised from the outcomes of the qualitative investigation (interviews conducted at the destination) and the results are used to infer the network metrics at different times in order to obtain a dynamic network growth model.

6.1 The Elba network, a static characterisation

The reciprocal relationships of the almost one thousand tourism operators located in Elba are graphically depicted in Figure 6.1. The picture shows a dense connected central core surrounded by a sparsely linked set of nodes. A significant number of isolated elements completes the network.

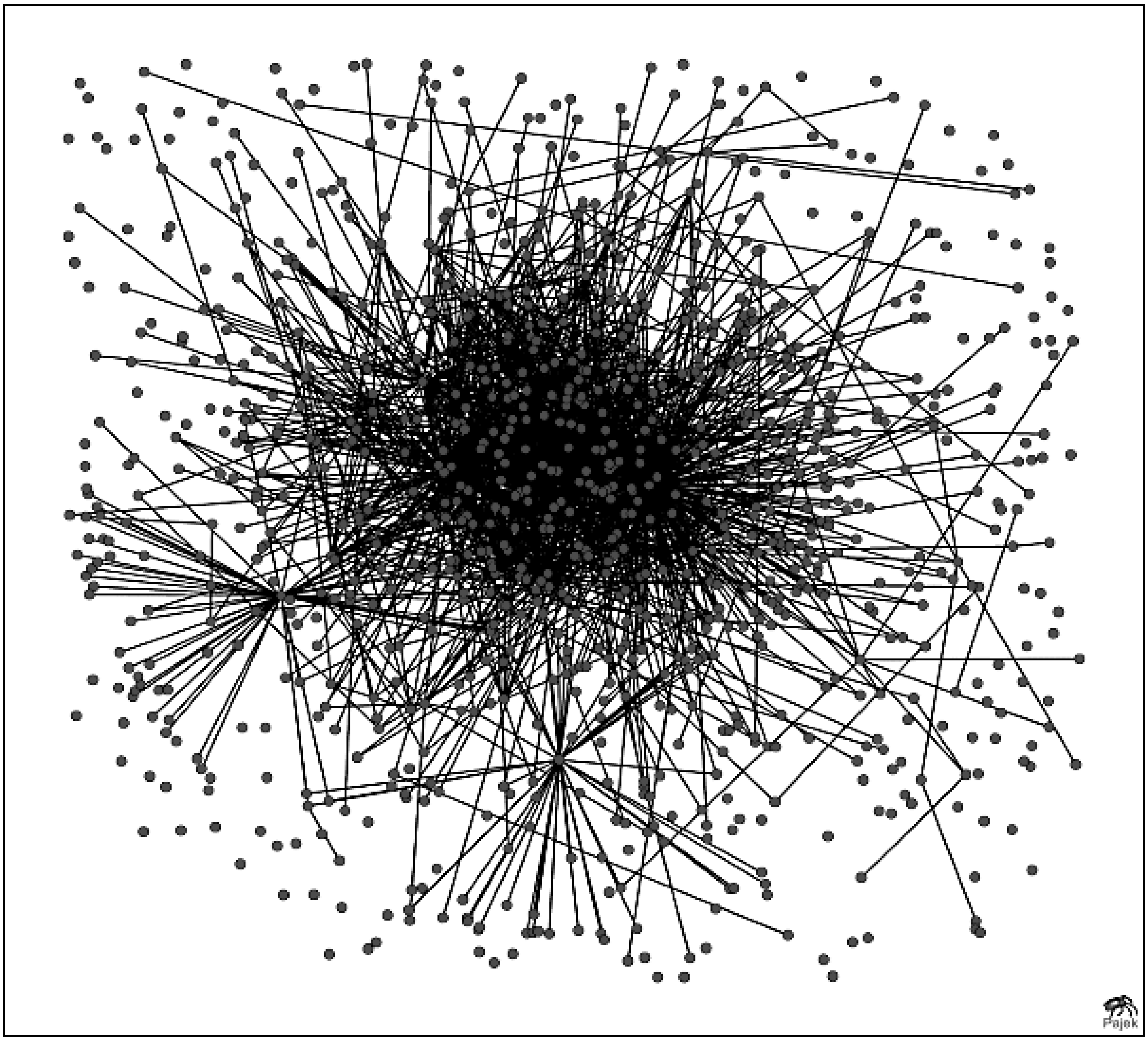


Figure 6.1 The network of Elban tourism operators

The dense central core is shown enlarged in Figure 6.2.

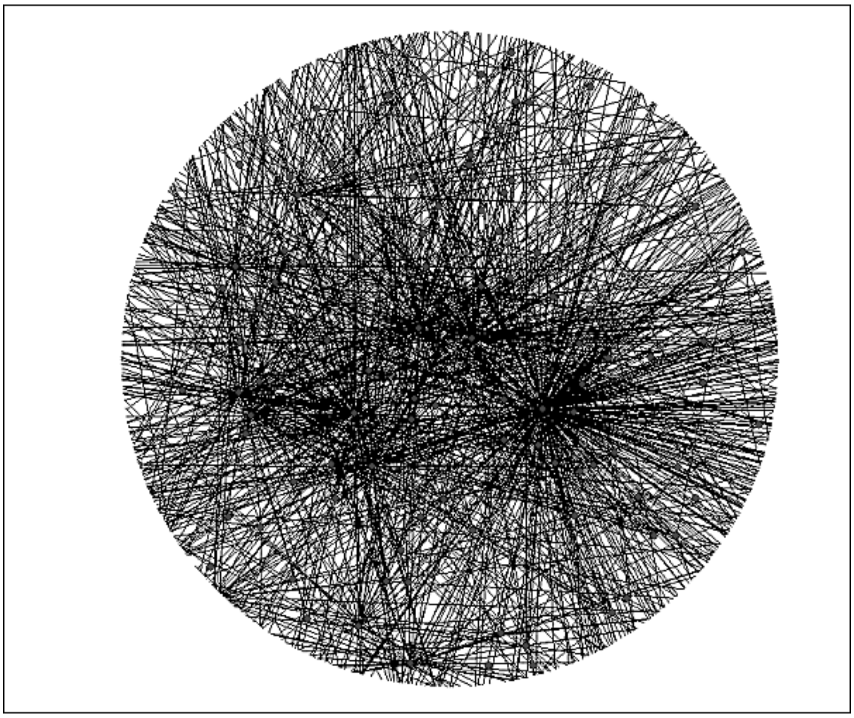


Figure 6.2 A close-up of the central region shown in Figure 6.1

Observation indicates that the network clearly exhibits a ‘structure’ and this can be highlighted by comparing the previous pictures with Figure 6.3 which shows a network of the same size (both number of nodes and edges) with a random distribution of the links.

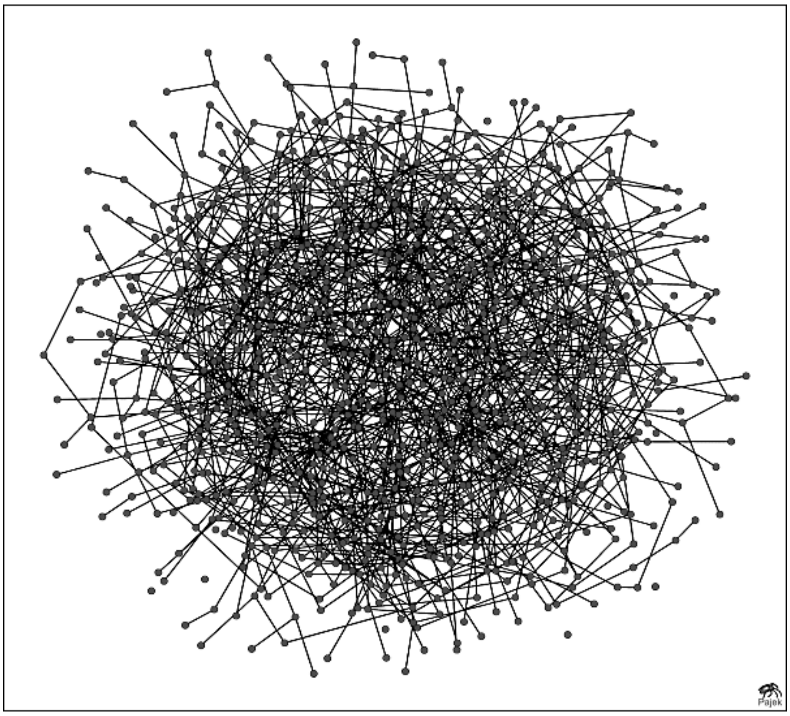


Figure 6.3 A network with a random distribution of links

The main static metrics for the Elba network were calculated following the methods described previously (Chapters 2 and 5). The network is considered to be undirected, i.e. once a link between two actors has been identified, this is considered to be a symmetrical connection. The values are shown in Table 6.1. Most of these metrics do not have an ‘absolute’ meaning, but can provide useful information on the differences between the network structures found and a reference model (null model). A frequently used null model is the ensemble of graphs with the same number of nodes and links, in which the links are distributed randomly. In what follows, the values for the reference model were calculated by averaging 10 realisations of such a *random* network. The last column of Table 6.1 shows also the order of magnitude of some significant values pertaining to real social networks as reported in the literature (Boccaletti et al., 2006; Dorogovtsev & Mendes, 2002; Newman, 2003b).

Table 6.1 The values of the main metrics for the Elba network (*random* values are averages over 10 realisations)

Metric	**Elba network**	**Random**	**Social networks**
No. of nodes	1028	1028	
No. of links	1642	1642	
Density	0.003	0.003	10^{-1} - 10^{-2}
Disconnected nodes	37%	3%	
Diameter	8	13	10
Average path length	3.16	5.86	10
Clustering coefficient	0.050	0.003	10^{-1}
Proximity ratio	34.09	N/A	10^{2} - 10^{3}
Average degree	3.19	3.25	
Sum of degrees	3284	3246	
Average closeness	0.121	0.157	
Average betweenness	0.001	0.004	
Global efficiency	0.131	0.169	10^{-1}
Local efficiency	0.062	0.003	10^{-1}
Assortativity coefficient	-0.164±0.022	0.031±0.033	10^{-1} (>0)
Topological entropy	4.527	2.199	

The density of links found is quite low, considering that the values found in the literature for the social networks studied are typically of the order of 10^{-1} – 10^{-2}. Moreover, the percentage of nodes that do not have any connections with other elements of the system is very high (39%). This results in a sparse network. This sparseness is also confirmed by the small value of the clustering coefficient. The efficiency of the Elban network is consequently quite low, both at a global and a local level.

Another value which is different from what would have been expected in considering a socio-economic network such as Elba is the assortativity coefficient. This, as seen in Chapter 2, represents the tendency of a vertex to connect with vertices having similar degrees. The correlation has been found positive for many of the social networks examined by the literature (Newman, 2002b; Whitney & Alderson, 2006), and, while debated by some authors (Mitzenmacher, 2004), this positivity is generally considered to be a distinguishing characteristic of social networks with respect to technological systems (such as the Internet or the World Wide Web networks).

On the other hand, the calculated values for diameter and average path length seem to be in line with those of other real social systems and sensibly smaller than those exhibited by a random network, indicating a certain level of compactness of the Elban network, at least for what concerns its central connected core. This is also confirmed by the proximity ratio which indicates a good level of 'small-worldness' of the network (see Chapter 2).

As discussed in Chapter 2, the most striking characteristic of a complex network is the distribution of the degrees of its vertices $P(k)$. The degree distribution and its cumulative version of the Elba tourism network are shown in Figure 6.4 and Figure 6.5. The scale-free behaviour, and the power-law scaling, of the two distributions is evident; a few actors possess the majority of the connections. More than 40% of the total links are due to the most connected 20 nodes. The scaling exponent is usually calculated by using an OLS (ordinary least squares) linear regression on the log-log plot of the data. It is also customary to fit the plot obtained by logarithmically binning the data or by considering the cumulative distribution, in order to avoid an excessive weight of the points belonging to the long tail of the distribution (Adamic, 2000; Mitzenmacher, 2004; Newman, 2005). By using this technique, the exponent of the distribution $P(k) \sim k^{-\alpha}$ is $\alpha = 2.17$.

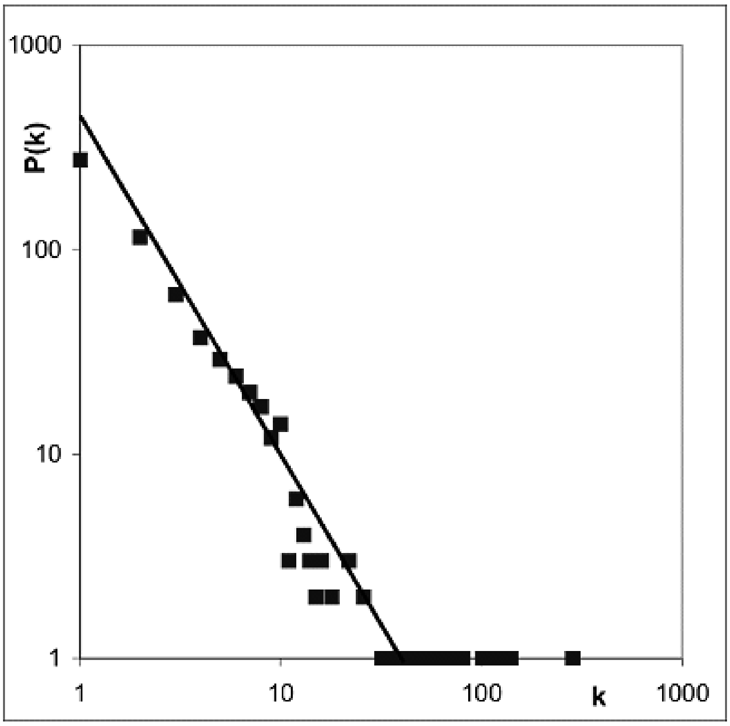


Figure 6.4 Degree distribution of the Elba network

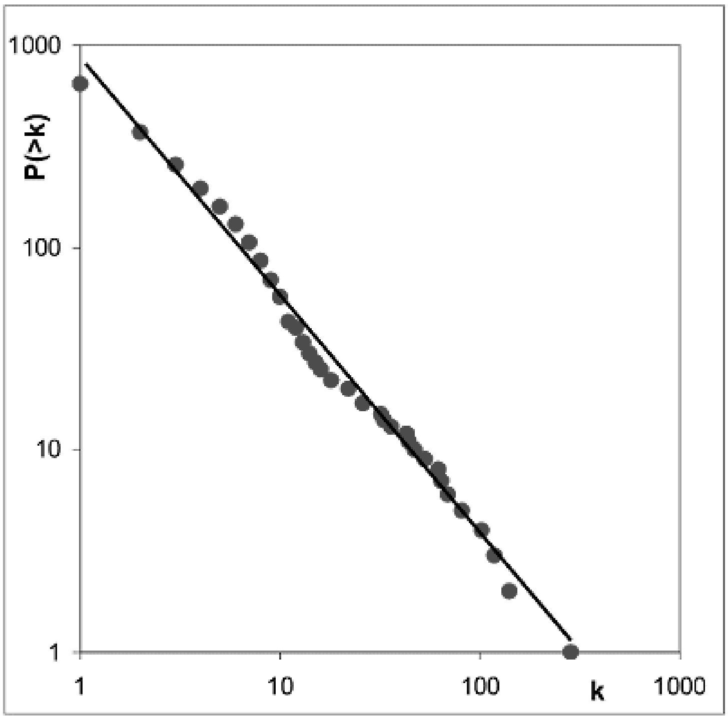


Figure 6.5 Cumulative degree distribution of the Elba network

More recently, however, some authors have argued that these fitting methods may produce systematically biased estimates of the degree distribution parameters (Clauset et al., 2007; Goldstein et al., 2004). A more sophisticated method proposed is to use maximum likelihood methods for the parameters determination and the Kolmogorov-Smirnov statistic to estimate the experimental errors. With these algorithms (Clauset et al., 2007) the value calculated in the present case is: $\alpha = 2.32 \pm 0.27$.

The spectrum analysis of a complex network can provide useful insights in its characteristics. As discussed in Chapter 2, the shape of the spectral density $\rho(\kappa)$ of a graph (i.e.: the distribution of the eigenvalues of its adjacency matrix) is known to be an indicator of the topological properties of the network. For random graphs with a giant connected component it converges to a semicircle following Wigner's law (Wigner, 1955, 1958). All other cases produce different distributions (de Aguiar & Bar-Yam, 2005): a highly skewed multi-peaked structure for a small-world network and a triangular shape for scale-free graphs. As Figure 6.6 shows, the Elban network demonstrates scale free properties.

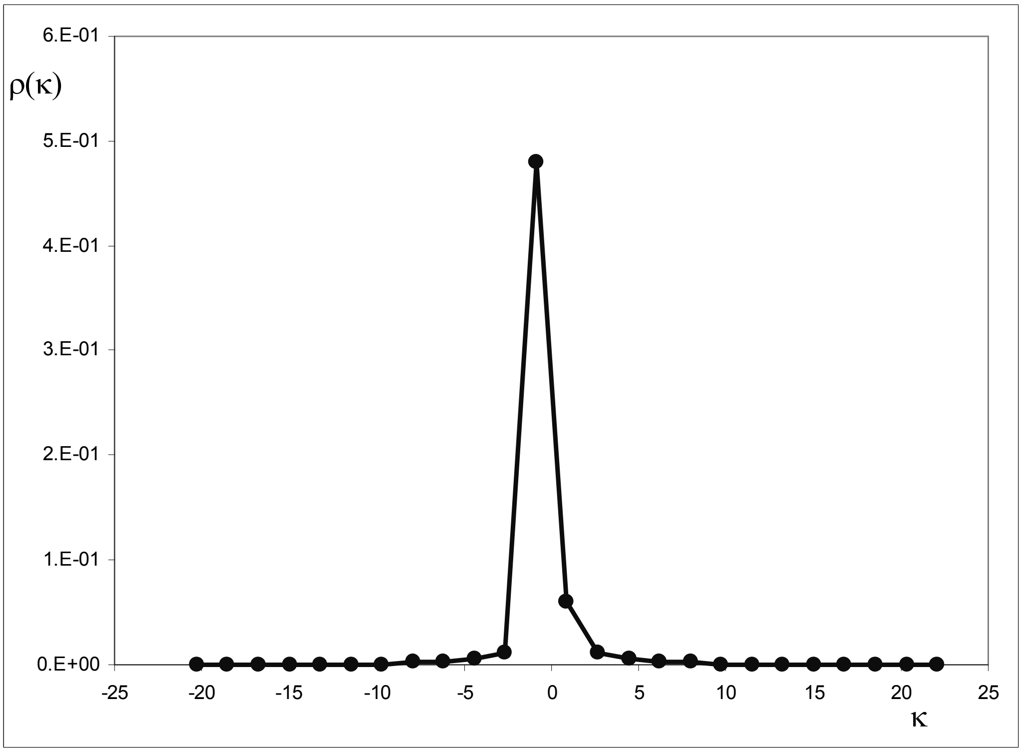


Figure 6.6 The adjacency matrix eigenvalues spectrum of the Elba network

The $\rho(\lambda)$ spectral distribution of the Laplacian matrix associated with the network is shown in Figure 6.7. It can be noted that a high number of the Laplacian eigenvalues are null. This provides one more clear indication of the scarce connectedness of the network. The multiplicity of the null eigenvalue, in fact (see section 2.3.4 and 2.6.2), corresponds to the number of connected components of the network.

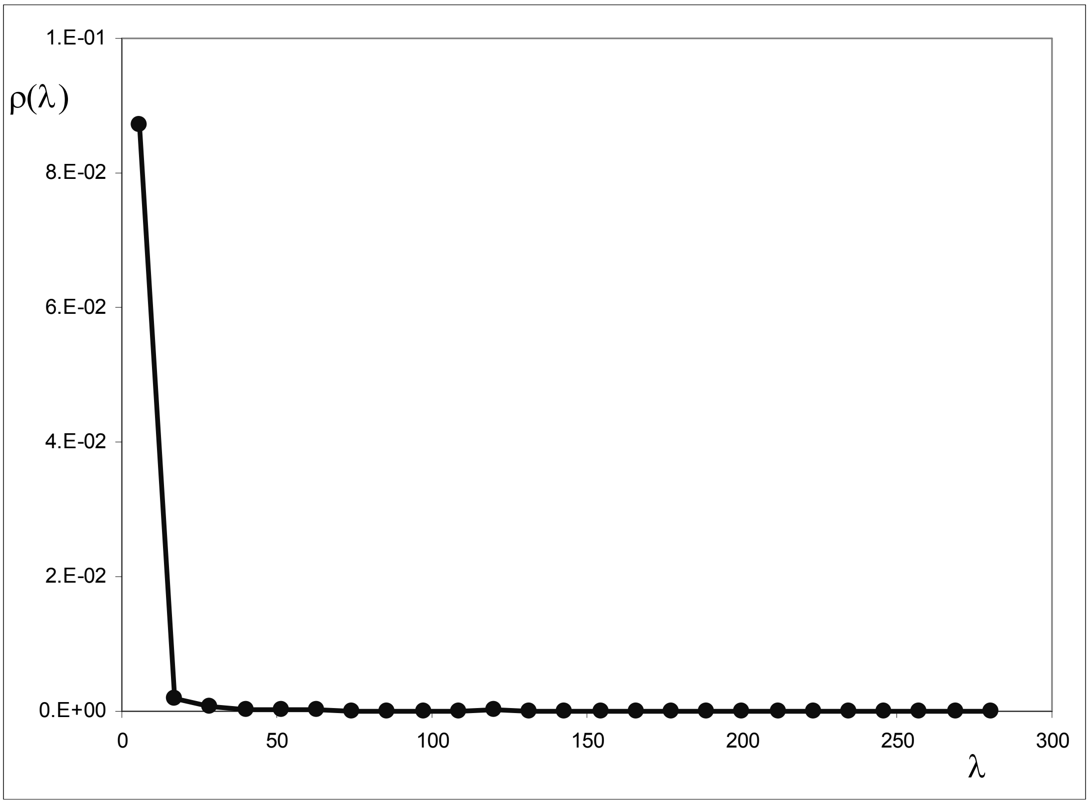


Figure 6.7 The Laplacian matrix eigenvalues spectrum of the Elba network

The modularity (see sections 3.2.5 and 5.5) of the network was calculated by dividing its actors with respect to the type of business (hospitality, associations, food and beverage services etc.) and to the geographical location (the eight Elba municipalities). As a comparison, the modularity was investigated by using the Newman-Girvan algorithm as modified by Clauset, Newman and Moore. (2004). The results are shown in Table 6.2. The table shows the number of clusters identified (groups) and the modularity index. CNM indicates the calculations were performed following the Clauset-Newman-Moore algorithm. The last row reports the values calculated (CNM) for a network of the same order as the Elba network with a random distribution of links (values are averages over 10 realisations). To better compare the different results, the last column of the table contains the average modularity over the groups (modularity/number of groups).

Table 6.2 Modularity for the Elba network

Grouping	No. of groups	Modularity	Average Modularity
Geography	9	0.047	0.0052
Type	8	-0.255	-0.0319
CNM	11	0.396	0.0360
CNM (random)	23	0.606	0.0263

All networks have a very low modularity. In one case (grouping by type), the negative value indicates that the actors tend to have more connections outside the group to which they belong than with similar businesses. The higher values found by the CNM algorithm confirm that division by geography or by type of business does not imply any form of clusterisation in these groups. The values obtained for the random network, however, are a further confirmation of the very low degree of modularity of the Elba network.

6.2 Data completeness evaluation

As discussed in Chapter 5 (section 5.4), the completeness of the data collected is an important issue as it may affect the validity of the different measurements in a completely different way than in the case of a 'randomly' distributed characteristic of a population. It has been shown that sampling some of the network elements (nodes or edges) in different ways can provide different results. This is even more important when the network under consideration exhibits a scale-free topology as in our case.

Two aspects must be considered: the completeness in collecting the vertices of the graph and the completeness in listing the links among them. In the first case, the data collected can be considered nearly complete. Looking at Table 5.4 (section 5.4.2) we may conclude that the set examined contains almost all the elements (about 95%). The uncertainty in the official listings, at most, concerns a limited number of very small businesses which can be reasonably assumed not to influence in a sensible way the calculation of the main network parameters.

The connections among the operators were inspected by conducting the interviews mentioned in section 5.4.2. With respect to the initial data collection, these interviews have revealed a very limited number of connections that were not discovered at first. Even if quantifying this difference and extending this estimate to the whole network is very difficult, it seems reasonable to assume that the final layout has a completeness of about 90%.

Conservatively, we may therefore assume a 10% confidence level for the completeness of the whole ensemble (nodes and edges). Comparing this estimate with the results reported in the literature on the subject (Han et al., 2005; Kossinets, 2006; Lee et al., 2006; Rafiei & Curial, 2005; Smith et al., 2003; Stumpf et al., 2005a; Stumpf & Wiuf, 2005; Stumpf et al., 2005b) it is possible to conclude that the calculations of the network parameters are not affected in sensible ways by this incompleteness.

6.3 The technological network

The network formed by the websites belonging to the Elban tourism operators has been studied as described in section 5.5.4 (see also Baggio, 2007, Baggio et al., 2007a).

The shape of the webgraph is shown in Figure 6.8. The network is considered here a directed network. The scale-free structure is, again, clear and is further confirmed by the inspection of the distribution of the degrees. Both the in- and out-degree cumulative distributions are shown in Figure 6.9. They follow a power law $P(k) \sim k^{-\alpha}$; the exponents calculated are: $\alpha_{in} = 2.96$ and $\alpha_{out} = 1.89$.

The main characteristics of this network are listed in Table 6.3. The values are compared with common values published in the literature for the World Wide Web (Albert & Barabási, 2002; Boccaletti et al., 2006; Caldarelli, 2007; Dorogovtsev & Mendes, 2002, Newman, 2003b).

In general, it is noted that the general connectivity (density, clustering coefficient, proportion of disconnected elements, efficiencies) indicate a high sparseness of the webspace for the Elban destination.

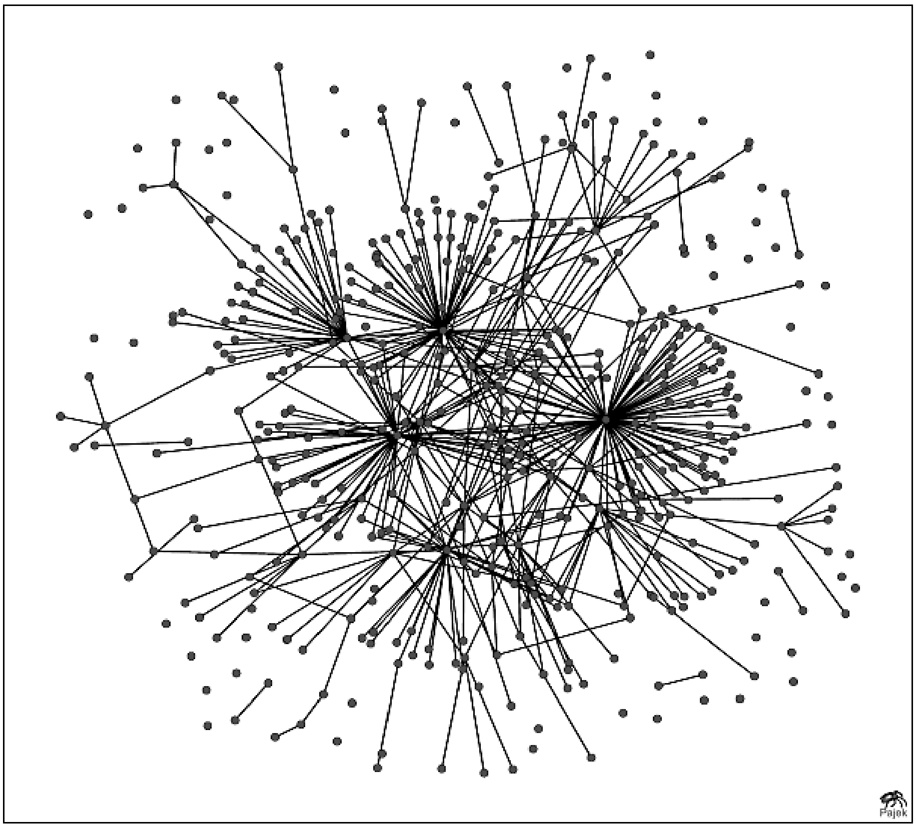


Figure 6.8 The Elban web network

Table 6.3 The main characteristics of the Elba web network compared with common values for the whole WWW

Metric	Elba	WWW
No. of nodes	468	
No. of links	507	
Degree distribution		
In-degree	2.96	2.1
Out-degree	1.89	2.7
Density	0.002	$O(10^{-1})$
Disconnected nodes	21%	DCC < 9%
Average path length	4.5	16
Diameter	11	28 (in SCC)
Clustering coefficient	0.003	0.11
Local efficiency	0.0145	0.36
Global efficiency	0.1698	0.28
Assortativity coefficient	-0.101 ± 0.094	$O(10^{-1})$ - negative

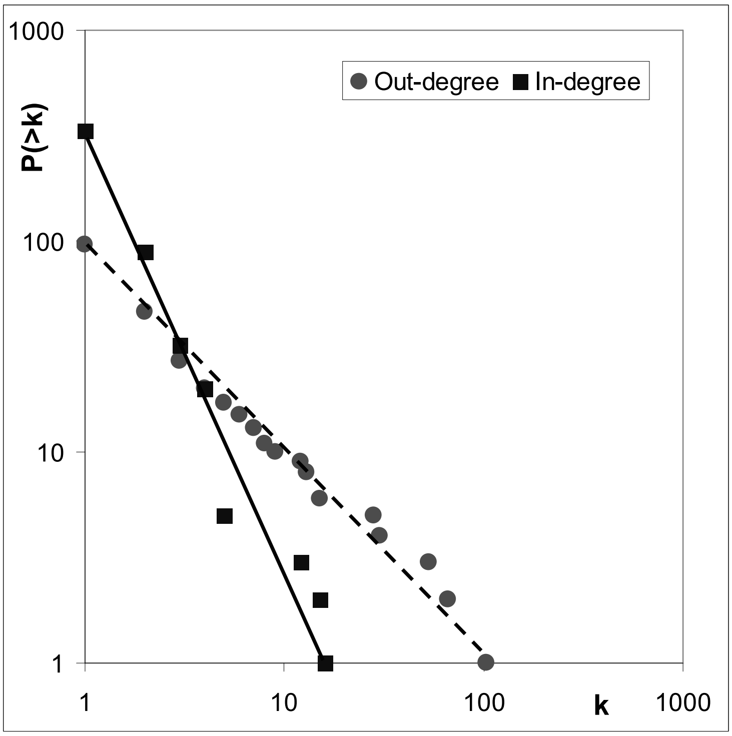


Figure 6.9 In- and out-degree cumulative distributions for the Elba web network

The in-degree distribution exponent is higher than the one measured for the Web: $\alpha_{in} \sim 2.1$ (Broder et al., 2000). This implies a greater concentration. On the contrary, the out-degree distribution exponent is much lower (typical Web value is $\alpha_{out} \sim 2.7$); i.e. the distribution of the links is much flatter and more spread.

According to the current understanding, the WWW is thought to exhibit a macro-structure known as a bow-tie (see section 3.3). By applying this model to our network we obtain the results shown in Table 6.4.

Table 6.4 Bow-tie components calculated for the Elba Web and general values for the WWW (as reported by Broder et al., 2000; Kumar et al., 2000b)

Bowtie components	**Elba WebNet**	**WWW**
SCC	3%	28%
IN	2%	21%
OUT	52%	14%
TUBE	1%	7%
TENDRIL	16%	9%
DCC	25%	

The comparison with the global values reported for the WWW (Broder et al., 2000; Kumar et al., 2000b) shows again that the general connectivity of the tourism destination websites is much lower.

6.3.1 The technological and socio-economic networks: a comparison

Both networks, the Web and the socio-economic network of Elban tourism stakeholders are expressions of the same system: the Elban tourism destination. It is then interesting to compare the characteristics of the two graphs to investigate possible similarities. In fact, as discussed in the previous chapters, a strand of literature argues that the connections among the websites (hyperlinks) may be considered not simply as a technological manifestation but also (and probably mainly) as a reflection of social processes. The structure of hyperlinks, it is argued, forms patterns based on the designs and aspirations of the individuals or organisations who own websites. A growing literature suggests that these networks reflect offline connections among social actors and support specific social or communicative functions (Jackson, 1997; Park, 2003; Park & Thelwall, 2003).

Even considering the warnings and the limitations on the validity of this type of interpretation, as discussed by Thelwall (2006), in the case of commercial companies' websites, it is reasonable to assume, also for the importance given to the practice of hyperlinking (Walker, 2002), that the layout of the network is a reflection of the structural characteristics of the social network from which it originates. This relationship between cyberspace and the physical world is two-way: on one side, the online linkages represent and complement social relations in the offline world; on the other side, offline interactions can influence the way in which online relationships are established and developed (Birnie & Horvath, 2002; Wellman, 2001).

Table 6.1 shows a comparison between the metrics calculated for the Elban networks. As it can be seen, apart from scale factors, most of the values have differences which are lower than an order of magnitude.

Table 6.5 Comparison between the main network metrics for the socio-economic (TN) and the web (WN) networks of the Elba tourism destination

Metric	TN	WN
Number of nodes	1028	468
Number of edges	1642	495
Density	0.003	0.005
Disconnected nodes	37%	21%
Diameter	8	10
Average path length	3.16	3.70
Clustering coefficient	0.050	0.014
Degree distribution exponent	2.32	2.17
Proximity ratio	34.10	12.21
Average degree	3.19	2.12
Average closeness	0.121	0.155
Average betweenness	0.001	0.003
Global efficiency	0.131	0.170
Local efficiency	0.062	0.015
Assortativity coefficient	-0.164	-0.167

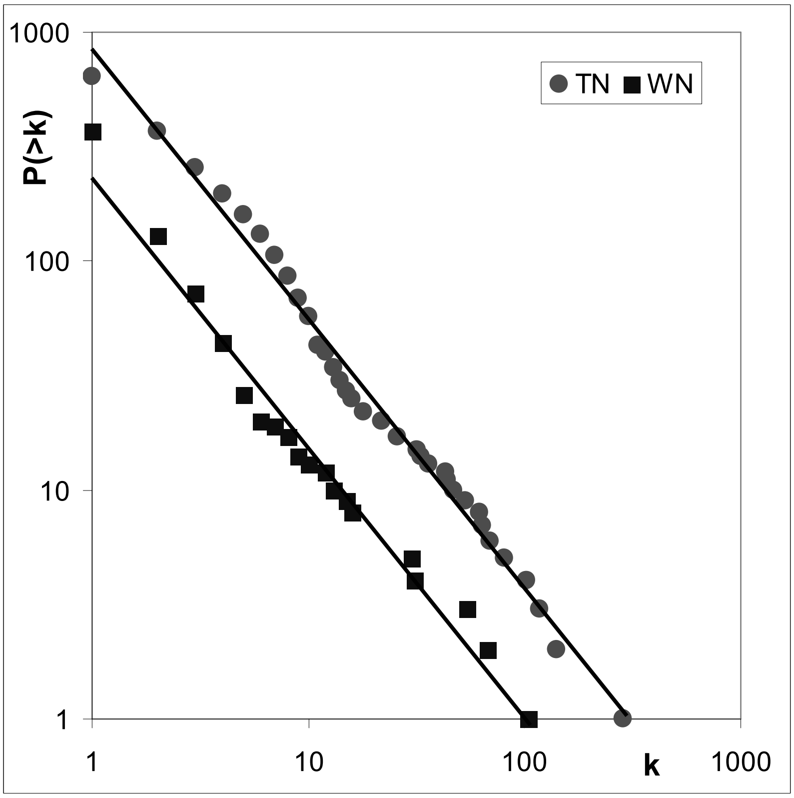


Figure 6.10 Cumulative degree distributions for the tourism stakeholders (TN) and the tourism websites (WN) networks

Another important indicator is the degree distribution which is commonly regarded as a signature of the network topology. The cumulative degree distributions are shown in Figure 6.10. Considering the undirected version of the web network for compatibility with the real one, the exponents of the power-law degree distribution calculated from this data are α_{TN} = 2.32±0.269; α_{WN}: 2.19±0.109 (calculations were performed according to Clauset et al., 2007). Thus, they can be considered identical within the statistical uncertainty of their determination.

It is known (see Chapter 2) that in most cases, the various quantities characterising the topology of a complex network can hardly be regarded as normally distributed, and the simple comparison of their averages (arithmetic means) may look insufficient.

In these cases, as already proposed by some authors (Clauset et al., 2007; Leskovec & Faloutsos, 2006), the Kolmogorov-Smirnov (KS) statistic is quite effective. The KS D-statistic gives the maximum distance between the cumulative probability distributions of the empirical data *F(x)* and *G(x)* over the entire *x* range:

$$D = \max_x \left| F(x) - G(x) \right|$$

The statistic is nonparametric and it is insensitive to scaling issues. It compares only the shapes of the empirical distributions (Siegel & Castellan, 1988).

Table 6.6 The Kolmogorov-Smirnov D-statistics for the different network metrics (WN = web network, TN = real tourism destination network, RN = random network)

Metric	**WN vs. TN**	**RN vs. TN**
Degrees	0.119	0.147
Clustering coefficient	0.147	0.178
Closeness	0.044	0.083
Betweenness	0.030	0.077
Local efficiency	0.125	0.184

Table 6.6 shows the values for the D-statistics calculated when comparing the quantities of the Web network with those of the real network (WN vs. TN). As a reference, the same values were calculated for a random sample of the same size as WN, extracted from the real network (RN vs. TN: the values are averages over 10 realisations). The consistently lower values of the D-statistic in the case of the Web network can be considered as a good

confirmation of the similarities of the two topologies (Clauset et al., 2007; Leskovec & Faloutsos, 2006).

6.4 Dynamic processes

Networks seem to be natural candidates for computer simulations. They can be modelled with sound mathematical techniques (graph theory) and they are a sufficiently general object with which we represent a wide variety of natural and artificial systems of different complexity. Moreover, as has been seen previously (section 3.2), many dynamic processes can be defined and studied, all of them quite interesting for scientists and practitioners.

The first simulation made here concerns the transfer of information or knowledge across the network. The objective here is to assess the present situation and to test the capability of the destination network in absorbing the knowledge transferred when changing some of its structural parameters.

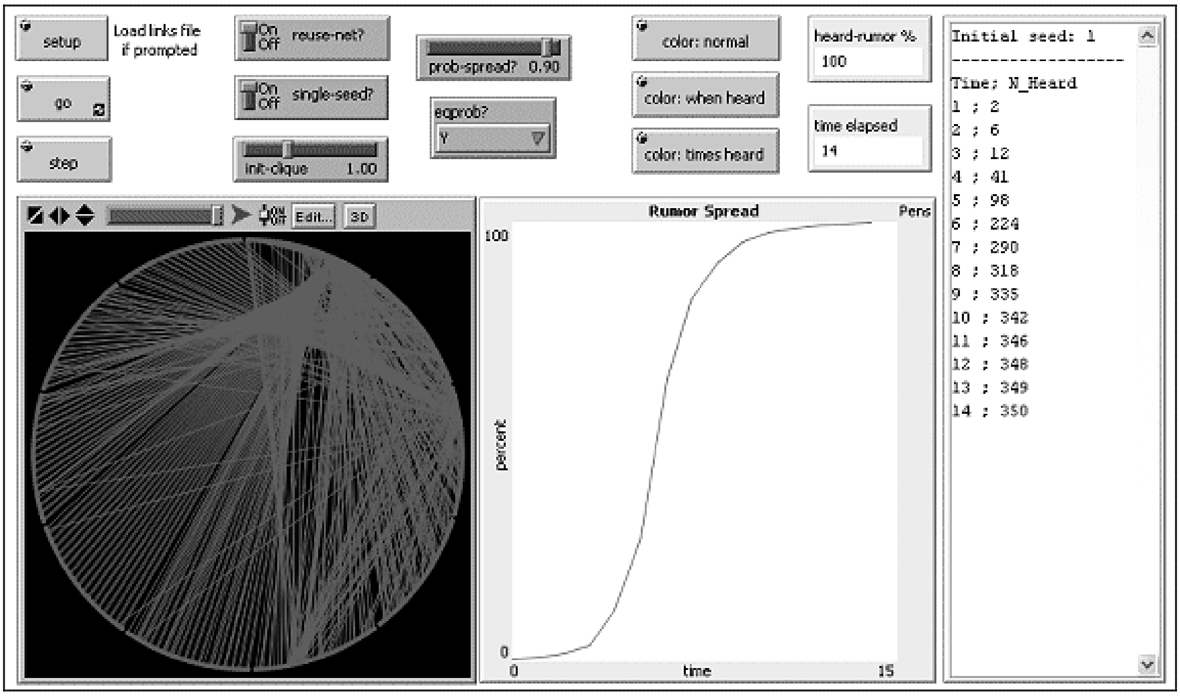


Figure 6.11 The NetLogo diffusion model

The NetLogo simulation (Figure 6.11) has been set according to the description in section 5.5.2. It may be useful to remember here that the model allows the variation of a number of parameters: the initial rumour can be spread by a single element or by a certain number of elements. At each step a node can tell the rumour to all of its neighbours or only to a certain percentage of them. At the end of a simulation run, it is also possible to visualise how often the nodes have been told the rumour or how recently they have been aware of it.

Two series of runs were performed: the first one considered the network in its basic configuration and the actors were divided into three categories according to their estimated size (large, medium and small). The hypothesis, in this case, was that the capability to absorb and transfer information or knowledge is connected with the size of the organisation considered (Inkpen & Tsang, 2005; Tsai, 2001). The numbers of tourism operators in the different classes is shown in Table 6.7. The size distribution is quite typical of Italian and European tourism destinations (EUROSTAT, 2006; ISTAT, 2006).

Table 6.7 Size distribution of Elban tourism operators

Size	%
Large	7%
Medium	16%
Small	77%

Almost arbitrarily, it is assumed that the values for the proportion of neighbours informed for each time step in the simulations run are set as: $p_{large} = 1$, $p_{medium} = 0.8$, and $p_{small} = 0.6$.

The second series of runs concerns an 'optimised' network. By using the algorithms described in section 5.5.2, the destination network was modified in order to achieve a higher clustering coefficient and local efficiency. A small fraction (5%) of randomly chosen edges was added to the network and the network was rewired until a threefold increase in local efficiency and clustering was obtained (see Table 6.8). The network used in the simulations was the main connected component of the original Elba network (i.e. the network built by excluding all the disconnected nodes after the addition of the new edges).

Table 6.8 Network parameters used in the simulation

Metric	Elba net	Optimised net
Clustering coefficient	0.084	0.274
Local efficiency	0.104	0.334

A random network (same size and density and random distribution of edges) was then used as a further comparison. All the results reported are averages over 10 realisations.

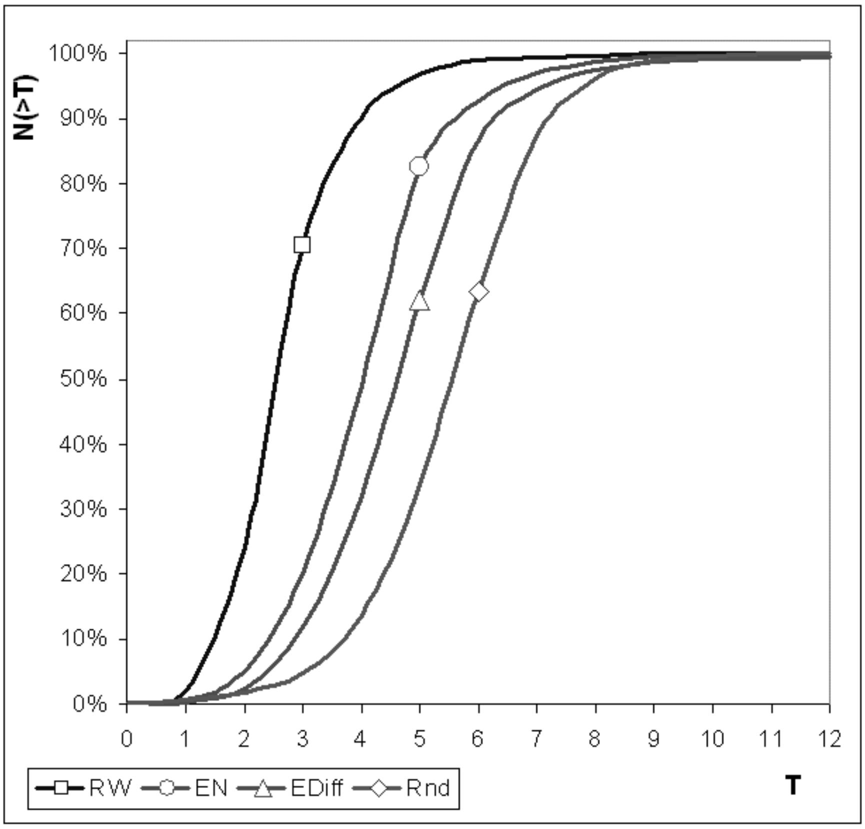


Figure 6.12 Cumulative diffusion curves (Rnd = random network, EDiff = differential actors capabilities, EN = original Elba network, RW = rewired optimised network).

The results of the simulations are shown in Figure 6.12. The curves represent the cumulative numbers of 'infected' actors as a function of time *N(>T)* for the different network configurations used.

The scale-free topology effect is evident in the difference of diffusion speed. It also makes clear the difference between the simulations obtained by assigning different capabilities to the Elban tourism operators and the one obtained by optimising the network with a threefold increase in local efficiency and clustering.

This effect is clearer when looking at Figure 6.13 which depicts the differential diffusion curves *N(T)*. The separation of the peaks is quite remarkable. If we assume that the position of this peak is a good indicator of the speed and the 'efficiency' of the whole process, it is possible to quantify its value by taking the random case as a comparison. The results are shown in Table 6.9.

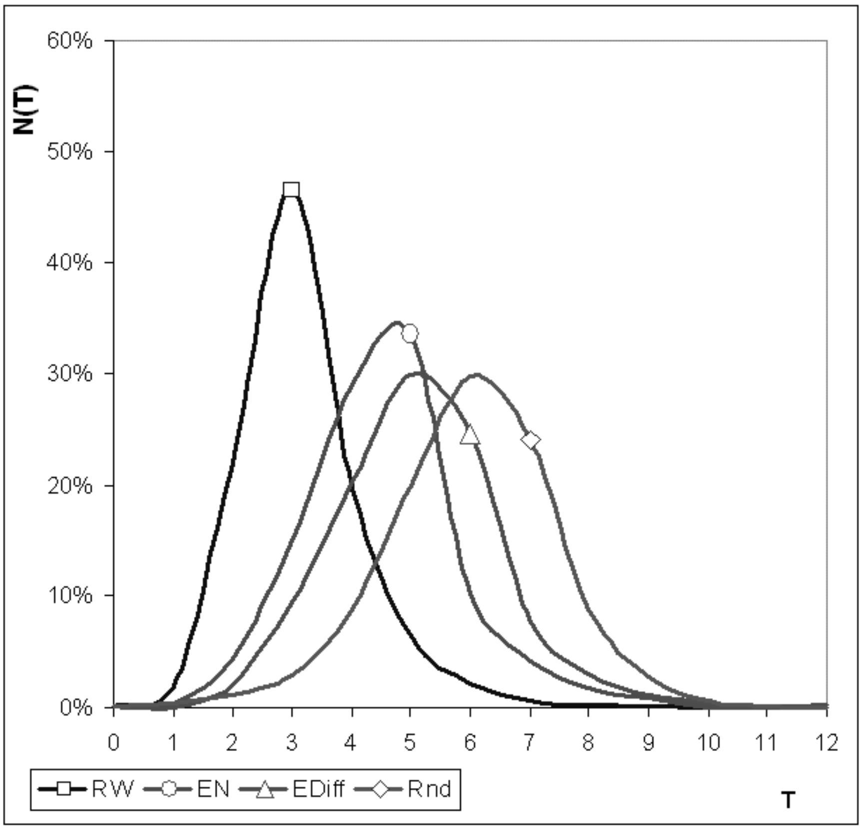


Figure 6.13 Differential diffusion curves (Rnd = random network, EDiff = differential actors capabilities, EN = original Elba network, RW = rewired optimised network).

Table 6.9 Speed improvements in diffusion over different network topologies (Rnd = random network, EDiff = differential actors capabilities, EN = original Elba network, RW = rewired optimised network).

Network	T(peak)	% improvement
Rnd	6.1	----
EDiff	5.2	16%
EN	4.8	22%
RW	2.9	52%

An important feature of a complex system such as the Elba tourism destination is the capacity to preserve its main characteristics when subject to some disrupting event. The robustness of the network can be assessed, as discussed in sections 3.2.1 and 5.5.2, by randomly removing the connections among the actors and measuring some significant parameter at each deletion. In the simulation performed, the parameters chosen were the local and the global efficiencies of the system. As described in section 3.2, the simulation applies the model and the procedure discussed by Albert et al. (2000) and Latora and

Marchiori (2001). After having calculated the local E_{loc} and the global E_{glob} efficiencies for the whole network, a random number of links (3%) were removed at each step and the parameters calculated again. The findings were compared with a similar simulation performed on a random network.

The results are shown in Figure 6.14 and Figure 6.15. In both, the fraction (*f*) of the original networks after a removal and the efficiency values calculated as a ratio over the value for the initial entire network are shown. The efficiency values are the averages over 10 realisations. As would have been expected (scale-free vs. random networks) the Elba system shows a higher robustness in cases of random removals of connections. In particular, there is a sudden transition in E_{loc} at $f \sim 55\%$ for the random network. At this value the network can be considered completely disrupted.

It must be noted, however, that the difference in the behaviours of the two networks is not as large as would have been expected (Crucitti et al., 2003). This, again, can be explained by considering the characteristics of the Elba network and mainly the fact that a *relatively flat* degree distribution makes it closer to the one of a random network.

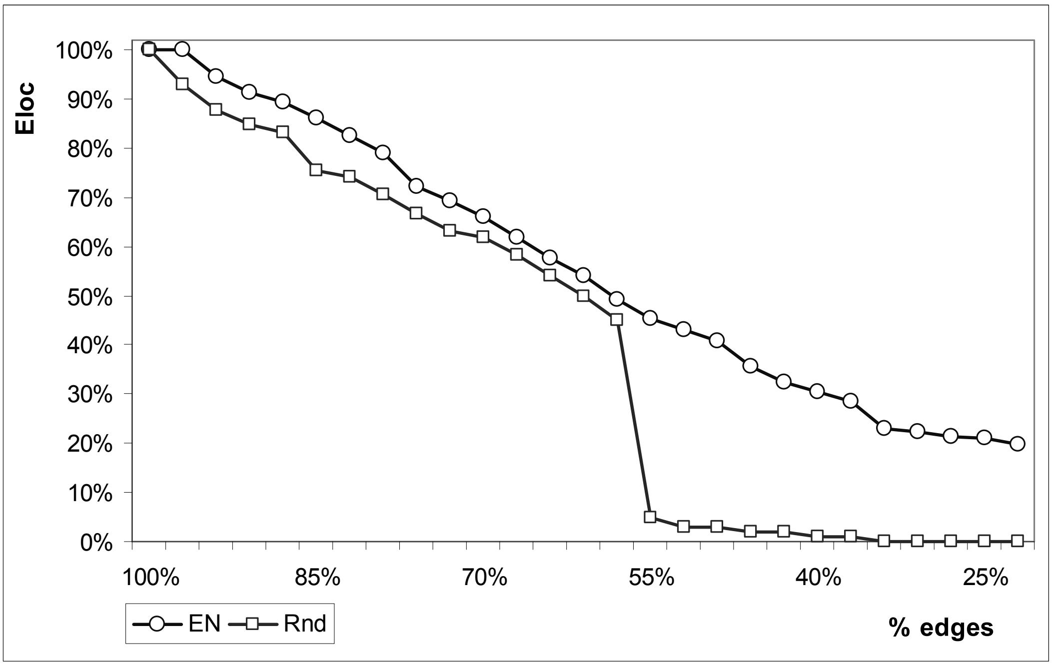


Figure 6.14 Local efficiency variation as a function of random removal of edges

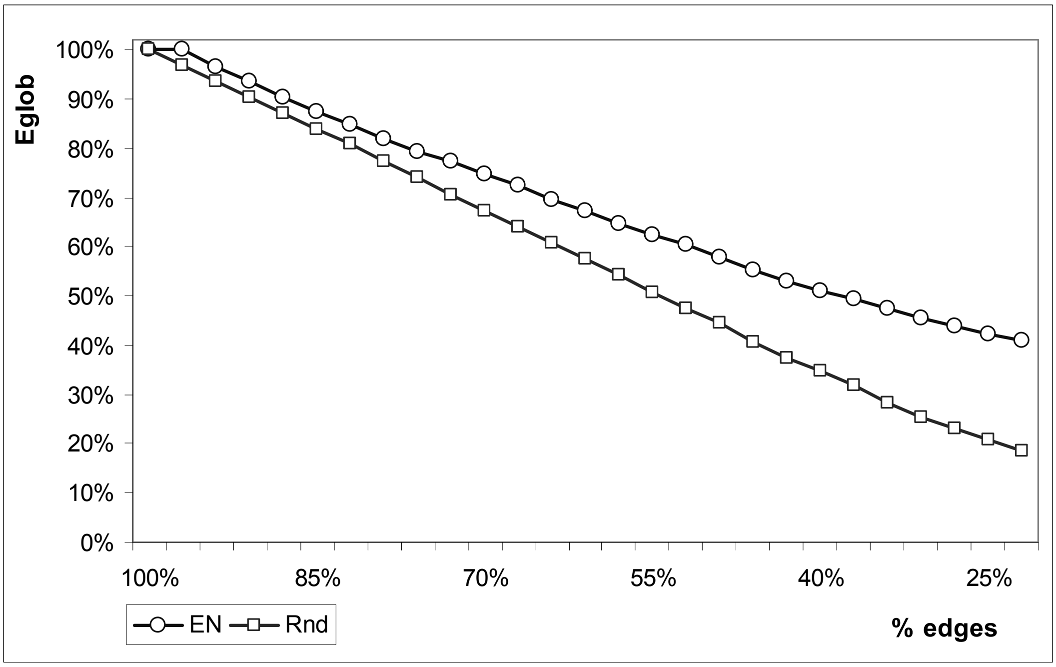


Figure 6.15 Global efficiency variation as a function of random edges removal

In conclusion, the results of the dynamic analysis reconfirm the fact, known in the literature (see for example Boccaletti et al., 2006; Durrett, 2006), that the topological structure of the network has a significant effect on both its robustness characteristics and the course of the diffusion processes which occur on it.

6.5 Historical evolution of the tourism destination network

The Elba history, seen through the discussions with the interviewees, can be analysed by looking at the tourists' overnight stays time series (Figure 6.16) and can be read through the lenses of the Butler's model (see section 4.2.3).

From the analysis of Figure 6.16 it is possible to identify three main phases. The first phase, from the mid 1950s to the mid 1980s, is an expansion phase. After a few years (up to circa 1956) in which Elba is discovered by domestic and international travellers, the tourist flux starts increasing and the supply side tries to keep the pace of the growing number of visitors, although in a relatively confused and non-coordinated way according to the interviewees. This is a golden age for the Elban tourism operators, mainly in the hospitality sector, despite their limited skills, competencies and capacity to manage their businesses. In this period the small enterprises induce the transition to an almost pure tourism economy for the whole island (Tallinucci & Testa, 2006). Practically the whole tourism infrastructure present today is built in these years.

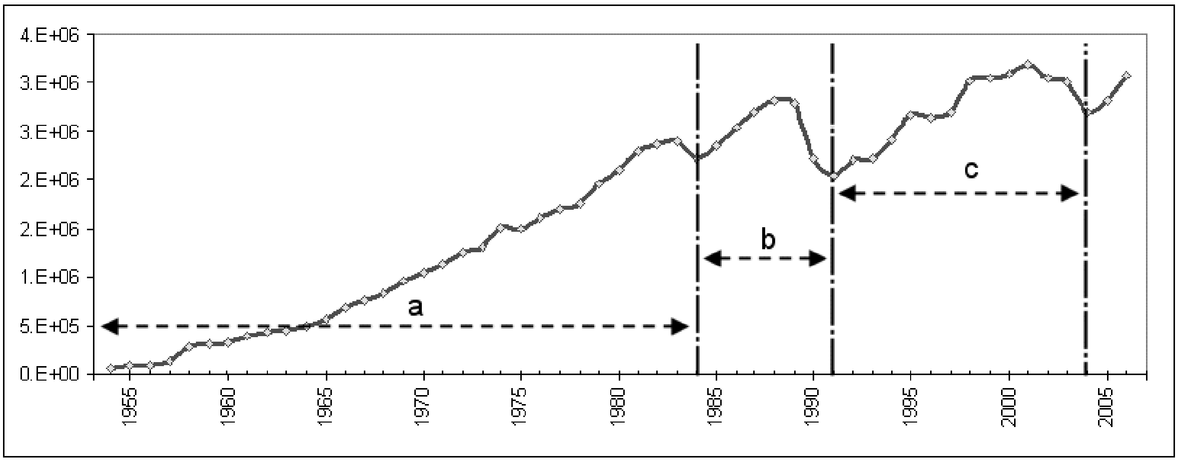


Figure 6.16 Elba overnight stays for the period 1954-2006. The three evolution phases a, b, c are discussed in the text (data source: Elba Tourism Board).

The fast growth reaches a maximum in the early 1980s. Then, a large influx and disordered crowding causes a decrease in the number of international visitors, who seem to dislike the excesses of mass tourism. This crisis induces a revision, or better a reorganisation, of the promotional activities. The local tourism board starts a restyling of the destination brand, giving focus to a deeper contact with the natural and cultural resources and putting to the background the overexploited ‘sun and sand’ image.

The second phase brings the destination to a peak, attained in the late 1980s with about 2,800,000 overnight stays, which is very close to today’s levels. During this phase some limits become evident. A very short duration of the seasonal cycle and a lack in redistribution of income highlight the weaknesses of an unplanned development for the tourism industry. This imbalance is worsened by capital investments made by non-residents which further reduce the available income distributed on the territory with respect to the one generated by the tourists. The fragmentation of the industry is very high with the vast majority of operators consisting of small family-run businesses. In the early 1990s, then, the destination faces a series of negative national and international events: an economic crisis and an increase in the Italian inflation rate, also connected with the exit of Italy from the SME (the European agreement on currencies), the German reunification process after the fall of Berlin wall, and the war in the Balcans. All these produce a drastic decrease in the numbers of domestic and international tourists, the highest proportion of which are Germans.

The Elban tourism entrepreneurs acknowledge the need for professional retraining and a diversification of the local offering with a widened perspective. The Elba Promotion Consortium is formed and is financed with European Community funds. It is the first time that the destination stakeholders recognise the necessity of a close public-private partnership. The process also generates also a series of new initiatives leading to the birth of other consortia and to the strengthening of existing associations.

This push towards collaboration and cooperation is deemed to have caused the subsequent growth phase (Tallinucci & Testa, 2006). The end of EU financing, at the beginning of the 2000s has stopped the Elba Promotion consortium operations. Moreover, in recent years, the fragmentation and individualistic culture seems to have gotten the upper hand. The poor results obtained in 2002-2004 emphasise these problems, although they seem not to be acknowledged by the Elban tourism enterprises. Very recent initiatives may help in reversing this negative phase: a new project for the design of a new market-oriented product offer, Elba-felix, financed in 2004 by a local bank (Banca di Credito Cooperativo dell'Isola d'Elba), and the constitution of a public-private permanent coordinating committee for the development of the island. The data in Figure 6.16 seem to show a positive tendency for the very last period, but, in the opinion of the interviewees, it is too early to be able to draw unambiguous conclusions.

The very brief history sketched above, allows us to consider Elba a 'mature' tourism destination (always looking at its history through Butler's model). It has a long history and has gone through a number of different expansion and reorganisation cycles.

6.5.1 A network based assessment of the destination evolution

As stated in section 5.5.3, a quantitative assessment of the historical development of the Elba network would require a precise knowledge of the configurations at different times in the past. This knowledge is not available, at least not at the level required for exact calculations, therefore the layout of the network has been estimated by considering the available data on the tourism supply and the interviews conducted as described above (section 6.5).

The period considered (see section 5.5.3) is the beginning of the 1990s (the beginning of phase c in Figure 6.16), when the destination started to recover from a period of *failure* in performance concerning the arrivals and overnight stays of tourists.

First of all, we notice that the basic infrastructure of the tourism supply was not much different in the early 1990s than it is today. Figure 6.17 depicts the historical evolution of the hospitality industry infrastructure at Elba (data source: Elba Tourism Board). As can be seen, the last substantial increase in the number of hotels and bed places dates back to the early 1980s. After that period, the supply can be considered constant.

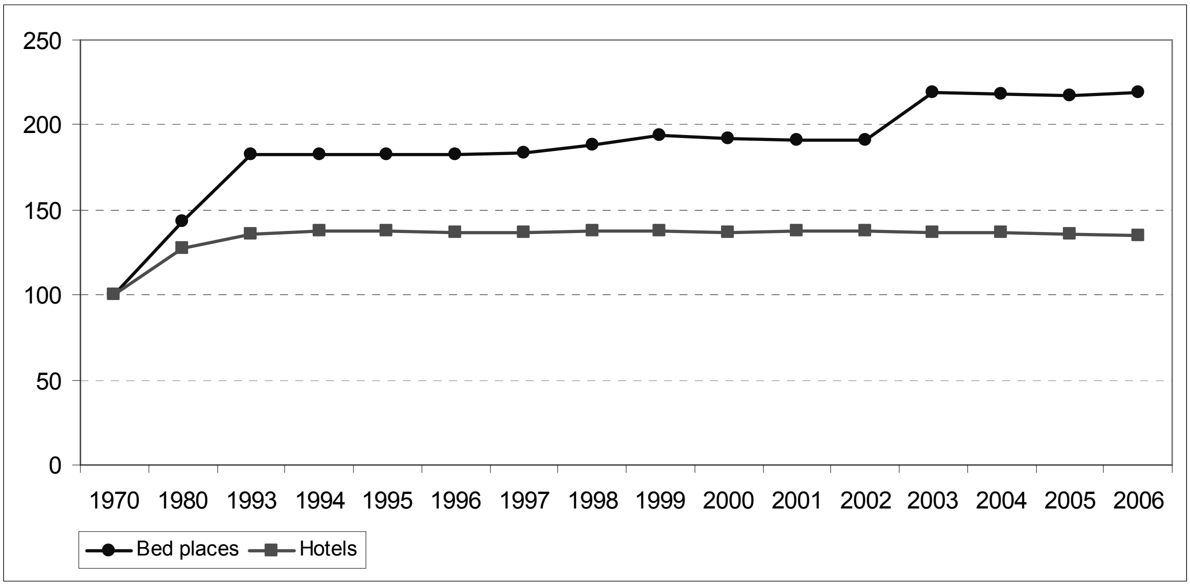


Figure 6.17 Evolution of the hospitality infrastructure at Elba (1970 = 100, data source: Elba Tourism Board)

The discussions and interviews conducted at Elba (in the period between November 2006 and March 2007) have highlighted that most of the associations and consortia present today arose in the early 1990s. This has been verified also by checking the foundation dates of these organisations in the public records (see Table 6.10).

With these elements we may formulate the hypothesis that the configuration of the 1990s system is a network with the same number of vertices. The edges of the organisations quoted above (17 associations and consortia) have been reduced to 5% of today's values. This is to account for the fact that, even if these were not existing, some part of the connections among the operators that would have formed them, was already present. The reduction has been performed by randomly choosing the edges to be removed. Ten different realisations of this presumed 1990s network were obtained and analysed. Although arbitrary and somehow questionable, this seems to be the only possible reasonable reconstruction.

Table 6.10 Foundation dates for the Elban associations and consortia (Note 1: these organisations although formally founded as shown, were greatly extended and revitalised in the early 1990s; source: public records, see text)

Association/consortium	**Year**
Progetto Benvenuti all'Elba	2004
Trek&Bike Hotels	2004
Vetrina Toscana a Tavola	2000
Associazione Albergatori	1960s-1990s [1]
Confcommercio	1960s-1990s [1]
Confesercenti	1960s-1990s [1]
FAITA Associazione Campeggi	1960s-1990s [1]
Legambiente Turismo	2003
Charme Camping	1995
Cons. Caposantandrea	1990
Cons. Costa del Sole	1998
Cons. Elba Promotion	1994
Cons. L'Elbavoglio	2002
Cons. servizio.albergatori	1994
La strada del vino - Costa degli Etruschi	2001
Consorzio Baia Blu	2005
Consorzio Laltraisola	2004

The degree distribution of the links in the original network (E_100) and the reconstructed ensemble (E_90) is taken as an indicator of the system's topology. The results are shown in Figure 6.18.

Besides a size scaling effect, the most striking difference is the cut-off at high degrees k. This is due to the limitations we have imposed on the connectivity of the network during the reconstruction process. The scaling exponents for the two parts of the distribution are $\alpha_1 = 2.22$ (the upper segment); $\alpha_2 = 3.58$ (the lower part).

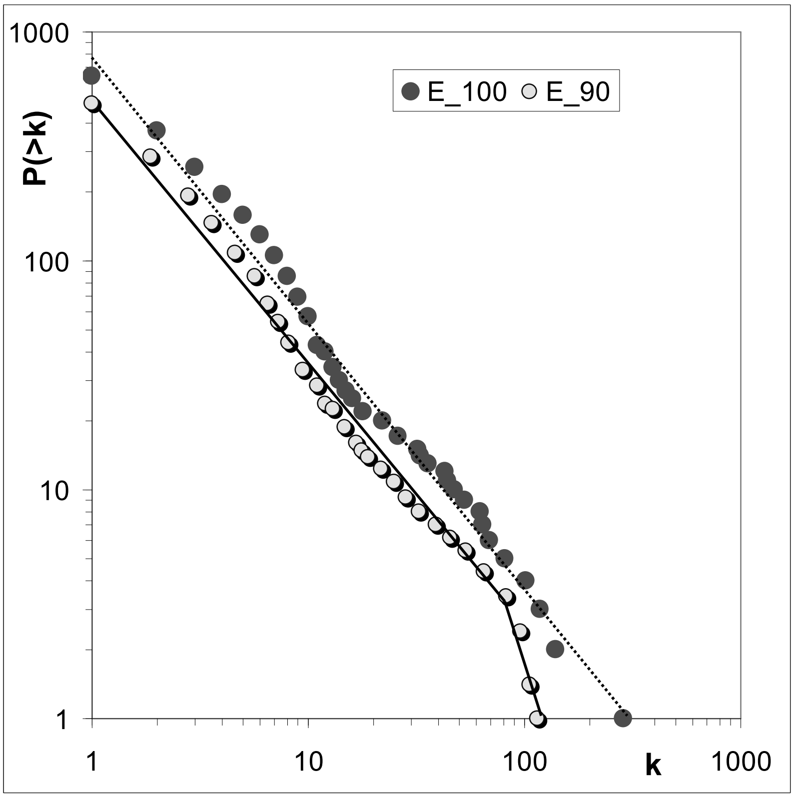


Figure 6.18 Cumulative degree distributions for the original labelled as (E_100) and the reconstructed 1990s networks labelled as (E_90)

The first part of the distribution behaviour is consistent with that of today's network; the second one is much higher, indicating a greater concentration in this segment. Since the 5% limit was arbitrarily set, a further simulation was performed by setting different reductions. The results are shown in Figure 6.19. The different limits set are: 10% (E_09), 20% (E_08) and 30% (E_07). These distributions are plotted along with the E_100 and E_90 described above. For the sake of clarity, only the lower part of the distributions are plotted.

An obvious difference in the scaling is noticeable, as is a difference in the 'bending' of the curves. The important point to notice, however, is that limitations in the ability or capacity of high degree nodes to connect other nodes creates some kind of departure from a power-law behaviour in the degree distributions generating a cut-off at high degrees. This result is consistent with some of the models presented in the literature (Caldarelli, 2007; Dorogovtsev & Mendes, 2003; Mossa et al., 2002; Pastor-Satorras & Vespignani, 2004) and will be further discussed in the next chapter.

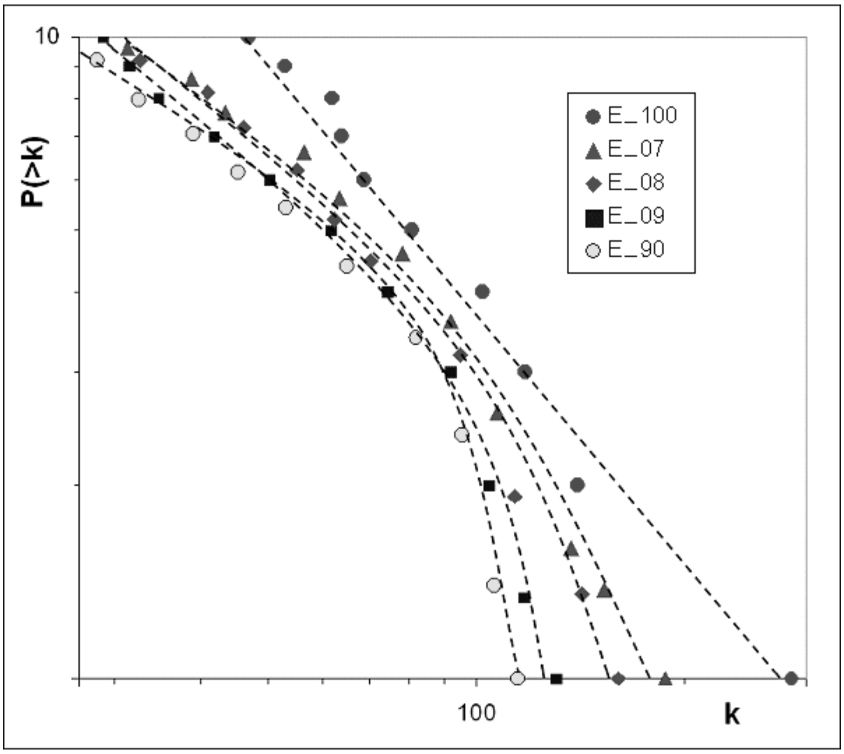


Figure 6.19 Cumulative degree distributions for the different network reconstructions (lower parts)

6.6 Conclusions: a summary of the main results

The main characteristics of the network formed by the tourism stakeholders of the Elba destination have been assessed (RQ1[11]). The main results can be summarised as follows:

- the general connectivity is very low (link density) with a very large proportion of disconnected elements;
- clustering is quite limited, and the efficiency is also limited, both at a local and at a global level;
- a very small and negative correlation among the degrees of the nodes has been found (assortativity coefficient), i.e. nodes with high degrees tend to connect with nodes with low degrees;
- the modularity of the network does not show any particular grouping and is, however, very low;

[11] Reference is to the research questions listed in the introduction to Chapter 5

- the network shows a scale-free topology (power-law behaviour of the degree distributions) which is consistent with the one generally ascribed to many artificial and natural complex networks.

The technological system formed by the websites belonging to Elban tourism operators has been examined. Its topological characteristics were found to be very similar to those of the ‘real’ network (RQ6).

A number of numerical simulations were performed and the process of information diffusion on the network has been examined. The main outcome was the great improvement in speed and efficiency of the process when the network was optimised by increasing its local efficiency. Changes in the abilities of the single network actors in retransmitting the information received had, conversely, a lower impact on the course of the diffusion process. A further simulation showed that the robustness of the system with respect to random removals of connections is reasonably high (RQ3 and RQ4).

Finally, the dynamic evolution of the destination network was simulated by reconstructing the state of the system at a past period. The results show a limitation effect on the high degree portion of the degree distribution curve which shows up as a cut-off in the distribution (RQ2 and RQ5).

7 Discussion and interpretation

The results obtained in the analysis of the Elban networks described in Chapter 6 are discussed and interpreted here. This provides an answer to the initial research questions (see Chapters 1 and 5).

It is worth recalling here, before starting the discussion, that the main objective of this work is to study a tourism destination network from a structural point of view. Therefore, all the connections examined are considered un-weighted, i.e. a link between two actors does not possess any specific value or characteristic, it merely exists. Moreover, although careful attention was given in collecting the maximum amount of data, the resultant network (and the results obtained) may still suffer from 'missing data' problems (section 5.4.2). However, even considering these two limitations, the results obtained, as the discussion in this chapter will show, are relevant and valid.

The layout of this chapter follows the presentation of the results reported in Chapter 6. First of all (section 7.1) the discussion analyses the static structural characteristics of the destination network. The most important issue concerns the assessment of the degree of collaboration and cooperation among the different stakeholders. A quantitative measurement for this feature is naturally derived from the metrics used for the network analysis. The qualitative data gathered during the interviews are used to complete and substantiate the interpretation given. The virtual network (websites of the destination stakeholders) is considered next (section 7.2). The topological similarity of this network with the 'real' one is examined and the implications of this result for the analysis of a socio-economic system are discussed. The virtual network analysis also allows us to draw conclusions on the usage of information and communication technologies by the Elban tourism operators.

Section 7.3 explores the outcomes of the dynamic analysis. Robustness of the network and the information diffusion process are reviewed with the assessment of the effects of a topology optimisation based on network efficiency. Next, the reconstruction of the historical evolution of the destination, based on the qualitative investigation performed, is related to simulated network configurations at different times (section 7.4). This allows us to draw a possible network dynamic growth model which is based on the literature on the subject (see for example Boccaletti et al., 2006; Caldarelli, 2007) with the modifications needed to be adapted to the present situation.

Finally, section 7.5 examines the main implications of the results presented here for the management and governance of the destination and for the organisations involved.

The chapter closes by summarising the results obtained in this work and by relating them to the specific research questions.

7.1 The Elba tourism destination network: a static characterisation

The arguments developed in the first chapters (2, 3, and 4) have set the general background for the work which has led to this thesis. The object of study is the Elba Island tourism destination. As argued, and following a recent and important strand of research (Baggio, 2006a, 2008; Farrell & Twining-Ward, 2004; Faulkner & Russell, 1997; Russell, 2006; Scott et al., 2008b), a tourism destination is a complex adaptive system, and one of the possible representations of these systems is a network. With this approach, a model of the network formed by the destination stakeholders and their linkages was built (Chapter 6) and analysed.

The static structure exhibited by this model (see section 5.5) can be classified as a scale-free network. It is a system in which a few actors possess the majority of the connections. More than 40% of the total links are due to the most connected 20 nodes. These are the most important travel agencies and associations or consortia present on the island from a network perspective. They are the main organisations responsible for the connectivity of the network which, otherwise, would be completely disjointed.

This topology is quite common in many artificial and natural networks and has been observed also in most of the social networks studied in the literature (see for example Boccaletti et al., 2006; Caldarelli, 2007). On the other hand, it is today well known that many economic systems, and especially financial systems, exhibit complex fractal properties (Bouchaud, 2001; Gopikrishnan et al., 1999; Mandelbrot, 1963; Mantegna, 1999; Mantegna & Stanley, 2000). Fractal and self-similarity characteristics are typical symptoms of complex non-linear behaviours and are statistically represented by a power-law relationship (Komulainen, 2004). Zipf's law and Pareto's law, for example, are well known principles, both exhibiting a power-law behaviour and are commonly considered to be tell-tale signs of complex systems. As would have been reasonably expected, the Elba network shows a clear power-law behaviour in the distribution of its links. This is a further indication of the complex nature of this socio-economic system.

Despite this agreement with the general topology of many complex networks, the data presented in the previous chapter gives a picture that may look contradictory. While some values (diameter, average distance or power-law scaling in the degree distribution) are consistent with those generally found in social network analyses, other values exhibit significant differences.

The general connectivity (density of links) is quite low, and a very high proportion of disconnected nodes exist. On the other hand the network has a good degree of small-worldness, and the distribution of the distances among the nodes (in the largest connected component) further proves this compactness. The tourism stakeholders appear divided into two categories. A good proportion (37%) mainly 'live by themselves', disconnected (at least apparently) from the others, and pursuing their objectives individually and almost independently from the rest of the community. The others (63%) seem to appreciate the advantages of collaborating with diverse operators.

These first conclusions can be studied in more depth by analysing the clustering characteristics of the network. Two parameters will guide us in this analysis: the clustering and the assortativity coefficients. The clustering coefficient provides a measure of the presence of a substructure in a network by emphasising local variations of the link density. The phenomenon is close to the one modelled by Watts and Strogatz (1998) in their analysis of small-world networks. These are highly clustered networks where any two neighbours of a given node have a probability of being themselves connected, which is much larger than in a graph where the links are completely randomly distributed.

Small-world networks have been found to be the structure of a wide range of social and economic systems such as scientific collaborations, corporate alliances, friendships, Hollywood actors, Internet web pages, Broadway musicians and production teams in business firms (Boccaletti et al., 2006; Caldarelli, 2007; Dorogovtsev & Mendes, 2003; Pastor-Satorras & Vespignani, 2004).

Sociology, and social network analysis, has acknowledged this phenomenon for a long time, and concepts such as 'cliques' have been employed several times to describe such social systems (Scott, 2000; Wasserman & Faust, 1994). Sets of actors in a network who are connected to one another by strong relations forming densely interconnected groupings are deemed to be representative of some kinds of social structures (Burt, 1980; Fredericks & Durland, 2005; Freeman, 2004).

7.1.1 Assessing collaboration and cooperation characteristics

In a socio-economic system such as a tourism destination, the concept of clustering translates into one of collaboration or cooperation among the network actors. High clustering coefficients indicate the presence of local denser groups of linkages implying some form of stronger (with respect to the overall average) relationships among the elements involved (see for example Boguna et al., 2004; Caldarelli, 2007) .

The issues of collaboration and cooperation among economic entities (companies, firms, associations etc.) have been extensively investigated (see the discussion on industrial clusters in Chapter 4). There is a general consensus among researchers and practitioners on the benefits of such cooperative arrangements. Even at an individual level, the ability to establish and develop successful relationships with other companies is thought to be an important competitive advantage. Managing a good set of relationships is a critical factor for the achievement of economic development objectives and, as Ritter et al. write (2004: 181):

> The challenge for managers is to develop a networking ability that enables them to connect their resources to those of other actors. This development is hindered by the lack of understanding of the construct, but this does not mean that networking cannot be developed or is unimportant. We see the major challenge in cross-relational task development and in organizational development towards an open, networked firm.

Social structures have a significant impact on considering economic outcomes such as the productivity of the single or the group, the generation and diffusion of innovative products or of effective governance practices, the exploitation of resources, or the spread of implicit or explicit knowledge (Dyer & Singh, 1998; Granovetter, 2005).

The same phenomenon has been shown to have an even wider importance. It has been found, in fact, that in other complex systems such as multicellular organisms or insect colonies, efficient information sharing can result in sensible advantages both in reducing individual efforts and in improving group organisation, even for cases in which the information acquisition has a cost (Lachmann et al., 2000).

These advantages are particularly important in socio-economic aggregations such as industrial districts and clusters (see Chapter 4).

As discussed previously (section 4.2), these issues have acquired an even higher importance in the tourism field. More than other economic sectors, tourism involves the development of various forms of collaboration and partnerships (Bramwell & Lane, 2000). These interorganisational networks, which can also be seen as composite groups of independent suppliers who link together to deliver the final product or service, are the essence of a tourism destination. Cooperative and competitive linkages in a destination are shaped by both their internal capabilities and by the effects of the external environment (Tremblay, 1998, 1999).

However, the lack of cooperation among the actors of these destination networks is well known for the fragmented nature of the tourism industry and the diversity of activities and organisations involved. This fragmented nature is considered a main reason for the need for cooperation. Many authors (Huxham, 2003; Jamal & Getz, 1995; Roberts & Simpson, 1999; Timothy, 1998) claim that a tourism system can have a balanced and planned evolution only through a process of shared information and decision making with all the stakeholders involved. This much sought-after capacity to work together is considered to be a crucial element for the success of a destination (Bramwell & Lane, 2000).

It is clear, then, that the capacity to assess collaborative or cooperative groupings in the most rigorous and objective way is important for all those interested in the development of a tourism destination. So far, these characteristics have been assessed by using qualitative methods such as surveys or focus groups (see for example Jamal et al., 2002; Ladkin &

Bertramini, 2002). The limitations in reliability of these approaches are well known (Bernard et al., 1984; Killworth & Bernard, 1976).

When representing a tourism destination as a complex network, the clustering coefficient is a natural candidate to provide a quantitative measurement of the existence and the extent of tightly connected neighbourhoods. This measurement, obviously, can hardly deliver definitive conclusions as it considers only the topological structure of the destination network. The coupling of this metric with more traditional qualitative assessments, however, can provide a powerful combination able to gauge this phenomenon and to complement (and validate) other possible investigations (see Chapter 2).

One more quantity measurable for a network can complete this picture: the assortativity coefficient. It measures the extent to which a correlation exists between the degree of a node and those of its neighbours. In assortative networks, well-connected elements tend to be linked to each other. Most social networks exhibit this characteristic. Someone who is particularly sociable and friendly is very likely to know also the other most sociable people. If a company has an operational or economic connection with some other organisations, it is probable that these latter will have common connections among themselves (Chapters 2 and 3). These relationships look to be a distinctive trait of a social system, as McPherson et al. write (2001: 415):

> Similarity breeds connection. This principle - the homophily principle - structures network ties of every type, including marriage, friendship, work, advice, support, information transfer, exchange, comembership, and other types of relationship. The result is that people’s personal networks are homogeneous with regard to many sociodemographic, behavioral, and intrapersonal characteristics.

The results presented in Chapter 6 give a very low value for the clustering coefficient of the Elba network and a low and negative value for the assortativity coefficient. The conclusion is, therefore, that our destination shows a substantial lack of cooperation or collaboration among the stakeholders.

This inference is in agreement with a series of considerations that emerged during the interviews conducted at the destination and with the results of previous studies (Pechlaner et al., 2003; Tallinucci & Testa, 2006). The origins of Elban tourism activities go back to the early 1960s. They arose as a complementary source of revenue for many families still, at that time, employed in the mining industry (see also Chapter 4). The growth of tourist numbers was almost spontaneous up to the 1980s, due mainly to the attractiveness of Elba's natural resources. As Tallinucci and Testa say (2006: 33): "it was not the island to choose tourism, it was tourism to choose the island".

A diffuse geographic distribution of small family-run businesses, with a strong independence and a high rate of loyal guests producing good word-of-mouth promotion, are the main original characters of the tourism industry at Elba, and they have survived up to the present time. These conditions have supported the spontaneous, unregulated development of the destination, in which possible incentives to aggregation have traditionally been contrasted by the individualism of the single operators. It was only after crisis periods, such as those that occurred in the mid 1980s and in the early 1990s, that coordination and planning activities acquired some momentum.

The low value for local efficiency is a direct consequence of today's sparse structure of the network. It is, in fact, closely related to the clustering coefficient (Latora & Marchiori, 2001). The low efficiency translates directly into a low capability of the network actors to transmit information and knowledge (see Chapter 2). This fact threatens to worsen the capability of the destination stakeholders to be able to react, as a system, to the effects of external events. In essence, a low efficiency of the system is detrimental for its capacity to achieve good results. With this assumption it might be possible to partially explain the decreasing number of overnight stays in recent years (see Figure 6.16).

The substantial agreement of the interpretation given to the clustering and assortativity coefficients with the general picture of the relationships among the Elban tourism operators supports the use of these coefficients as quantitative assessments of the extent to which the tourism organisations work together collaborating or cooperating, i.e. forming cohesive communities inside the destination. The clustering coefficient can be thought of as a *static* measurement, while the assortativity coefficient can be interpreted as expressing the *tendency* to form such communities (see also Baggio, 2007).

7.1.2 Modularity features

A further way to assess the clustering characteristics of the Elban network is to study its modularity characteristics. The homophily principle quoted above, interpreted in an interorganisational network as similarity between companies with respect to ‘routines and competencies’ (Staber, 2001: 546) can be verified by looking for some structural characteristic of the network.

Many complex systems have shown the presence of subgroups, in which the density of linkages varies. Multiple clusters of actors can occur as cliques where intracluster relations are denser than intercluster relations. This modularity characteristic, as discussed in Chapter 3, has no common definition, but, generally, it is assessed with some numerical algorithm designed to maximise the modularity of the network when adopting different partitionings of the actors.

In the case of Elba, first of all, the study examined if ‘natural’ grouping of the tourism operators exhibits significant modularity characteristics and, in particular, if they tend to group together based on the similarity of their businesses (hotels, intermediaries, service companies etc.) or on geographical proximity (the eight municipalities in which Elba is administratively divided).

The results obtained show a very low level of modularity. Elban operators look, even from this point of view, quite disconnected. When looking at the modularity of the ‘type of business’ division, a negative value for the modularity index was found (-0.255). A negative value tells us that operators of the same typology tend to have more links connecting them with a different group than among themselves. It is possible to see in this lack of will to connect similar entities a clear sign of a strong ‘competitive’ attitude.

This apparent lack of collaboration among operators belonging to the same type has proved to be detrimental when thinking about the capacity of innovation which might help them face the challenges of the contemporary highly competitive and globalised market. It has been shown, in fact, that a collaborative approach and intense exchanges, even in seemingly competitive organisations such as the group of Sydney hotels described by

Ingram and Roberts (2000), may allow a valuable amalgamation of best practices, with the result of improving the performance and profitability of the whole group and its members.

It is interesting to note, in the results of the analysis (Table 6.2), that the highest modularity value is obtained with the usage of a 'generic' numeric algorithm, the one proposed by Clauset et al. (2004). The topology of the network shows limited modularisation (i.e. the values of the modularity indices are quite low). This community structure, in the common understanding of the phenomenon, can be considered better than the others which can be found based on different criteria (Arenas et al., 2004; Danon et al., 2005).

As can be seen in Table 6.2 (Chapter 6), the number of clusters found is different from the number found when considering the other types of grouping used (mainly the other positive one, the geographic division). Moreover, the clusters are composed of members belonging to different groups (as for geography and typology). The system, in other words, exhibits self-organisation properties which lead to the formation, to some extent, of some agglomeration of ties and produces a number of informal communities and an informal community structure.

It can be concluded that the information contained in the geographical or business typology data does not represent fully the communality characteristics. The network topology shows how these data provide non-optimal modularity solutions when geographic or typology groupings are used. This evidence has been also found in other social networks (Minerba et al., 2007).

From a destination management viewpoint, this result is important. It can provide indications on how to optimise some performance, for example, optimal communication pathways or even productivity in collaborations, overcoming rigid traditional subdivisions (administrative, for example). It can provide a more practical tool to go along with the ideas and practices of an adaptive approach to the management of a tourism destination which has been advocated by some scholars as discussed in Chapter 4.

A word of caution is necessary, mainly when considering extending the considerations made on network clustering and modularity to other cases. It has been shown, for example, that the value of the clustering coefficient in a network can be also accounted for by a simple random graph model in which edges are placed at random, under the constraint of a fixed degree distribution *P(k)*. In other words, the emergence of this effect can also be a kind of 'statistical fluctuation', a mere topological property due to the form of the degree distribution that disappears in the thermodynamic limit (number of vertices $N \rightarrow \infty$), but able to produce noticeable values in finite networks (Newman, 2003a; Newman et al., 2001). A correct interpretation of the result, therefore, can only be achieved by complementing the quantitative assessment with a deep knowledge of the social system under study, which typically comes from a tradition of qualitative investigations.

7.2 The technological network

The analysis of the website network belonging to the Elban tourism operators has shown results which are quite similar, from a qualitative point of view, to those of the 'real' network. Summarising, it is possible to say that:

- the results show a general agreement with similar results obtained by studying the Web and website configurations, with the exception of the bow-tie model (section 6.3);
- the network's most striking characteristic is its low connectivity and high sparseness (section 6.3);
- moreover, a very small and negative correlation among the degrees of the nodes has been found (assortativity coefficient, see section 6.3).

As happens in many tourism destinations in the world, Elba's tourist enterprises consider the Internet as an open window on the world, able to improve their communications with actual or potential guests and to favour their promotional activities.

A 2002 survey of Elban operators (Pechlaner et al., 2003) revealed a diffuse usage of Internet technologies with an extremely high rate of self-made websites (31.9%) and a sound preference for direct contact with the customers (80%) in order to increase loyalty and create individual brands to link with the destination image of Elba. According to a recent survey by the statistics bureau of the Toscana region (SSSR, 2006), almost 80% of the hospitality structures make extensive use of ICTs, 90% have some kind of visibility on the WWW either because they own a website (62%) or because they have some pages contained in portals belonging to institutions or other private companies.

When considering the data presented about the technological network, it is possible to interpret the low connectivity and modularity (i.e. low and sparse number of connected communities or clusters) and the low efficiency as a waste of resources both from a technical and an organisational point of view. The Elban tourism operators appear to be missing out on the advantages of collaboration and cooperation mechanisms on the Internet that could greatly improve the organisation and management of the destination and its capacity to face a highly competitive globalised market. The benefits of sharing technological resources and functionalities should be well known and have been emphasised several times (Barua et al., 2000; Hackathorn, 2003; McLaren et al., 2002; Walker, 2002). The operators of Elba, on the contrary, look to have neglected these benefits. This inadequate usage of advance communication technologies may explain the relatively poor performance recorded. The Toscana region survey (SSSR, 2006), for example, measures a scant 37% of hotel electronic bookings for the 2006 Easter holidays. As a comparison, the estimate for the whole European industry in 2006 is about 50% (according to JDP, 2006).

The Web visibility of the tourism operators is also at risk. New generation search engines are already available online. They strive to supply more relevant responses to users' queries by analysing the modular structure of the Internet and trying to locate websites on the basis of their belonging to distinct thematic communities (Lehmann & Hansen, 2007; Liu et al., 2006). IGroup (igroup.msra.cn), a Web image search engine or Clusty (clusty.com), a textual metasearch engine, for example, work along these lines and organise the search results into semantic clusters (Jing et al., 2006; Moghaddam, 2007). From a strategic development perspective, the situation is even more problematic. Future search engines and recommendation systems, in fact, will be based on dynamic agents whose main task is the dynamic identification of connected communities on the Web (Adamic & Adar, 2003; Adomavicius & Tuzhilin, 2005; Lawrence, 2000; Skopal et al., 2003).

Those websites of tourism operators in destinations not forming well recognisable network ‘communities’ through dense interconnections will be difficult to reach by a casual user. This will have, obviously, unfavourable consequences for the effectiveness of the marketing and communication activities and their economic outcomes.

7.2.1 The virtual and the real network, a comparison

In a paper published a few years ago, Wellman (2001: 2031) suggests a close relationship between a computer network and the social network of the people using it. In his words: “computer networks are inherently social networks, linking people, organizations, and knowledge”. Quite a number of other studies have, in one way or another, maintained a similar interpretation (Haythornthwaite & Hagar, 2005; Matei & Ball-Rokeach, 2001; Park, 2003).

In many other fields, the Internet and the WWW network have provided the basic materials with which network scientists have significantly improved our comprehension of how systems of any kind, not only computerised, are structured or behave (Barabási & Albert, 1999; Faloutsos et al., 1999; Watts & Strogatz, 1998). Examples are the analyses of social networks deduced by the exchange of e-mail messages (Tyler et al., 2005), by hyperlinked individual webpages (Adamic & Adar, 2003; Park, 2003) or by the websites of similar organisations (Pennock et al., 2002).

Social network analysis (SNA) has for many years collected results on the formation and the evolution of human, social and economic relationships, on the importance of some positions in the web of connections we have, and on how to use these outcomes to steer and to encourage the development of a community, a company or a society (Wasserman & Faust, 1994).

As discussed in section 5.4, one major problem faced by social network analysis has always been the collection of the data needed for the analyses. Many methods have been devised and many techniques have been proposed to allow the extraction of meaningful insights from the sometimes scarce records a researcher is able to collect on the elements and linkages of a social network (Marsden, 1990). On the other hand, the outcomes obtained by the contemporary network scientists could be used to improve this work. But these are typically coming from huge quantities of data and at least an adequate and reliable amount of it is needed to highlight structures, differences, patterns and to trace evolutionary processes and developments.

The Web has been seen as a good candidate to provide this amount of data and several studies have followed this direction creating a discipline called *hyperlink analysis* (Park, 2003). It is reasonable to assume, also for the importance given to the practice of hyperlinking (Walker, 2002), that the layout of the network can be a reflection of the structural characteristics of the social network from which it originates. This relationship between cyberspace and the physical world is two-way: on one side, the online linkages represent and complement social relations in the offline world; on the other side, offline interactions can influence the way in which online relationships are established and developed (Birnie & Horvath, 2002; Wellman, 2001).

Some have argued and warned of the risks and dangers of this approach. They claim that the links are built in a rather unpredictable way, and it is not possible to find unambiguous meanings (Thelwall, 2006). While this can be true when thinking of webpages built by individuals, the situation looks different for private or public organisations. We see more and more cases in which the practice of hyperlinking is regulated. In these, the presence of a link reflects a specific choice made by the website owner. A link is considered to be a strategic resource, and the possible variations in the structure of the 'corporate' WWW are shaped by specific communicative aims, rather than by random technological processes (Birnie & Horvath, 2002; Vaughan et al., 2006).

A scrutiny of corporate websites belonging to diverse industries (140 of them have been inspected) shows that a vast majority (87%) publishes some form of 'instructions' and legal disclaimers about the website and its contents. All of them specifically address the issue of linking. They state how and when permissions must be sought before a connection can be made to the organisation's website. Moreover, the analysis of a number of internal documents shows that carefully defined policies regulate the publication process for all the contents of a website, including the hyperlinks to other organisations. In a series of enquiries conducted elsewhere in Italy, more than 400 tourism operators, grouped in a number of associations or consortia, were informally asked to describe the process in place to update their websites. All of them, even those relying on some form of technical outsourcing, stated that the insertion of a link towards a different 'entity' is decided largely on a *business* basis. More specifically, links to companies belonging to the same sector of activity are only present when some kind of collaboration agreement exists (Antonioli Corigliano et al., 2007; Baggio, 2006b, 2006c).

A more rigorous study on the web networks of two tourism destinations, the island of Elba (Italy) and the Fiji islands (Baggio et al., 2007b), shows that 68% of their links connect other tourism companies in the same area, 3% link non-tourism companies in the same area, 22% link tourism organisations outside the area, and the remaining 7% are generic broad interest links.

These outcomes, however, could be classified as *presumptive evidence*. A better way to specify the relationship between the virtual and the real network of a socio-economic system (a TD in the present instance) is to look for similarities and differences in their topological structures. The comparison made on the characteristics of the real Elba network and the virtual one (section 6.3.1) reveals a very good topological similarity of the two networks.

Clearly, full equality cannot be claimed, but the results legitimate the web as a significant sample of the underlying socio-economic network. The obvious limitation is that the comparison holds when considering 'institutional' websites, belonging to companies, associations or other institutions. Additionally, the area taken into account must show a quite high usage of the Internet and the Web. Yet nowadays, for a large part of the world, this is not a severe limitation.

By carefully applying the considerations made in the literature (see for example Kossinets, 2006; Lee et al., 2006) on the handling of samples of networks, the WWW can provide an effective environment to study the characteristics and behaviour of social and economic systems formed by companies, corporations, associations and other such entities. Tools and methods of the science of networks can thus be extended to these important components of today's world.

This result is even more important for the field of tourism. Quantitative network analysis methods can provide more possibilities to improve and complement our qualitative knowledge of these structures. Moreover, the network approach can be extended to implement simulation models with which different scenarios can be obtained in order to explore the possible effects of different managerial activities. Finally, regular scanning of a tourism destination webspace can provide evidence of dynamic transformations of the network, thus easing the analysis which, otherwise, would be quite difficult given the complexity of performing a traditional investigation. This can give all people interested in the life of a tourism destination powerful tools to inform their policy or management actions.

7.3 Dynamic processes

Network dynamic processes are strongly influenced by the underlying topology (see Chapter 3). The small-world characteristics have been shown to enhance the performance of many processes as compared to regular or random lattices. This is a direct consequence of the presence of shortcuts that accelerate the communication between otherwise distant nodes and of the relatively short distance between any two vertices (Latora & Marchiori, 2001; Zanette, 2002). On the other hand, the heterogeneity of the link distribution in scale-free networks strongly affects the development of processes that they support. For example (probably the most remarkable one), it is known that an epidemic disease will spread completely in a scale-free network (at the thermodynamic limit, i.e. for an infinite number of nodes) regardless of its infection rate, while in a random network, a critical threshold density of connections must be achieved before being sure that the diffusion process can reach all the nodes of the network (Barthélemy et al., 2005; Pastor-Satorras & Vespignani, 2003).

Once the basic topology of the network is established, the models discussed in the literature can provide a guideline for understanding the behaviour of the system when some important dynamic processes take place. In particular, for a socio-economic system such as a tourism destination, two phenomena are of particular interest. The first is the ability to maintain its fundamental properties when it is subject to some disrupting event, and the second involves the capacity of the system to act as a substrate for the diffusion of information and knowledge.

One of the advantages of the network representation, as discussed in the previous chapters (see Chapters 2 and 5), is the opportunity to investigate these phenomena by using computer simulations. More than giving accurate descriptions and explanations of phenomena, numerical simulations allow experiments to be performed in fields where these would be not feasible for both theoretical or practical reasons. The power of a model lies in its capacity to allow 'playing' with it, exploring different possibilities, different configurations, and to derive useful predictions on some characteristic of the system.

Numerical simulations have been used to analyse the two phenomena quoted above. The soundness of this approach comes from a methodological point of view. It is known from the literature, in fact, that the most important determinant for the validation of a model lies in the design of the conceptual model (Adrion et al., 1982; Balci, 2003). The models used in our case are well known and have been described and used several times to assess similar situations (see Chapters 4 and 5).

The robustness of a complex system and the preservation of its fundamental properties even in presence of external (or internal) disastrous events is, as seen in Chapter 2, one of the features of such (complex) systems. In the case of Elba, a certain degree of resilience (the term usually employed to define this property) has been estimated already by using different methods (Baggio, 2008). Here a 'network' approach was used. The progressive deletion of linkages among the destination actors confirmed this result. It also confirmed the common understanding of this phenomenon in complex scale-free networks.

In the results obtained, however, it must be noticed that the difference in the behaviours of the networks considered (Elba and a random network) is not as large as would have been expected from the literature (Albert et al., 2000; Crucitti et al., 2003, 2004b). This has been ascribed to the characteristics of the Elba network and mainly to the flattish degree distribution which makes it closer to a random network than a 'typical' scale-free graph.

Elba's network is more robust than a random network, but its robustness is quite fragile. In this case, both theory and models predict that the only possibility for achieving a better capability of the Elba network to resist random shocks which might reduce its information transfer efficiency is a thorough reconsideration of its connectivity characteristics. In other words, a higher resilience of the system can be obtained by causing an expansion in the number and type of linkages among the tourism operators so as to increase the *scale-freeness* of the network and its clustering properties. This can be directly translated into a clear direction for the design and implementation of policies for the destination.

The same indication, reinforcing this conclusion, comes from the second set of simulations performed on the diffusion mechanisms. In this case, two conceptually different situations were simulated. The first one considered the stakeholders of the destination as elements with different capabilities to acquire and consequently retransmit information or knowledge. The second one assessed the effects of a change in the topology of the network obtained by optimising it with respect to its efficiency.

The basis for the first simulation is the idea that particular actors can have difficulties in acquiring and retaining all the knowledge available to them due to their internal functioning or because of the associated costs. This feature is usually called absorptive capacity, an important construct introduced to organisational studies by Cohen and Levinthal (1990). It refers to the capability of a firm to recognise, absorb and exploit knowledge from the environment. Building up and sustaining absorptive capacity is deemed to be a critical factor for the development and success of an organisation (Tsai, 2001). A rich literature has studied the effects and motivations for the establishment of a good level of absorptive capacity. At a basic level, this is, without doubt, connected with the organisational capacities and with the extent to which a firm is able to exploit all the available resources. It can also be connected with the firm's size (see the review in Lane et al., 2006). In a very concise way, "firms with higher absorptive capacity are more likely to establish knowledge linkages with other local firms" (Giuliani & Bell, 2004: 4).

The measurement of the speed of diffusion, and the form of the diffusion curve obtained (Chapter 6, Figure 6.12 and Figure 6.13) confirm, first of all, the improvement in the efficiency of the whole process due to the existence of a structured network in place of a randomly linked system. The results are also a confirmation of the importance of the capacities of the single operators. There is a clear improvement in diffusion speed when all the actors are considered to have the same capacity to transfer information or knowledge. This is an important indication for a destination manager; putting in place measures and actions aimed at reducing the differences in the absorptive capacities of the destination stakeholders can have a highly beneficial impact on the overall system.

However, the second simulation results indicate that a similar effect, but with an even higher magnitude, can be obtained by optimising the network efficiency. The exchange of information among the nodes, that is the 'physical' meaning of efficiency (see Chapter 2), is much improved if the connectivity of the network is modified so as to increase the local efficiency, and consequently the clustering coefficient.

This shows that a very important determinant for the spread of knowledge in a socio-economic system such as a tourism destination is the presence of a structured topology in the network of relations that connect the different stakeholders, and more than that, the existence of a well-identified degree of local cohesion. This supports the notion that destination stakeholders should be encouraged to form clusters and to both compete and cooperate in order to exchange knowledge and hence to raise the overall competitiveness of the destination. Often the public sector intervenes to initiate such cooperative processes, given the competitive nature of SMEs. Public sector support can facilitate a network and provide ongoing support, but it is the destination stakeholders who must operate the network.

The network simulation has thus reconfirmed a known principle in industrial economics and destination management studies. As discussed in Chapter 4 (see also Becattini, 1990; Bramwell & Lane, 2000; Porter, 1998), a large number of researchers maintain that the most significant determinants for the competitive advantage of a district (or an industrial cluster) are in the speed and the efficiency of the circulation of information and knowledge. The Marshallian *industrial atmosphere* (Marshall, 1920) has an essential role in the creation of the entrepreneurial culture within cohesive groups of firms, with the final effect of expanding and strengthening the system and of generating positive externalities.

Quantitative network methods, such as those used here, can, therefore, not only assess this effect, but, more importantly, give practical indications on how to improve the process. By performing different simulations with different sets of initial parameters (distribution of absorptive capacities or different levels of clustering), it is possible to obtain different settings and evaluate the effects of the choice of parameters on the final result.

From the *network science* point of view, these results are not completely new. The effect was identified by Granovetter (1973, 1983) as the *strength of weak ties* and reconfirmed by the more recent works on *small-world networks* (Latora & Marchiori, 2001; Uzzi & Spiro, 2005; Watts & Strogatz, 1998), and by several other authors empirically (Reagans & McEvily, 2003; Sorenson et al., 2005). Here, for the first time, a tourism destination is used as test case.

By coupling the theoretical and analytical approach with a detailed understanding of the destination and its stakeholders, it has been possible to diagnose the efficiency of the destination's network structure and its implications for competitiveness. It is also possible to utilise policy instruments to intervene and to make the network more efficient. In other words, in our case, the simulations can be used to create development scenarios in which the efforts to move towards strong forms of collaboration are increased, even if at a very 'local' level. This can be highly beneficial not only for the stakeholders involved, but for the whole destination.

Moreover, the same simulations can be employed in the understanding of how deep, fast and efficient the spread of a message through the system can be. This can be of further help for a destination management organisation wishing to communicate effectively with the stakeholders.

7.4 Simulated evolution of the Elba destination

A simulation approach was used to trace the evolutionary history of the destination because of the difficulty in collecting precise data on the modifications of the network over time. The method used was a semi-qualitative one. As discussed in the previous chapters, the probable set of connections and actors was deduced from a number of sources, the first being the series of interviews conducted at the destination (see Chapter 6). The others were the checks made on the evolution of the Elban tourism infrastructures made on the data provided by the local tourism board.

The framework for the interpretation of the destination's evolution follows (see Chapter 6) the generally accepted model proposed by Butler (1980). The analysis performed on the reconstructed network of Elba tourism stakeholders can enter this framework to explain at least some of the mechanisms which are at the basis of the evolution of such a network.

First of all it must be noted that the vast majority of the network growth models reported in the literature, whose outcomes are scale-free networks such as the Elban one, utilise the basic 'preferential attachment' mechanism proposed by Barabási and Albert (1999). In this model the network grows with the creation of new nodes and the probability for a new node to create a link to an existing one depends on the degree of the attached node, i.e. the higher the number of links a node has, the higher the probability to form new connections. This basic mechanism can be modified by introducing a number of possible variations. A node can have a high 'fitness', which increases the probability of receiving links even if it is a newcomer. Factors such as cost limitations in forming connections, aging of nodes that stop creating links after a certain period of time, spatial confinements of the network or finite lifetimes can be introduced. The result is, in all these cases, a degree distribution which scales as a power law, while some of the constraints can generate slight modifications such as changes in the slope of the curve in some region (Boccaletti et al., 2006; da Fontoura Costa et al., 2007).

By examining the results of the simulated reconstruction, the most striking result is the difference in the shape of the degree distribution. The 'old' network, although exhibiting a general power-law behaviour, shows a clear cut-off at high nodal degrees. This cut-off disappears in the degree distribution of the contemporary network.

From a topological perspective this means that at a certain point in time, high degree nodes have connected with only a fraction of the nodes they would have connected with in a pure preferential attachment condition. In the subsequent period of time, the network grows with the addition of nodes and links, and the high degree nodes complete their 'links budget'.

As already stated, to be meaningful, this quantitative idea must be coupled with the knowledge of the system under study. In our case, among the different network growth mechanisms reported in the literature, the constraint mechanisms which look more suitable are those related to the level of information network actors have about the rest of the system. A number of studies (Mossa et al., 2002; Štefančić & Zlatić, 2005) have highlighted that information filterings are responsible for truncations in the power-law behaviours of the degree distributions.

This phenomenon is known in the economic literature as *bounded rationality*. Introduced by Simon (1955a; 1997), it contrasts the game-theoretical idea of an economic agent able to make decisions based on rational examination of the possibilities it faces. The concept serves to overcome some of the limitations of the (over)simplifications of this approach which disregards the "psychological limits of the organism (particularly with respect to computational and predictive ability)" (Simon, 1955a: 101). On the contrary, Williamson (1981: 553) states: "boundedly rational agents experience limits in formulating and solving complex problems and in processing (receiving, storing, retrieving, transmitting) information".

In our case, we can use this interpretation by stating that, during an early stage of the destination's evolution, high fitness actors exist, but they have not yet connected to others. This happens because they probably do not feel such a necessity or because they have not yet recognised the existence of other stakeholders or because the social rules for interaction have changed over time. Larger organisations or associations, generally responsible for the higher degrees in the network, still have to establish a link with the newer nodes (following Mossa et al., 2002). Or, somehow symmetrically, the new nodes create links by using a preferential attachment method, but exert this 'option' only on a fraction of all the existing nodes because of a limited knowledge of the magnitude of the whole network due to their 'young age' (as in the proposal by Štefančić & Zlatić, 2005). Whatever it is, the limitation in (some of) the nodes' ability to process information about all the other nodes of the network is able to generate the exponential truncation found in the degree distribution of the network.

It is interesting to note, as does Mossa et al. (2002: 4), that:

> In the context of network growth, the impossibility of knowing the degrees of all the nodes comprising the network due to the filtering process – and, hence, the inability to make the optimal, rational, choice – is not altogether unlike the 'bounded rationality' concept of Simon (1997). Remarkably, it appears that, for the description of WWW growth, the preferential attachment mechanism, originally proposed by Simon (1955b), must be modified along the lines of another concept also introduced by him – bounded rationality (1997).

This interpretation has been used to explain the difference (similar to the one found for the Elba network) between the networks of two different tourism destinations (today's Elba island and Fiji) which are reputed to be at different evolutionary stages according to Butler's model (Baggio & Antonioli Corigliano, 2007; Baggio et al., 2007b).

With this theoretical apparatus, combined with the one coming from the study of tourism destinations, the history of the Elban network can be described in the following way.

The destination starts its life cycle when it is *discovered* by the visitors and some reception facilities are implemented: the *exploration* and *involvement* phases of Butler's model. The destination enters then a *development* stage in which the supply infrastructures spread over the territory. At this point it is reasonable to suppose that the network is a rather sparse and disconnected system in which a few links exist, and they are mainly established in a random way. This is in agreement with the initial stages of most network growth models.

The *consolidation* and possible *stagnation* phases see a relatively slow increase of the network's linkages. In this situation, the network starts assuming its scale-free characteristics. The preferential attachment mechanism drives the evolution. Destination stakeholders (mainly the small ones) establish connections with some of the other fitter and larger actors, typically intermediaries, and self-organise forming (embryonic) associations. In this phase the cohesiveness of the network is still very low, showing very limited degrees of collaboration and cooperation among the tourism operators. New elements of the network still arise but they or the most connected nodes have only incomplete information about the overall network size and, therefore, do not still create all the possible connections. The network shows a clear tendency towards a scale-free topology, but the information limitations produce a bending in the degree distribution at high degree values.

The number of nodes is now stable. The stagnation phase, in Butler's model, can give rise to a rejuvenation or to the decline of the destination. In the case of Elba, this corresponds to the 1990s break. The combination of both external (improvement in the economic conditions and the start of EU financed projects) and internal (basically the realisation of the need for new forms of cooperation) events produce, as seen above, a new growth phase.

From a network viewpoint, this coincides with a growth in the density of links, the completion of the large actors linkage potential, the straightening of the degree distribution which tends to a 'purer' power law and an increase in the clustering of the network.

The scenario depicted above justifies well the empirical results. Moreover, it allows us to establish a relation between the evolution of the destination network and the outcomes in terms of the measured tourist fluxes.

These results reconfirm those obtained by other authors who have examined the influence of the structure of a destination network on its development. For example, Pavlovich, studying the New Zealand Waitomo destination (Pavlovich, 2001, 2003a; Pavlovich & Kearins, 2004) has shown how the relationships between stakeholders can act as a self-organising mechanism for the destination, and demonstrated the role and function of the aggregated patterns of interconnectivity in understanding the formation and development of strategic capabilities that belong to the network. The structural embeddedness proves to be able to assist or inhibit organisation within networks and to favour, through the formation of heterogeneity and interconnections, the reciprocal flows of resources and information. These greatly affect the development of strategic capabilities and, consequently, the growth of the system.

Obviously, the mere topological structure evolution cannot be used as a full justification for the destination growth. More data and information on the other several possible determinants should be thoroughly investigated. This is, however, outside the objectives of this work.

7.5 Network thinking: implications for management and governance

The exploratory investigation presented in this thesis has, as already stated, a number of important implications in the field of tourism studies. They have been already expressed while discussing the outcomes of this investigation, but it may be worth (even at cost of some repetition) to summarise them. First of all it must be noted that the results discussed here have been obtained by combining a quantitative and a qualitative approach. As in many other disciplines, in fact, the results of numerical analyses, to be 'useful' and significant, need a solid layer of interpretation which can only be possible by turning to the outcomes of qualitative studies on the same subject. Conversely, qualitative investigations can greatly benefit from the possibility of an 'independent' quantitative confirmation of the results. This is the case, for example, of the proposal made to use specific network metrics for the assessment of the extent of cooperation in a destination.

The application of network analysis to the management of both single and groups of organisations is not a new idea. The framework, the methods for both data collection and analysis and the implications in organisation theory and research have been known for a long time, at least among some (although not many) scholars (Tichy et al., 1979b).

Today's strong focus on issues such as partnership, collaboration, cooperation and the benefits of the tools available for the investigation of the relationships between the elements of a socio-economic system have been discussed several times in the area of general management studies. The implications, it is argued, go well beyond the simple study of networks. These methods are recognised to have a strong potential to inform a wide number of concerns such as the use of technology, the study of epidemiological diffusion (from diseases to marketing messages), the formation of consensual opinions and the impacts of these on organisational structure and performance (Parkhe et al., 2006).

Elban network analysis considerations can help destination stakeholders in understanding who is involved in the network, what (if any) significant division into clusters exists or what new connections are worth developing. Most importantly, it can make it possible to decide to what extent actions aimed at increasing the density of relationships should be put forward and to derive the effort needed in managing and coordinating such a network. The diffusion simulations performed, along with some more knowledge of the attitudes of the single stakeholders and the experience of past activities in this field, can provide insights into the outcomes obtainable. This may allow the completion of a 'balance sheet' in order to make well informed decisions. It is known, in fact, that excessive involvement or collaboration might not always be a desirable objective (Provan & Kenis, 2007; Provan et al., 2005). At Elba, this could ease the process of forming a 'control room', which many of the interviewees have called for, with the objective of coordinating shared actions for marketing the destination and for improving the effectiveness of the customer relations management practices implemented so far (see also Tallinucci & Testa, 2006: 192). An increase in efficiency of the overall communication process in the destination is felt necessary by many Elban operators, even if not much is done in this regard. Undeniably, optimised communication channels can greatly enhance the diffusion of knowledge, of best practices, of 'good attitudes' and can greatly help in fostering a more collaborative atmosphere which is one of the premises for a higher competitiveness of the whole destination (see section 4.2).

Managers willing to encourage effective networking practices should be first aware of the structure and general dynamic behaviour of a network. The structure can be studied by using methods such as those presented in this thesis, and the patterns of change can be identified and measured as has been done by comparing the different maturity stages of Elba. With such environmental knowledge, it may be more comfortable to evaluate the stakeholder's role in the network and to assess the congruence between strategic goals and possible evolutionary scenarios. The acknowledgement of the path of the system's evolution and its implications may put a manager in a better position to take full advantage of the opportunities while avoiding the threats related to the configuration (Koka et al., 2006).

Rather obviously, a great part of the effectiveness of a networked system is due to the actions of individual actors, and single stakeholders may benefit from network involvement apart from the form of governance. However (see also the discussion in Chapter 4), collective outcomes, effectiveness and behaviours are significantly influenced by governance forms based on a systemic approach (Berry et al., 2004; Farrell & Twining-

Ward, 2004; Lazzeretti & Petrillo, 2006). And this is more valid when considering an inherently networked system such as a tourism destination (Aas et al., 2005; Agostinho & Teixeira de Castro, 2003; Bickerdyke, 1996; Capone, 2004; Carlsen, 1999; Pechlaner et al., 2002; Walker et al., 1998).

From a policy perspective, "when governments, communities, foundations, or regional industry groups think about how they can improve their economy, disaster preparedness, competitiveness, health and well-being of citizens, and so on, collaboration through an interorganizational network is an approach that is increasingly utilized" (Provan et al., 2007: 512). The governance style of a network system can have significant implications for its overall effectiveness. A valuable management requires the capability to appreciate and react to both internal and external demands, mainly when dealing with the tensions that may arise (Provan & Kenis, 2007). In this respect, the methods of network science can prove highly beneficial in deepening the knowledge of the whole system and, coupled with more traditional procedures, can provide powerful tools to enable those *adaptive management* practices considered by many the only practical way to steer the collective efforts of multiple organisations (Bankes, 2002; Farrell & Twining-Ward, 2004; Holling, 1978; Ritter & Gemünden, 2003; Ritter et al., 2004). The possibility to analyse in a quantitative way the relationships between tourism operators opens new paths for the researcher interested in the structure, the evolution, the outcomes, the effectiveness and the governance of the system. More than that, these methods and the emphasis on the whole network rather than only on the particular relations that any organisation might have, offer effective tools to practitioners, both at a single operator and whole destination level, to strengthen their capacity of intervention. Most interviewees have agreed on this perspective, recognising the need for stronger guidance for the destination which, however, should take into due account the 'autonomous' attitude of the majority of Elban tourism operators in conducting their businesses.

7.6 An answer to the research questions

The main research question posed at the beginning of this thesis was: *How do tourism destination networks evolve over time?*

This question has been articulated in a number of subquestions which have formed the main guideline for the research work presented in this thesis. The results and their discussion in Chapters 6 and 7 give an answer to these questions. This section is a summary of the answers found.

RQ1: What is the structure of the tourism destination network? and RQ6: How do the real and technological networks compare in a tourism destination?

The investigation of the island of Elba has shown the implications of assuming that a complex system can be represented by a network. The topology has been found in general agreement with the one exhibited by many other complex networks. The differences found, mainly in the density and clustering, have been related to the sociological characteristics of the destination. The metrics (clustering and assortativity coefficients) have been used as quantitative assessments of the degree of collaboration and cooperation among the stakeholders. The different modularity properties have also been discussed, showing the tendency of the system to self-organise beyond some basic 'induced' subdivision such as the geographical or administrative ones.

The analysis of the network formed by the websites of the tourism operators has given indications about the usage of technology and warned about the possible strategic directions in this important field. More importantly, it has provided evidence of the possibility of using this technological network as a significant sample of the *real* network. This result may greatly lower the burden of data collection and favour an increase in the number of network studies of tourism destinations.

RQ3: How does the network structure affect dynamic processes? and *RQ4: How can the network be optimised?*

An epidemic information diffusion mechanism has shown the different effects of the actor's capabilities and of the network topology. A general qualitative agreement with similar works reported in the literature has been found, while the peculiar characteristics of the system studied can explain the differences. The results demonstrate clearly the importance of the distribution of links in a simulated network and principally of the clustering characteristics. To this extent, the rewiring procedure with its increase in the efficiency of the network, has given indication of a possible modification (optimisation) which can greatly improve the diffusion mechanism. In the cases in which this 'mechanical procedure' can be translated into feasible modifications of the real destination network structure, this is undoubtedly an additional possibility offered to a destination management body in order to improve the flow of knowledge and information. In fact, it has been shown that the effects of network optimisations on the course of a diffusion process outperform other possible modifications such as changes in the actors' abilities to transfer information.

RQ2: How can we build a model for network evolution that reproduces the present structure of the network? and *RQ5: How does this model explain the history of the destination?*

The examination of the recorded history of the Elba destination and the comparison with the characteristics of the stakeholders network at different times has allowed the building of a growth model for the network. This is a variation of the basic preferential attachment model which is able to reproduce the scale-free topology of today's network.

The variations concern the limitations of high degree nodes which, at some early stage of development, have only a limited knowledge of the rest of the network and, as such, produce the degree distribution observed. At subsequent times this knowledge is completed and the final configuration can be explained. The impact of the network modifications (growth in density and clustering) on the development of the destination have been highlighted. This model has been related to the widely accepted qualitative destination evolution model used in the field: Butler's life cycle model. The general agreement demonstrates the reliability of the model proposed.

8 Conclusions

The original motivation for this research was to satisfy a need to explore the structure and the dynamics of a tourism destination by applying the recent advances made in the area of network analysis. The central research question formulated in the introductory chapter (Chapter 1) of this thesis was: *'How do tourism destination networks evolve over time?'*

This main research question has been divided into a number of subquestions:

- RQ1: What is the structure of the tourism destination network?
- RQ2: How can we build an evolutionary model that reproduces the present structure?
- RQ3: How does the network structure affect dynamic processes?
- RQ4: How can the network be optimised?
- RQ5: How does this model relate to the history of the destination?
- RQ6: How do the real and technological networks compare in a tourism destination?

8.1 A summary of the thesis

Different disciplines have been called into play and the models and techniques chosen for this study have been discussed in Chapters 2, 3 and 4. The quantitative network approach used stems from the mixed tradition formed by the social network analysis methods linked to the very recent instruments made available by the community of mathematicians and physicists who have worked on the analysis of complex networks. The general framework of complexity and chaos theory, combined with a statistical physics approach, has proved very effective in providing a strong theoretical basis for these methods. More specifically, Chapter 2 has introduced basic notions from graph theory and presented methods to quantify various properties of network graphs from a static point of view. Chapter 3 has discussed the behaviour of evolving and growing networked systems and several dynamic processes which may occur on them, along with their use to simulate real world networks. The main ideas which come from the tradition of industrial economics studies and the most important aspects of a tourism destination have been described in Chapter 4.

The theoretical background for the research work has been finally defined as the combination, or better the intersection, of these three main areas: complex network theory, industrial economics and tourism destination studies.

The choice and the discussion of data collection methods, network static and dynamic metrics and simulation algorithms which form the methodology used in this study have been presented in Chapter 5. The main criteria which led to the selection of the tourism destination used as test bench in this work, the island of Elba (Italy) have been highlighted in the same chapter.

The outcomes of the investigation and the discussion of their interpretation have been given in Chapters 6 and 7; they have also provided the specific answers to the research questions.

The Elba tourism destination network has been characterised as a complex network whose main traits are common to many other natural and artificial systems. Its *scale-freeness* has been assessed. Despite this similarity, the structure differs from those exhibited by other complex systems mainly in its degree of connectivity. Even the modularity of the whole network seems to be quite low. The interpretation given, based on the extent of collaboration among the destination stakeholders, has been cross verified with the outcomes of other studies (RQ1). The simulations performed have shown how and to what extent the topology of the network is responsible for the robustness of the system, the evolution of dynamic diffusion processes and the effects of network optimisation on these processes (RQ2 and RQ5). The comparison between the web and the real networks has allowed us to gather some insights into the usage of information and communication technologies and to assess the substantial similarity between the two topologies (RQ6). Finally, a model for the network's dynamic evolution has been built and compared with the historical development of the destination (RQ3 and RQ4).

8.1.1 Research limitations and boundaries

The boundaries of this work have already been uncovered in the introduction to this thesis and in the Methods chapter (Chapter 5). Basically, only 'formal' relationships between tourism stakeholders have been taken into account, and they have been considered all of equal importance. No attempt has been made to differentiate them and no consideration has been given to a different type of network that, undoubtedly, can influence the behaviour of a tourism system such as personal social relationships (e.g. kinship or friendship) amongst people working in the field. Being interested in the topological characteristics of a tourism destination network, and in its dynamic evolution, the functioning of the single elements (the network actors) has been disregarded along with the influence this may have on the properties studied here. The only exception was the study on the diffusion process in which actors were assigned different properties (capability to transfer information).

The most important point to note, however, is that all the results presented here concern a single case. Although the usage of a single case study is well acknowledged in the sociology and economic literature, this poses a limit to the general validity of the outcomes of this research. Before being able to generalise some or all of the outcomes, a much more extensive collection of similar studies should be assembled. Nonetheless, given the basic methodological nature of the work presented here, this is not seen as a major limitation for the present thesis.

8.2 Contributions of this thesis and further research

Besides answering to the research questions posed at the beginning of this work, this thesis has made a number of significant contributions to the study of tourism destinations.

The first, and most important contribution, is of methodological nature. The combination of models and techniques drawn from different disciplines has been synthesised in order to develop a uniform set of tools which has proved effective for the structural analysis of a tourism destination. As has been noted several times in the discussion contained in this chapter, both quantitative and qualitative instruments are necessary to fully exploit the potential of the methods presented here. This work, therefore, strongly supports the idea, already expressed by many scholars, that triangulation of research methods can give the clues necessary to (Davies, 2003: 110):

> lead to a 'truer analysis' of business behaviour and hence more purposeful investigation of hotels, tour operators, travel agents and the business of tourism in general.

In this general framework, three contributions have been made. The first one is the successful use of quantitative measurement for the assessment of the extent of collaborative and cooperative characteristics of the activities of the stakeholders in a tourism destination. For the scholar, this can greatly help in confirming those models which see this phenomenon as a crucial determinant for the balanced development of the system. For those interested or involved in managing a destination, the combination of both traditional qualitative evaluations and quantitative measurements can give more strength to the decisions made and better inform the actions and policies needed.

The methods proposed here rely on a set of data collection techniques which can be replicated without major difficulties. However, as mentioned previously, the process is still long, cumbersome and complex. In this respect, demonstrating the possibility of using web derived data to perform the analyses is an important contribution. This means that there is an effective and relatively fast way to gather a significant sample useful for the analysis. Even considering the limitations inherent in a 'network sampling' scheme, this result can greatly enhance the capabilities to efficiently use the methods proposed here. Furthermore, the analysis of the technological network has proved valuable in providing an assessment of the use of these tools, fundamental in today's competitive globalised market.

As has been discussed, the numerical network representation allows a wide range of simulations. Even a very simplified model, such as the diffusion mechanism presented in this work, is able to provide insightful results. This is an important outcome. Tourism is a complex phenomenon and this complexity soundly characterises the socio-economic system denoted as a tourism destination, which is the essential unit of analysis for the understanding of the whole sector. Its behaviour can be well considered to be in that ideal phase space region between a completely ordered conduct and a completely disordered one which is also known as the *edge of chaos*. This idea has been intuitively with us for a long time. However, only in recent years has a group of scholars considered that a linear deterministic description is largely insufficient to explain the behaviour of a system whose components interact in so many different ways (Farrell & Twining-Ward, 2004; Faulkner & Russell, 1997; McKercher, 1999; Russell, 2006; Russell & Faulkner, 2004). The relationships among the different tourism operators can be highly nonlinear and the whole ensemble can exhibit features which cannot be (or can be with enormous difficulty) derived by meaningful compositions of those of the single components. This is even more noticeable when considering dynamic behaviours. It is quite impossible to explain in simple terms why for some cases the system is able to resist huge external shocks (natural disasters, for example) when a similar system can be disrupted after an avalanche created by some seemingly insignificant event. It is difficult to understand why the action of some entrepreneur can act as a catalyst for incredible socio-economic growth in some cases, while in other situations similar behaviours do not have any recognisable effects. These motivations are those cited by the growing and important literature on the subject.

Probably the most important result of this vision is the claimed impossibility of fully predicting the dynamic evolution of the system and of recognising that a successful management of a complex adaptive system needs to be adaptive itself. Numerical simulations seem to be the only real possibilities to overcome, at least partially, these difficulties and provide a range of solutions which will (Bankes, 2002: 7266) "allow users to iterate with the computer to gradually evolve policy schemas that have particular policy instances with desirable properties".

The dynamic progress of the destination has been related to the modifications of the network topology and proved models have been invoked to explain this evolution. Again, even with the limitations discussed, this result can prove extremely useful, mainly when combined with simulation techniques, to help understand (theoretically) the evolutionary mechanisms and to derive practical insights useful in informing a destination's management and supervision activities.

The limitations considered in section 8.1.1 can draw a path for possible future research. The established use of analogy and the general framework of statistical mechanics in which network theories are embedded give us a solid justification for extending the results obtained in diverse fields to the tourism one. However, there is little doubt that the compilation of a richer catalogue of cases, studied with network science methods, will help in assessing (or confuting) the results of this thesis and better explore the possibilities, both from a theoretical and a practical viewpoint, of this approach. First and foremost, the relations between the evolution of a destination and the time modifications of the underlying network, combined with more sophisticated simulation models, can provide very influential and authoritative means of helping those fundamental forecasting activities which are at the basis of all strategic and tactical management practices of a tourism destination. The effects of different linkage typologies and intensities, and the presence of a directionality in them, can provide deeper understanding of the aggregation mechanisms in a destination. Finally, the relationships between a tourism destination and other similar networks, or those which can exist in other sectors of the economy can be examined through the analysis of the reciprocal connections. For example, prior research (Bouncken & Sungsoo, 2002; Cooper, 2006; Pechlaner et al., 2002) has shown that knowledge management is an important element in the management of tourism destinations and that the beneficial effects of effective exchanges of information can be greatly improved in networks with certain characteristics and with certain preferential connective paths to the external socio-economic environment (Dredge, 2005; Nordin, 2003; Novelli et al., 2006; Stamboulis & Skayannis, 2003).

8.3 A concluding remark

In the last few years, the techniques and methods of complex networks analysis have developed significantly. An increasing literature has studied a great variety of theoretical and empirical aspects of networks of many types. When this research work started, tourism systems, in spite of their recognised *networked* characteristics, were almost absent from these investigations. The present contribution is one of the first attempts at using quantitative network analysis models in the study of tourism destinations.

As has been discussed, the exploratory nature of the analysis performed here and the restriction to a single case put an obvious limit on the generalisation of our results. However, this thesis has given evidence of the possibility to assess quantitatively the structural characteristics of a network of tourism stakeholders and to simulate its evolutionary path. The coupling of sound quantitative analysis methods with qualitative interpretations of the main characteristics of a tourism destination has proved extremely useful. Future and more extensive work, already under way, will reinforce the conclusions of this thesis, mainly from a methodological point of view, and will confirm the effectiveness of the combined approach used here.

As a final point, it is a firm conviction of the author that a more rigorous establishment of methodological tools such as those used in this work, can be a powerful way to help a transition towards a less *undisciplined* set of theories and models in the tourism arena, and that this can be greatly beneficial for the understanding of the structure and behaviour of this system and its components, so important in today's social and economic setting.

... et, quaecumque viam dederit Fortuna, sequatur

Publius Vergilius Maro , Aeneis X.49

References

Aas, C., Ladkin, A., & Fletcher, J. (2005). Stakeholder collaboration and heritage management. *Annals of Tourism Research, 32*(1), 28-48.

Adamic, L. A. (2000). *Zipf, Power-laws, and Pareto - a ranking tutorial* (Working paper): Xerox Palo Alto Research Center, Palo Alto, CA, USA. Retrieved May 2006, from http://www.hpl.hp.com/research/idl/papers/ranking/.

Adamic, L. A., & Adar, E. (2003). Friends and Neighbors on the Web. *Social Networks, 25*(3), 211-230.

Adomavicius, G., & Tuzhilin, A. (2005). Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering, 17*(6), 734-749.

Adrion, W. R., Branstad, M. A., & Cherniavsky, J. C. (1982). Validation, Verification, and Testing of Computer Software. *ACM Computing Surveys, 14*(2), 159-192.

Agarwal, S. (1994). The resort cycle revisited: Implications for resorts. In C. Cooper & A. Lockwood (Eds.), *Progress in Tourism, Recreation and Hospitality Management* (Vol. 5, pp. 194-208). Chichester, UK: Wiley.

Agarwal, S. (1997). The resort cycle and seaside tourism: an assessment of its applicability and validity *Tourism Management, 18*(2), 65-73.

Agarwal, S. (2002). Restructuring Seaside Tourism: The Resort Lifecyle. *Annals of Tourism Research, 29*(1), 25-55.

Agostinho, M. E., & Teixeira de Castro, G. (2003). *Co-Creating a Self-Organizing Management System: A Brazilian Experience.* Paper presented at the Complexity, Ethics and Creativity Conference, London School of Economics, Retrieved 17-18 Sept.2003, from http://www.psych.lse.ac.uk/complexity/Conference/AgostinhoCastro.pdf.

Ahuja, M. K., & Carley, K. M. (1999). Network Structure in Virtual Organizations. *Organization Science, 10*(6), 741-757.

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Review of Modern Physics, 74*, 47-91.

Albert, R., Jeong, H., & Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *Nature, 406*, 378–382.

Amaral, L. A. N., & Ottino, J. M. (2004). Complex networks - Augmenting the framework for the study of complex systems. *The European Physical Journal B, 38*, 147-162.

Amaral, L. A. N., Scala, A., Barthélémy, M., & Stanley, H. E. (2000). Classes of small world networks. *Proceedings of the National Academy of the Sciences of the USA, 97*, 11149-11152.

Anderson, R. M., & May, R. M. (1992). *Infectious diseases in humans*. Oxford: Oxford University Press.

Andersson, T., Schwaag-Serger, S., Sörvik, J., & Hansson, E. W. (2004). *The Cluster Policies Whitebook*. Malmö, S: IKED - International Organisation for Knowledge Economy and Enterprise Development.

Angeloudis, P., & Fisk, D. (2006). Large subway systems as complex networks. *Physica A 367* 553-558.

Antonioli Corigliano, M. (1999a). I distretti turistici e le aggregazioni tra attori per lo sviluppo del prodotto-destinazione. In M. Colantoni (Ed.), *Turismo: una tappa per la ricerca* (pp. 99-140). Bologna: Patron Editore.

Antonioli Corigliano, M. (1999b). *Strade del vino ed enoturismo. Distretti turistici e vie di comunicazione*. Milano: Franco Angeli.

Antonioli Corigliano, M. (Ed.). (2000). *L' osservatorio turistico: obiettivi, metodologie, strumenti*. Milano: EGEA.

Antonioli Corigliano, M., & Baggio, R. (2003). Italian tourism virtual communities: empirical research and model building. In A. J. Frew, M. Hitz & P. O'Connor (Eds.), *Information and Communication Technologies in Tourism 2003 - Proceedings of the International Conference in Helsinki, Finland* (pp. 86-95). Wien: Springer.

Antonioli Corigliano, M., Viganò, G., & Baggio, R. (2007). *Le strade dei vini e dei sapori a una svolta* (unpublished report).

Archdale, A. (1995). New frontiers for tourism technology. *Insights: Tourism Marketing Intelligence Service, 6*(9), D27- 31.

Arenas, A., Danon, L., Díaz-Guilera, A., Gleiser, A. P., & Guimera, R. (2004). Community analysis in social networks. *The European Physics Journal B, 38*(2), 373-380.

Arthur, W. B. (1989). Competing Technologies, Increasing Returns, and Lock-In by Historical Events. *Economic Journal, 99*(394), 116-131.

Arthur, W. B., Durlauf, S. N., & Lane, D. (Eds.). (1997). *The Economy as an Evolving Complex System II*. Reading, MA: Addison-Wesley.

Axelrod, R. (1997). The Dissemination of Culture: A Model with Local Convergence and Global Polarization. *The Journal of Conflict Resolution, 41*(2), 203-226.

Axelrod, R., Mitchell, W., Thomas, R. E., Bennett, D. S., & Bruderer, E. (1995). Coalition formation in standard-setting alliances. *Management Science*(41), 1493-1508.

Aydalot, P. (1985). *Économie régionale et urbaine*. Paris: Économica.

Aydalot, P. (1986). *Milieux innovateurs en Europe*. Paris: Groupe de Recherche Européen sur les Milieux Innovateurs (GREMI).

Baggio, R. (2006a). Complex systems, information technologies and tourism: a network point of view. *Information Technology and Tourism, 8*(1), 15-29.

Baggio, R. (2006b). *Consorzio L' Altra Maremma: la strategia Internet di promozione* (unpublished report).

Baggio, R. (2006c). *Mizar: Programma USA: analisi del mercato high-end e promozione online* (unpublished report).

Baggio, R. (2007). The Web Graph of a Tourism System. *Physica A, 379*(2), 727-734.

Baggio, R. (2008). Symptoms of complexity in a tourism system. *Tourism Analysis, 13*(1), 1-20.

Baggio, R., & Antonioli Corigliano, M. (2007). Tourism destinations: a network analysis of the Web space. *Proceedings of the Advances in Tourism Marketing Conference (ATMC), Valencia, Spain, 10-12 September*.

Baggio, R., Antonioli Corigliano, M., & Tallinucci, V. (2007a). The websites of a tourism destination: a network analysis. In M. Sigala, L. Mich & J. Murphy (Eds.), *Information and Communication Technologies in Tourism 2007 - Proceedings of the International Conference in Ljubljana, Slovenia* (pp. 279-288). Wien: Springer.

Baggio, R., Scott, N., & Wang, Z. (2007b). What network analysis of the WWW can tell us about the organisation of tourism destinations. *Proceedings of the CAUTHE 2007, Sydney, Australia, 11-14 February*.

Bagnasco, A. (1977). *Tre Italie. La problematica territoriale dello sviluppo italiano*. Bologna: Il Mulino.

Baiesi, M., & Manna, S. S. (2003). Scale free networks from a Hamiltonian dynamics. *Physical Review E, 68*, 047103.

Bailey, N. (1975). *The Mathematical Theory of Infectious Diseases and its Applications* (2nd ed.). London: Griffin.

Bak, P. (1996). *How nature works. The science of self-organized criticality*. New York: Springer.

Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-Organized Criticality. An Explanation of 1/f Noise. *Physical Review Letters, 59*, 381-384.

Bak, P., Tang, C., & Wiesenfeld, K. (1988). Self-Organized Criticality. *Physical Review A, 38*, 364-374.

Balázs, K., Hastings, M. B., & Korniss, G. (2005). Diffusion Processes on Power-Law Small-World Networks. *Physical Review Letters, 95*, 018701.

Balci, O. (2003). Verification, Validation, and Certification of Modeling and Simulation Applications. *Proceedings of the 2003 Winter Simulation Conference, Piscataway, NJ*, 150-158.

Baldwin, R. E., & Forslid, R. (2000). The Core-Periphery Model and Endogenous Growth. *Economica 67*, 307-324.

Ball, P. (2002). The physics of society. *Nature 415*, 371.

Ball, P. (2003). The Physical Modelling of Human Social Systems. *Complexus 1*, 190-206.

Bankes, S. C. (2002). Tools and techniques for developing policies for complex and uncertain systems. *Proceedings of the National Academy of the Sciences of the USA, 99* (suppl. 3), 7263-7266.

Bar-Yam, Y. (1997). *Dynamics of Complex Systems*. Reading, MA: Addison-Wesley.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286*, 509-512.

Barrat, A., & Weigt, M. (2000). On the properties of small-world network models. *European Physical Journal B, 13*, 547-560.

Barrell, R., & Pain, N. (1999). Domestic Institutions, Agglomerations, and Foreign Direct Investment in Europe. *European Economic Review, 43*, 925-934.

Barthélemy, M., Barrat, A., Pastor-Satorras, R., & Vespignani, A. (2005). Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. *Journal of Theoretical Biology, 235*, 275-288.

Barua, A., Whinston, A. B., & Yin, F. (2000). Value and Productivity in the Internet Economy. *Computer, 33*(5), 102-105.

Batagelj, V., & Mrvar, A. (1998). PAJEK (Version 1.15 - 2006). Ljubljana, Slovenia: available at: vlado.fmf.uni-lj.si/pub/networks/pajek/.

Baum, T. (1998). Taking the Exit Route: Extending the Tourism Area Life Cycle Model. *Current Issues in Tourism, 1*(2), 167-175.

Becattini, G. (1979). Dal Settore Industriale al Distretto Industriale. Alcune considerazioni sull'unità di indagine dell'economia industriale. *Rivista di Economia e Politica Industriale, 1*, 1-8.

Becattini, G. (1990). The Marshallian industrial district as a socio-economic notion. In F. Pyke, G. Becattini & W. Sengenberger (Eds.), *Industrial Districts and Inter-firm Co-operation in Italy* (pp. 37-51). Geneva: International Institute for Labour Studies.

Becattini, G. (Ed.). (1987). *Mercato e forze locali: il distretto industriale*. Bologna: Il Mulino.

Belussi, F. (2004). *In search of a useful theory of spatial clustering.* Paper presented at the The 4th Congress on Proximity Economics: Proximity, Networks and Co-ordination, Marseille, France, Retrieved June 17 - 18, 2004, Belussi, F., Gottardi, G., & Rullani, E. (Eds.). (2003). *The Technological Evolution of Industrial Districts*. Boston: Kluwer.

Bernard, H. R., Killworth, P. D., Kronenfeld, D., & Sailer, L. (1984). The problem of Informant Accuracy: The Validity of Retrospective Data. *Annual Review of Anthropology, 13*, 495-517.

Bernardes, A. T., Stauffer, D., & Kertész, J. (2002). Election results and the Sznajd model on Barabasi network. *The European Physical Journal B, 25*, 123-127.

Bernstein, S., Lebow, R. N., Stein, J. G., & Weber, S. (2000). God Gave Physics the Easy Problems: Adapting Social Science to an Unpredictable World. *European Journal of International Relations, 6*(1), 43-76.

Berry, F., Brower, R., Choi, S. O., Goa, W. X., Jang, H. S., Kwon, M., & Word, J. (2004). Three Traditions of Network Research: What the Public Management Research Agenda Can Learn from Other Research Communities. *Public Administration Review, 64*(5), 539-552.

Bettencourt, L. M. A., Cintrón-Arias, A., Kaiser, D. I., & Castillo-Chávez, C. (2006). The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. *Physica A, 364*, 513-536.

Bianconi, G., & Barabási, A.-L. (2001). Bose-Einstein Condensation in Complex Networks. *Physical Review Letters, 86*(24), 5632-5635.

Bickerdyke, I. (1996). Australia: The evolving structure and strategies of business networks. In OECD (Ed.), *Networks of Enterprises and Local Development: Competing and Co-operating in Local productive Systems* (pp. 203-216). Paris: Organisation for Economic Co-operation and Development.

Biggiero, L. (1999). Markets, hierarchies, networks, districts: A cybernetic approach. *Human Systems Management, 18*, 71-86.

Biggs, N. L., Lloyd, E. K., & Wilson, R. J. (1976). *Graph Theory 1736-1936*. Oxford: Clarendon Press.

Birnie, S. A., & Horvath, P. (2002). Psychological predictors of Internet social communication. *Journal of Computer-Mediated Communication [On-line] 7*(4). Retrieved April, 2006, from http://jcmc.indiana.edu/vol7/issue4/horvath.html.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics Reports, 424*(4-5), 175-308.

Boguna, M., Pastor-Satorras, R., Diaz-Guilera, A., & Arenas, A. (2004). Models of social networks based on social distance attachment. *Physical Review E, 70*(5), 056122-056128.

Bollobás, B. (1998). *Modern Graph Theory*. New York: Springer.

Bonacich, P. (1972). Factoring and Weighting Approaches to Status Scores and Clique Identification. *Journal of Mathematical Sociology, 2*, 113-120.

Bonacich, P. (1987). Power and Centrality: A Family of Measures. *American Journal of Sociology, 92*, 1170-1182.

Bondy, J. A., & Murty, U. S. R. (1982). *Graph Theory with Applications* (5th ed.). New York: Elsevier.

Bönke, D., Ammann, E., & Zabel, J. (2001). Knowledge Engineering in Virtual Organisations. In B. Stanford-Smith & E. Chiozza (Eds.), *E-work and E-commerce - Novel solutions and practices for a global networked economy* (Vol. 2, pp. 728-734). Amsterdam: IOS Press.

Boote, D. N., & Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. *Educational Researcher, 34*(6), 3-15.

Borgatti, S. P., Everett, M. G., & Freeman, L. C. (1992). UCINET (Version 6 - 2006). Harvard, MA: Analytic Technologies (www.analytictech.com).

Bornholdt, S., & Schuster, H. G. (Eds.). (2002). *Handbook of Graphs and Networks - From Genome to the Internet*. Berlin: Wiley-VCH.

Bouchaud, J.-P. (2001). Power laws in economics and finance: some ideas from physics. *Quantitative Finance, 1*, 105-112.

Bouncken, R., & Sungsoo, P. (Eds.). (2002). *Knowledge Management in Hospitality and Tourism*. New York: The Haworth Hospitality Press.

Bramwell, B., & Lane, B. (2000). *Tourism Collaboration and Partnerships: Politics Practice and Sustainability*. Clevedon, UK: Channel View Publications.

Brenner, T. (2000). *The evolution of localised industrial clusters: identifying the processes of self-organisation* (Papers on economics & evolution, no. 0011). Jena: Max Planck Institute.

Brenner, T., & Weigelt, N. (2001). The Evolution of Industrial Clusters - Simulating Spatial Dynamics. *Advances in Complex Systems, 4*, 127-147.

Brigandt, I. (2004). *The role a concept plays in science: The case of homology* (Manuscript). Pittsburgh, PA: University of Pittsburgh, Department of History and Philosophy of Science. Retrieved April 2006, from http://www.pitt.edu/~inb1/role.pdf.

Brin, S., & Page, L. (1982). The Anatomy of a Large-Scale Hypertextual (Web) Search Engine. *Computer Networks and ISDN Systems, 30*(1-7), 107-117.

Britton, J. N. H. (2003). Network Structure of an Industrial Cluster: Electronics in Toronto. *Environment and Planning A, 35*, 983-1006.

Broder, A. Z., Kumar, S. R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., & Wiener, J. L. (2000). Graph structure in the web. *Computer Networks, 33*(1-6), 309-320.

Buhalis, D. (2000). Marketing the Competitive Destination of the Future. *Tourism Management, 21*, 97-116

Buhalis, D. (2003). *eTourism: Information technology for strategic tourism management*. Harlow, UK: Pearson/Prentice-Hall.

Burns, P. M. (1999). *An introduction to tourism and anthropology*. London: Routledge.

Burt, R. S. (1980). Models of Network Structure. *Annual Review of Sociology, 6*, 79-141.

Butler, R. W. (1980). The Concept of a Tourist Area Cycle of Evolution: Implications for Management of Resources. *Canadian Geographer, 24* (1), 5-12.

Butler, R. W. (Ed.). (2005a). *The Tourism Area Life Cycle, Vol. 1: Applications and Modifications* (Vol. 1). Clevedon, UK: Channel View.

Butler, R. W. (Ed.). (2005b). *The Tourism Area Life Cycle, Vol. 2: Conceptual and Theoretical Issues* (Vol. 2). Clevedon, UK: Channel View.

Caldarelli, G. (2007). *Scale-Free networks: complex webs in nature and technology*. Oxford: Oxford University Press.

Capone, F. (2004). *Regional competitiveness in tourist local systems* (ERSA conference papers: ersa04p627). Porto, Portugal: 44th European Congress of the European Regional Science Association (ERSA). Retrieved December 2005, from http://www.ersa.org/ersaconfs/ersa04/PDF/627.pdf.

Carlsen, J. (1999). A systems approach to island tourism destination management. *Systems Research and Behavioral Science, 16*(4), 321-327.

Castellano, C., Fortunato, S., & Loreto, V. (2007). *Statistical physics of social dynamics* (arXiv:physics/0710.3256). Retrieved October 2007, from http://arxiv.org/abs/0710.3256.

Castellano, C., Marsili, M., & Vespignani, A. (2000). Nonequilibrium phase transition in a model for social influence. *Physical Review Letters, 85*, 3536-3539.

Castells, M. (2000). *The Information Age: Economy, Society and Culture: The Rise of the Network Society* (2nd ed.). Oxford: Blackwell.

Castilla, E. J., Hokyu, H., Granovetter, E., & Granovetter, M. (2000). Social networks in Silicon Valley. In C.-M. Lee, W. F. Miller, M. G. Hancock & H. S. Rower (Eds.), *The Silicon Valley Edge: A Habitat for Innovation and Entrepreneurship* (pp. 218-247). Stanford: Stanford University Press.

Catanzaro, M., Caldarelli, G., & Pietronero, L. (2004). Assortative model for social networks. *Physical Review E, 70*, 037101

Chen, C., & Hicks, D. (2004). Tracing knowledge diffusion. *Scientometrics, 59*(2), 199-211.

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E, 70*, 066111.

Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2007). *Power-law distributions in empirical data* (arXiv/physics/0706.1062). Retrieved July 2007, from http://arxiv.org/abs/0706.1062.

Cochran, W. G. (1977). *Sampling Techniques* (3rd ed.). New York: John Wiley.

Cohen, E. (1984). The Sociology of Tourism: Approaches, Issues, and Findings. *Annual Review of Sociology, 10*, 373-392.

Cohen, W., & Levinthal, D. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly, 35*, 128-152.

Cooper, C. (2006). Knowledge Management and Tourism. *Annals of Tourism research, 33* (1), 47-64.

Cooper, C., Fletcher, J., Gilbert, D., Fayall, A., & Wanhill, S. (2005). *Tourism Principles and Practice* (3rd ed.). Harlow, UK: Pearson Education.

Cooper, C., & Jackson, S. (1989). Destination Lifecycle: The Isle of Man Case Study. *Annals of Tourism Research, 16*(3), 377-398.

Cooper, C., & Scott, N. (2005). *Structure of the Australian Tourism Sector*. Gold Coast: Sustainable Tourism Cooperative Research Centre.

Costa Filho, R. N., Almeida, M. P., Andrade, J. S., & Moreira, J. E. (1999). Scaling behavior in a proportional voting process. *Physical Review E, 60*, 1067-1068.

Costenbader, E., & Valente, T. W. (2003). The stability of centrality measures when networks are sampled. *Social Networks, 25*, 283-307.

Coulon, F. (2005). *The use of social network analysis in innovation research: A literature review.* Paper presented at the DRUID Academy Winter 2005 PhD Conference on Industrial Evolution and Dynamics, Skørping, Denmark Retrieved January 27-29,

Creswell, J. W. (1998). *Qualitative inquiry and research design: choosing among five traditions*. Thousand Oaks, Calif.: Sage Publications.

Creswell, J. W. (2003). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Crucitti, P., Latora, V., & Marchiori, M. (2004a). Model for cascading failures in complex networks. *Physical Review E, 69*, 045104(R).

Crucitti, P., Latora, V., Marchiori, M., & Rapisarda, A. (2003). Efficiency of scale-free networks: error and attack tolerance. *Physica A, 320*, 622-642.

Crucitti, P., Latora, V., Marchiori, M., & Rapisarda, A. (2004b). Error and attack tolerance of complex networks. *Physica A, 340*, 388-394.

Crutchfield, J. P., & Young, K. (1990). Computation at the Onset of Chaos. In W. Zurek (Ed.), *Entropy, Complexity, and the Physics of Information [SFI Studies in the Sciences of Complexity]* (Vol. VIII, pp. 223-269). Reading, MA: Addison-Wesley.

Csermely, P. (2004). Strong links are important, but weak links stabilize them. *TRENDS in Biochemical Sciences, 29* (7), 331-334.

da Fontoura Costa, L., Rodrigues, A., Travieso, G., & Villas Boas, P. R. (2007). Characterization of complex networks: A survey of measurements. *Advances in Physics, 56*(1), 167-242.

Daniel, V. (1955). The Uses and Abuses of Analogy. *OR, 6*(1), 32-46.

Danon, L., Díaz-Guilera, A., Duch, J., & Arenas, A. (2005). Comparing community structure identification. *Journal of Statistical Mechanics*, P09008

Davies, B. (2003). The Role of Quantitative and Qualitative Research in Industrial Studies of Tourism. *International Journal of Tourism Research, 5*, 97-111.

de Aguiar, M. A. M., & Bar-Yam, Y. (2005). Spectral analysis and the dynamic response of complex networks. *Physical Review E, 71*, 016106.

de Arcangelis, L., & Herrmann, H. J. (2002). Self-organized Criticality on Small World Networks. *Physica A, 308*(1-4), 545-549.

Decrop, A. (1999). Triangulation in qualitative tourism research. *Tourism Management, 20*, 157-161.

Demetrius, L., & Manke, T. (2005). Robustness and network evolution - an entropic principle. *Physica A, 346* 682-696.

Denzin, N. K. (1989). *The research act: a theoretical introduction to sociological methods* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Deo, N., & Gupta, P. (2001). Graph-Theoretic Web Algorithms: An Overview. In T. Böhme & H. Unger (Eds.), *Lecture Notes in Computer Science* (Vol. 2026, pp. 91-102). Berlin: Springer.

Diekmann, O., & Heesterbeek, J. (2000). *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*. New York: John Wiley & Sons.

Diestel, R. (2005). *Graph Theory, Electronic edition 2005*. New York: Springer. Retrieved December, 2005, from http://www.math.uni-hamburg.de/home/diestel/books/graph.theory/GraphTheoryIII.pdf.

Dill, S., Kumar, S. R., McCurley, K., Rajagopalan, S., Sivakumar, D., & Tomkins, A. (2002). Self similarity in the web. *ACM Transactions on Internet Technology (TOIT), 2*(3 - August), 205-223.

Donetti, L., & Muñoz, M. A. (2004). Detecting network communities: a new systematic and efficient algorithm. *Journal of Statistical Mechanics, 10*, P10012.

Doreian, P., & Woodard, K. L. (1992). Fixed list versus snowball selection of social networks. *Social Science Research, 21*(2), 216-233.

Dorogovtsev, S. N., & Mendes, J. F. F. (2002). Evolution of networks. *Advances in Physics, 51*, 1079-1187.

Dorogovtsev, S. N., & Mendes, J. F. F. (2003). *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford: Oxford University Press.

Dredge, D. (2005). Networks and Innovation in Lake Macquarie. In D. Carson & J. Macbeth (Eds.), *Regional Tourism Cases: Innovation in regional tourism* (pp. 61-68). Gold Coast: STCRC.

Dredge, D. (2006). Networks, Conflict and Collaborative Communities. *Journal of Sustainable Tourism, 14*(6), 562-581.

Dubin, R. (1978). *Theory building*. New York: Free Press.

Dukes, S. (1984). Phenomenological methodology in the human sciences. *Journal of Religion and Health, 23*(3), 197 - 203.

Dunne, J. A., Williams, R. J., & Martinez, N. D. (2002). Food-web structure and network theory: The role of connectance and size. *Proceedings of the National Academy of the Sciences of the USA, 99*(20), 12917-12922.

Durlauf, S. N. (1997). Limits to Science or Limits to Epistemology? *Complexity, 2*(3), 31-37.

Durrett, R. (2006). *Random Graph Dynamics*. Cambridge: Cambridge University Press.

Dwyer, L., & Kim, C. W. (2003). Destination Competitiveness: a Model and Indicators. *Current Issues in Tourism, 6*(5), 369-414.

Dyer, J. H., & Singh, H. (1998). The relational view: Cooperative strategy and sources of interorganizational competitive advantage. *Academy of Management Review, 23*(4), 660–679.

Ebel, H., Davidsen, J., & Bornholdt, S. (2003). Dynamics of social networks. *Complexity, 8*(2), 24-27.

Echtner, C. M., & Jamal, T. B. (1997). The Disciplinary Dilemma of Tourism Studies. *Annals of Tourism Research, 24*(4), 868-883.

Einstein, A. (1949). Reply to Criticisms. In P. A. Schilpp (Ed.), *Albert Einstein: Philosopher-Scientist* (pp. 665-688). New York: Tudor.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review, 14*(4), 532-550.

Emirbayer, M. (1997). Manifesto for a Relational Sociology. *The American Journal of Sociology, 103*(2), 281-317.

Emirbayer, M., & Goodwin, J. (1994). Network Analysis, Culture, and the Problem of Agency. *The American Journal of Sociology, 99*(6), 1411-1454.

Erdös, P., & Rényi, A. (1959). On random graphs. *Publicationes Mathematicae (Debrecen), 6*, 290-297.

Erdös, P., & Rényi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences, 5*, 17-61.

Erdös, P., & Rényi, A. (1961). On the strength of connectedness of a random graph. *Acta Mathematica Academiae Scientiarum Hungaricae, 12*, 261-267.

Erickson, B. H., & Nosanchuk, T. A. (1983). Applied Network Sampling. *Social Networks, 5*, 367-382.

Euler, L. (1736). Solutio problematis ad geometriam situs pertinentis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae, 8*, 128-140.

European Commission. (2001). *Methodology for Regional and Transnational Technology Clusters: Learning with European Best Practices*. Brussels: European Commission, Enterprise Directorate General.

EUROSTAT. (2000). *Community methodology on tourism statistics*. Luxembourg: Office for Official Publications of the European Communities.

EUROSTAT. (2002). *Methodology of short-term business statistics: Interpretation and guidelines*. Luxembourg: Office for Official Publications of the European Communities.

EUROSTAT. (2006). *Panorama on Tourism*. Luxembourg: Office for Official Publications of the European Communities.

Faloutsos, M., Faloutsos, P., & Faloutsos, C. (1999). On power-law relationships of the Internet topology. *Computer Communication Review, 29*, 251-262.

Farkas, I. J., Derényi, I., Barabási, A.-L., & Vicsek, T. (2001). Spectra of "real-world" graphs: Beyond the semicircle law. *Physical Review E, 64*, 026704.

Farrell, B. H., & Twining-Ward, L. (2004). Reconceptualizing Tourism. *Annals of Tourism Research, 31*(2), 274-295.

Faulkner, B. (2000). *The future ain't what it used to be: Coping with Change, Turbulence and Disasters in Tourism Research and Destination Management* (Professorial Lecture Series No. 6, 2000): Griffith University. Retrieved July 2006, from http://www.gu.edu.au/ins/collections/proflects/faulkner00.pdf.

Faulkner, B., & Russell, R. (1997). Chaos and complexity in tourism: in search of a new perspective. *Pacific Tourism Review, 1*, 93-102.

Ferrer i Cancho, R., & Solé, R. V. (2003). Optimization in complex networks. In R. Pastor-Satorras, M. Rubi & A. Diaz-Guilera (Eds.), *Statistical Mechanics of Complex Networks, Lecture Notes in Physics* (Vol. 625, pp. 114-125). Berlin: Springer.

Ferro Luzzi, G., & Flückiger, Y. (2003). An Econometric Estimation of the Demand for Tourism: the Case of Switzerland. *Pacific Economic Review, 8*(3), 289-303.

Fiedler, M. (1973). Algebraic Connectivity of Graphs. *Czechoslovak Mathematical Journal, 23*(98), 298-305.

Fioretti, G. (2005). Agent-Based Models of Industrial Clusters and Districts. In F. Columbus (Ed.), *Contemporary Issues in Urban and Regional Economics (forthcoming)*. New York: Nova Science Publishers.

Fitch, W. M. (2000). Homology: a personal view on some of the problems. *Trends in Genetics, 16*(5), 227-231.

Flake, G. W., Lawrence, S., Giles, C. L., & Coetzee, F. M. (2002). Self-Organization of the Web and Identification of Communities. *IEEE Computer, 35*(3), 66-71.

Flick, U. (2004). Triangulation in qualitative research. In U. Flick, E. v. Kardorff & I. Steinke (Eds.), *A companion to qualitative research* (pp. 178-183). London: Sage Publications.

Font, X., & Ahjem, T. E. (1999). Searching for a balance in tourism development strategies. *International Journal of Contemporary Hospitality Management, 11*(2-3), 73-77.

Framke, W. (2002a). The Destination as a Concept: A Discussion of the Business-related Perspective versus the Socio-cultural Approach in Tourism Theory. *Scandinavian Journal of Hospitality and Tourism, 2*(2), 92-108.

Framke, W. (2002b). *The 'destination': a problematic concept* (Keynote paper presented at the 10th Nordiska Forskarsymposiet i Turism, Vasa, Finland, 18-20 october 2001). Retrieved January 2006, from http://www.wasa.shh.fi/konferens/abstract/a-framke.pdf.

Frank, O. (1971). *Statistical inference in graphs*. Stockholm: Försvarets forskningsanstalt (Research Institute of National Defense).

Frank, O. (1979). Estimation of population totals by use of snowball samples. In P. W. Holland & S. Leinhardt (Eds.), *Perspectives on Social Network Research* (pp. 319-348). New York: Academic Press.

Frank, O., & Snijders, T. A. B. (1994). Estimating the size of hidden populations using snowball sampling. *Journal of Official Statistics, 10*(1), 53-67.

Fredericks, K. A., & Durland, M. M. (2005). The Historical Evolution and Basic Concepts of Social Network Analysis. *New Directions for Evaluation, 107*, 15-23.

Freeman, L. C. (2004). *The Development of Social Network Analysis: A Study in the Sociology of Science*. Vancouver Empirical Press.

Galilei, G. (1623). The Assayer (S. Drake, Trans.). In S. Drake (Ed.), *Discoveries and Opinions of Galileo* (pp. 231-280). New York: Doubleday & Co, 1957.

Gardner, D. M. (1987). Product life cycle: A critical look at the literature. In M. Houston (Ed.), *Review of Marketing* (pp. 162-194). Chicago: American Marketing Association.

Garín-Muñoz, T., & Amaral, T. P. (2000). An econometric model for international tourism flows to Spain. *Applied Economics Letters, 7*, 525-529.

Gentle, J. E., Härdle, W., & Mori, Y. (Eds.). (2004). *Handbook of Computational Statistics: Concepts and Methods*. Heidelberg Springer.

Gentner, D. (1982). Are scientific analogies metaphors? In D. S. Miall (Ed.), *Metaphor: Problems and perspectives* (2nd ed., pp. 106-132). Brighton: Harvester Press Ltd.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*(2), 155-170.

Gentner, D. (2002). Analogy in Scientific Discovery: The Case of Johannes Kepler. In L. Magnani & N. J. Nersessian (Eds.), *Model-Based Reasoning: Science, Technology, Values* (pp. 21-39). New York: Kluwer Academic / Plenum Publishers.

Gentner, D., & Jeziorski, M. (1993). The shift from metaphor to analogy in western science. In A. Ortony (Ed.), *Metaphor and Thought* (2nd ed., pp. 447-480). Cambridge: Cambridge University Press.

Georgulas, N. (1970). Tourism Destination Features. *Journal of Town Planning Institute*(56), 442-446.

Gerring, J. (2004). What is a case study and what is it good for? *American Political Science Review, 98*(2), 341-354.

Getz, D. (1992). Tourism planning and the destination lifecycle. *Annals of Tourism Research, 19*(4), 752-770.

Geymonat, L. (Ed.). (1972). *Storia del pensiero filosofico e scientifico*. Milan: Garzanti.

Giaccaria, P. (1997). Modellizzazione di un distretto industriale mediante simulazioni fondate su agenti. *Sviluppo Locale, 4*, 131-171.

Gibson, D., Kleinberg, J., & Raghavan, P. (1998). Inferring Web communities from link topology. *Proceedings of the 9th ACM Conference on Hypertext and Hypermedia*, 225-234.

Giddens, A. (1979). *Central Problems in Social Theory: Action, Structure and Contradiction in Social Action*. London: Macmillan.

Gilder, G. (1993). Metcalfe's law and legacy. *Forbes (ASAP Supplement), 13 Settembre*.

Gilmore, A., Carson, D., & Grant, K. (2001). SME marketing in practice. *Marketing Intelligence & Planning, 19*(1), 6-11.

Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of the Sciences of the USA, 99*, 7821-7826.

Giuliani, E. (2006). *Networks and heterogeneous performance of cluster firms* (Papers in Evolutionary Economic Geography #06.02). Utrecht, NL: Utrecht University. Retrieved March 2006, from http://econ.geog.uu.nl/peeg/peeg0602.pdf.

Giuliani, E., & Bell, M. (2004). *When micro shapes the meso: Learning networks in a Chilean wine cluster* (SPRU Paper No. 115). University of Sussex: The Freeman Centre. Retrieved June 2007, from http://www.sussex.ac.uk/spru/.

Giuliani, E., Rabellotti, R., & van Dijk, M. P. (Eds.). (2005). *Clusters Facing Competition: The Importance of External Linkages*. Aldershot, UK: Ashgate.

Gleick, J. (1987). *Chaos: Making a New Science*. New York: Viking.

Gnoth, J. (2002). Leveraging export brands through a tourism destination brand. *Journal of Brand Management, 9*(4-5), 262-280.

Gnoth, J. (2004). Strengthening tourism SME brands. In W. Weber (Ed.), *Proceedings of Rencontres de St-Gall: Value Creation in Entrepreneurship and SMEs, [CD-ROM]*. St Gallen, Switzerland: Swiss Research Institute of Small Business and Entrepreneurship.

Gnoth, J. (2006). The Function and Dynamics of Clusters in Destinations as Virtual Service Firms: Responding to the Future. In K. Weiermair, H. Pechlaner & T. Bieger (Eds.), *Time Shift, Leisure and Tourism - Impacts of Time Allocation on Successful Products and Services* (pp. 353-364). Berlin: Erich Schmid-Verlag.

Godlewska, A. M. C. (2002). *Atlas of Napoleonic Cartography in Italy [web-based]*. Retrieved December, 2005, from http://www.geog.queensu.ca/napoleonatlas.

Godsyl, C., & Royle, G. (2001). *Algebraic Graph Theory*. New York: Springer.

Goldstein, M. L., A., M. S., & Yen, G. G. (2004). Problems with Fitting to the Power-Law Distribution. *The European Physical Journal B, 41*, 255-258.

Gollub, J., Hosier, A., & Woo, G. (2003). *Using Cluster-Based Economic Strategy to Minimize Tourism Leakages*. Madrid: UN World Tourism Organization.

Gomm, R., Hammersley, M., & Foster, P. (Eds.). (2000). *Case Study Method: Key Issues, Key Texts*. London: Sage.

Gopikrishnan, P., Plerou, V., Amaral, L. A. N., Meyer, M., & Stanley, H. E. (1999). Scaling of the distributions of fluctuations of financial market indices. *Physical Review E, 60*, 5305-5316.

Grabher, G. (1993). The weakness of strong ties: the lock-in of regional development in the Ruhr area. In G. Grabher (Ed.), *The Embedded Firm: On the Socioeconomics of Industrial Networks* (pp. 255-277). London: Routledge.

Grangsjo, Y. v. F. (2006). Hotel networks and social capital in destination marketing. *International Journal of Service Industry Management, 17*(1), 58-75.

Grannis, R. (2004). Sampling the Structure of Large-Scale Social Networks. *Proceedings of the Sixth International Conference on Social Science Methodology: Recent Developments and Applications in Social Research Methodology, Amsterdam, The Netherlands (August 16-20)*

Granovetter, M. (1973). The strength of weak ties. *The American Journal of Sociology, 78*(6), 1360-1380.

Granovetter, M. (1983). The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory, 1*, 201-233.

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology, 91*, 481-510.

Granovetter, M. (2005). The Impact of Social Structure on Economic Outcomes. *Journal of Economic Perspectives, 19*(1), 33-50.

Grönlund, A., & Holme, P. (2005). A network-based threshold model for the spreading of fads in society and markets. *Advances in Complex Systems, 8*, 261-273.

Guardiola, X., Díaz-Guilera, A., Pérez, C. J., Arenas, A., & Llas, M. (2002). Modeling diffusion of innovations in a social network. *Physical Review E, 66*, 026121

Guerrieri, P., & Pietrobelli, C. (2001). Models of Industrial Clusters' Evolution and Changes in Technological Regimes. In P. Guerrieri, S. Iammarino & C. Pietrobelli (Eds.), *The Global Challenge to Industrial Districts: Small and Mediumsized Enterprises in Italy and Taiwan*. Cheltenham, UK: Edward Elgar.

Guimerà, R., Danon, L., Díaz-Guilera, A., Giralt, F., & Arenas, A. (2003). Self-similar community structure in a network of human interactions. *Physical Review E, 68*, 065103(R).

Gunn, C. A. (1997). *Vacationscape, Developing Tourist Areas*. Washington, DC: Taylor & Francis.

Gunn, C. A., & Var, T. (2002). *Tourism planning: basics, concepts and cases* (4th ed.). New York: Routledge.

Hackathorn, R. (2003). The link is the King. *DM Review [On-line], August*. Retrieved December, 2003, from http://www.dmreview.com.

Hagel, J., & Armstrong, A. G. (1997). *Net Gain - Expanding markets through virtual communities*. Cambridge, MA: Harvard Business School Press.

Hall, C. M. (1999). Rethinking collaboration and partnership: A public policy perspective. *Journal of Sustainable Tourism, 7*(3&4), 274-289.

Hall, C. M., & Butler, R. W. (1995). In Search of Common Ground: Reflections on Sustainability: Complexity and Process in the Tourism System: A Discussion. *Journal of Sustainable Tourism, 3*(2), 99-105.

Hall, R. (1993). A framework linking intangible resources and capabilities to sustainable competitive advantage. *Strategic Management Journal, 14*, 607-618.

Han, J. J., Dupuy, D., Bertin, N., Cusick, M. E., & Vidal, M. (2005). Effect of sampling on topology predictions of protein-protein interaction networks. *Nature Biotechnology, 23*(7), 839-844.

Handford, P. (1999). Descifrando el pasado. *Ciencia Hoy, 9*(52), 44-60.

Hannemann, R. H. (2001). *Introduction to Social Network Methods [Online textbook]*. Riverside, CA: University of California. Retrieved April, 2005, from http://faculty.ucr.edu/~hanneman/SOC157/TEXT/TextIndex.html

Hansell, R. I. C., Craine, I. T., & Byers, R. E. (1997). Predicting Change in Non-linear Systems. *Environmental Monitoring and Assessment, 46*, 175-190.

Harrison, B. (1992). Industrial districts: old wine in new bottles? *Regional Studies, 26*(5), 469-483.

Hassink, R., & Shin, D.-H. (2005). Guest Editorial to the thematic issue: The restructuring of old industrial areas in Europe and Asia. *Environment and Planning A, 37*(4), 571-580.

Haythornthwaite, C., & Hagar, C. (2005). The Social Worlds of the Web. *Annual Review of Information Science and Technology, 39*(1), 311-346.

Heath, E., & Wall, G. (1992). *Marketing Tourism Destinations: Strategic Planning Approach*. New York: John Wiley & Sons.

Heckathorn, D. D. (1997). Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations. *Social Problems, 44*(2), 174-199.

Helbing, D., & Molnar, P. (1995). Social force model for pedestrian dynamics. *Physical Review E, 51*, 4282-4286.

Henderson, L. F. (1971). The statistics of crowd fluids. *Nature, 229*, 381-383.

Henrickson, L., & McKelvey, B. (2002). Foundations of ‘‘new’’ social science: Institutional legitimacy from philosophy, complexity science, postmodernism, and agent-based modeling. *Proceedings of the National Academy of the Sciences of the USA, 99* (suppl. 3), 7288-7295.

Hethcote, H. W. (2000). The Mathematics of Infectious Diseases *SIAM Review, 42*(4), 599-653.

Hjalager, A.-M. (1999). *Tourism destinations and the concept of industrial districts* (ERSA conference papers no. ersa99pa288): European Regional Science Association. Retrieved January 2006, from http://ideas.repec.org/p/wiw/wiwrsa/ersa99pa288.html.

Hoang, H., & Antoncic, B. (2003). Network-based research in entrepreneurship: a critical review. *Journal of Business Venturing, 18*(2), 165-187.

Hobbes, T. (1651). *Leviathan, or The Matter, Forme and Power of a Common-Wealth Ecclesiasticall and Civill*. Project Gutenberg, online version (2002).

Holling, C. S. (1986). Resilience of ecosystems; local surprise and global change. In W. C. Clark & R. E. Munn (Eds.), *Sustainable Development of the Biosphere* (pp. 292-317). Cambridge: Cambridge University Press.

Holling, C. S. (Ed.). (1978). *Adaptive Environmental Assessment and Management*. New York: John Wiley and Sons.

Holling, C. S., & Gunderson, L. H. (2002). Resilience and Adaptive Cycles. In L. H. Gunderson & C. S. Holling (Eds.), *Panarchy: Understanding Transformations in Human and Natural Systems* (pp. 25-62). Washington, DC: Island Press.

Hovinen, G. R. (2002). Revisiting the destination lifecycle model. *Annals of Tourism Research, 29*(1), 209-230.

Huberman, B. A., & Adamic, L. A. (2004). Information dynamics in a networked world. In E. Ben-Naim, H. Frauenfelder & Z. Toroczkai (Eds.), *Complex networks. Lecture Notes in Physics, Vol. 650* (pp. 371-398). Berlin: Springer-Verlag.

Hume, D. (1748). *An Enquiry Concerning Human Understanding*. Project Gutenberg, online version (2006).

Humphrey, J., & Schmitz, H. (2001). Governance in Global Value Chains. *IDS Bulletin - Institute of Development Studies, Brighton, UK 32*(3), 19-29.

Humphrey, J., & Schmitz, H. (2002). How does insertion in global value chains affect upgrading in industrial clusters? *Regional Studies, 36*(9), 1017-1027.

Hunt, E. S., & Murray, J. M. (1999). *A history of business in medieval Europe, 1200-1550*. Cambridge: Cambridge University Press.

Huxham, C. (2003). Theorizing Collaboration Practice. *Public Management Review, 5*(3), 401-423.

Ingram, P., & Roberts, P. W. (2000). Friendships among competitors in the Sydney hotel industry. *American Journal of Sociology, 106*(2), 387–423.

Inkpen, A. C., & Tsang, E. W. K. (2005). Social capital, networks, and knowledge transfer. *Academy of Management Review, 30*(1), 146-165.

Iribarren, J. L., & Moro, E. (2007). *Information diffusion epidemics in social networks* (arXiv:physics/0706.0641). Retrieved October 2007, from http://arxiv.org/abs/0706.0641.

ISTAT. (2006). *Annuario Statistico Italiano*. Roma: Istituto Nazionale di Statistica.

Jackson, M. H. (1997). Assessing the structure of communication on the world wide web. *Journal of Computer-Mediated Communication [On-line] 3*(1). Retrieved July, 2006, from http://jcmc.indiana.edu/vol3/issue1/jackson.html.

Jacobs, D., & de Man, A. P. (1996). Clusters, Industrial Policy and Firm Strategy: A Menu Approach. *Technology Analysis & Strategic Management, 8*(4), 425-437.

Jafari, J. (Ed.). (2000). *Encyclopedia of Tourism*. London: Routledge.

Jain, A. K., & Dubes, R. C. (1988). *Algorithms for Clustering Data*. Upper Saddle River, NJ: Prentice-Hall.

Jamal, T. B., & Getz, D. (1995). Collaboration Theory and Community Tourism Planning. *Annals of Tourism Research, 22*(1), 186-204.

Jamal, T. B., Stein, S., & Harper, T. (2002). Beyond Labels: Pragmatic Planning in Multi-Stakeholder Tourism-Environmental Conflicts. *Journal of Planning Education and Research, 22*(2), 164-177.

JDP. (2006). *European Hotel Guest Satisfaction Index Study*. Westlake Village, CA, USA: J.D. Power and Associates.

Jing, F., Wang, C., Yao, Y., Deng, K., Zhang, L., & Ma, W.-Y. (2006). IGroup: a web image search engine with semantic clustering of search results. *Proceedings of the 14th annual ACM international conference on Multimedia, Santa Barbara, CA, USA (October 23-27)*, 497-498.

Johns, N. (1999). What is this thing called service? . *European Journal of Marketing, 33*(9/10), 958-974.

Johnston, R. (2004). Clusters: A Review of their Basis and Development in Australia. *Innovation: Management, Policy & Practice, 6*(3), 380-391.

Jones, Q. (1997). Virtual-Communities, Virtual Settlement & Cyber Archaelogy: A Theoretical Outline. *Journal of Computer Mediated Communication [On-line], 3*(3). Retrieved August 2002, from http://www.ascusc.org/jcmc/vol3/issue3/jones.html.

Kantz, H., & Schreiber, T. (1997). *Nonlinear Time Series Analysis*. Cambridge: Cambridge University Press.

Katz, M. L., & Shapiro, C. (1985). Network Externalities, Competition and Compatibility. *American Economic Review, 75*(3), 424.

Kauffman, S. (1995). *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity*. New York: Oxford University Press.

Keller, P. (1996). Wieso verreisen wir? Metaeconomic Überlegungen. In G. Fischer & C. Laesser (Eds.), *Theorie und Praxis der Tourismus und Verkehrswirtschaft im Wertewandel*. St. Gallen: Paul Haupt.

Kerner, B. S., & Rehborn, H. (1996). Experimental features and characteristics of traffic jams. *Physical Review E, 53*, R1297-R1300.

Killworth, P. D., & Bernard, H. R. (1976). Informant Accuracy in Social Network Data. *Human Organization, 35*, 269-296.

Kirkpatrick, S., Gelatt Jr., C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science, 220*(4598), 671-680.

Kirkpatrick, S., Gelatt Jr., C. D., & Vecchi, M. P. (1984). Optimization by simulated annealing: Quantitative studies. *Journal of Statistical Physics, 34*, 975-986.

Kogut, B., & Walker, G. (2001). The small world of Germany and the durability of national networks. *American Sociological Review, 66*, 317-335.

Koka, B. R., Madhavan, R., & Prescott, J. E. (2006). The Evolution of Interfirm Networks: Environmental Effects on Patterns of Network Change. *The Academy of Management Review, 31*(3), 721-737.

Komulainen, T. (2004). Self-Similarity and Power Laws. In H. Hyötyniemi (Ed.), *Complex Systems - Science on the Edge of Chaos (Report 145, October 2004)*. Helsinki: Helsinki University of Technology, Control Engineering Laboratory.

König, D. (1936). *Theorie der endlichen und unendlichen Graphen*. Leipzig.

Kossinets, G. (2006). Effects of missing data in social networks. *Social Networks, 28*(3), 247-268.

Krieger, M. H. (2005). A 1940 Letter of André Weil on Analogy in Mathematics. *Notices of the American Mathematical Society, 52*(3), 334-341.

Krugman, P. (1991a). *Geography and Trade*. Cambridge, MA: MIT Press.

Krugman, P. (1991b). Increasing Returns and Economic Geography. *Journal of Political Economy, 99*, 483-499.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.

Kumar, R., Raghavan, P., Rajagopalan, S., Sivakumar, D., Tompkins, A. S., & Upfal, E. (2000a). Stochastic models for the Web graph. *Proceedings of the 41st Annual IEEE Symposium on the Foundations of Computer Science (FOCS), Redondo Beach, CA, USA (November 12-14)*, 57-65.

Kumar, R., Raghavan, P., Rajagopalan, S., Sivakumar, D., Tompkins, A. S., & Upfal, E. (2000b). The Web as a graph *Proceedings of the 9th ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems Dallas, TX, USA*, 1-10.

Kuperman, M., & Abramson, G. (2001). Small World Effect in an Epidemiological Model. *Physical Review Letters, 86*, 2909-2912.

Kvale, S. (1996). *Interviews: an introduction to qualitative research interviewing*. Thousand Oaks, Calif.: Sage Publications.

Kwon, Y., Lee, H.-K., Lee, S., & Lee, J. (2003). The Virtual Enterprise: Redefining the Concept. In C.-W. Chung, C.-K. Kim, W. Kim, T. W. Ling & K. H. Song (Eds.), *Lecture Notes in Computer Science* (Vol. 2713, pp. 249-258). Berlin: Springer.

Lachmann, M., Sella, G., & Jablonka, E. (2000). On the advantages of information sharing. *Proceedings of the Royal Society of London B, 267*, 1287-1293.

Ladkin, A., & Bertramini, A. M. (2002). Collaborative Tourism Planning: A Case Study of Cusco, Peru. *Current Issues in Tourism, 5*(2), 71-93.

Lagiewski, R. M. (2005). The Application of the TALC Model: A Literature Survey. In R. W. Butler (Ed.), *The Tourism Area Life Cycle: Applications and Modifications* (Vol. 1, pp. 27-50). Clevedon, UK: Channel View.

Laguna, M. F., Risau Gusman, S., Abramson, G., Gonçalves, S., & Iglesias, J. R. (2005). The dynamics of opinion in hierarchical organizations. *Physica A, 351*, 580-592.

Lakatos, I. (1974). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91-198). Cambridge, UK: Cambridge University Press.

Landau, L. D., & Lifshitz, E. M. (1980). *Statistical Physics - Part 1* (3rd ed.). Oxford: Pergamon Press.

Lane, P. J., Koka, B. R., & Pathak, S. (2006). The Reification of Absorptive Capacity: A Critical Review and Rejuvenation of the Construct. *Academy of Management Review, 31*(4), 833-863.

Lang, S. (1970). *Linear Algebra*. Reading, MA.: Addison-Wesley Pub. Co.

Langley, R. (1971). *Practical statistics*. New York: Dover

Latora, V., & Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical Review Letters, 87*, 198701.

Laudan, L. (1977). *Progress and its Problems*. Berkeley, CA: University of California Press.

Laudan, L. (1981). A Problem-Solving Approach to Scientific Progress. In I. Hacking (Ed.), *Scientific Revolutions* (pp. 144-155). Oxford: Oxford University Press.

Laudan, L. (1984). *Science and Values*. Berkeley, CA: University of California Press.

Laumann, E., Galaskiewicz, J., & Marsden, D. (1978). Community structure as interorganizational linkages. *Annual Review of Sociology, 4*, 455-484.

Laumann, E., Marsden, P. V., & Presnky, D. (1983). The boundary specification problem in network analysis. In R. Burt & M. Minor (Eds.), *Applied Network Analysis: A Methodological Introduction* (Vol. 18-34). London: Sage.

Lawrence, S. (2000). Context in Web Search. *IEEE Data Engineering Bulletin, 23*(3), 25-32.

Lazzeretti, L. (2003). City of art as a High Culture local system and cultural districtualization processes: the cluster of art restoration in Florence. *International Journal of Urban and Regional Research, 27*(3), 635-648.

Lazzeretti, L., & Petrillo, C. S. (Eds.). (2006). *Tourism Local Systems and Networking*. Amsterdam: Elsevier.

Lee, S. H., Kim, P.-J., & Jeong, H. (2006). Statistical properties of sampled networks. *Physical Review E, 73*, 016102

Lehmann, S., & Hansen, L. K. (2007). *Deterministic Modularity Optimization* (arXiv/physics/0701348). Retrieved July 2007, from http://arxiv.org/abs/0701348.

Leibniz, G. W. (1693). De Analysi situs. In C. I. Gerhardt (Ed.), *G.W. Leibniz Matematische Schriften* (Vol. V, pp. 178-183). Hildesheim: OLMS.

Leiper, N. (1979). The framework of tourism: Towards a definition of tourism, tourist, and the tourist industry. *Annals of Tourism Research, 6*(4), 390-407.

Leiper, N. (2000a). Are Destinations 'The Heart of Tourism'? The advantages of an Alternative Description. *Current Issues in Tourism, 3*(4), 364-368.

Leiper, N. (2000b). An emerging discipline. *Annals of Tourism Research, 27*(3), 805-809

Leskovec, J., & Faloutsos, C. (2006). Sampling from large graphs. *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining Philadelphia, PA, USA (August 20-23)*, 631-636.

Levin, S. A. (2003). Complex adaptive systems: Exploring the known, the unknown and the unknowable. *Bulletin of the American Mathematical Society, 40*(1), 3-19.

Lewin , R. (1999). *Complexity, life on the edge of chaos* (2nd ed.). Chicago The University of Chicago Press.

Lewin, R. (1999). *Complexity, life on the edge of chaos* (2nd ed.). Chicago The University of Chicago Press.

Leydesdorff, L. (1991). The static and dynamic analysis of network data using information theory. *Social Networks, 13*, 301-345.

Liebowitz, S. J., & Margolis, S. E. (1994). Network Externality: An Uncommon Tragedy. *Journal of Economic Perspectives, 8*(2), 133-150.

Lincoln, Y. S., & Guba, E. G. (1985). The only generalization is: there is no generalization. In Y. S. Lincoln & E. G. Guba (Eds.), *Naturalistic inquiry* (pp. 110-128). London: Sage.

Liu, Y.-C., Wang, X.-L., Liu, B.-Q., & Zhong, B.-B. (2006). Clustering analysis technology for information retrieval. *Journal of Electronics and Information Technology (Dianzi Yu Xinxi Xuebao), 28*(4), 606-609.

Lofland, J., Snow, D. A., Anderson, L., & Lofland, L. H. (2006). *Analyzing social settings*. Belmont, Calif.: Wadsworth/Thomson.

López-Pintado, D. (2004). *Diffusion in Complex Social Networks* (Working Papers No. AD 2004-33): Instituto Valenciano de Investigaciones Económicas, S.A. (Ivie). Retrieved December 2005, from http://www.ivie.es/downloads/docs/wpasad/wpasad-2004-33.pdf.

Loubser, J. J. (1976). Action and experience. In J. J. Loubser, R. C. Baum, A. Effrat & V. M. Lidz (Eds.), *Explorations in general theory in social science, vol. I. Essays in honor of Talcott Parsons* (pp. 240-263). New York: The Free Press.

Lozano, J., Gómez, C., & Rey-Maquieira, J. (2005). *An Analysis of the Evolution of Tourism Destinations from the Point of View of the Economic Growth Theory* (Nota di lavoro 146.2005). Milan: The Fondazione Eni Enrico Mattei. Retrieved March 2006, from http://www.feem.it/Feem/Pub/Publications/WPapers/default.htm.

Lynham, S. A. (2002). Quantitative Research and Theory Building: Dubin's Method. *Advances in Developing Human Resources, 4*(3), 242-276.

Maggioni, M. A. (2002). *Clustering Dynamics and the Location of High-Tech Firms*. Heidelberg Springer

Maggioni, M. A. (2004). *The rise and fall of industrial clusters: Technology and the life cycle of region* (Document de treball 2004/6). Barcelona, Spain: Institut d'Economia de Barcelona. Retrieved April 2006, from http://www.pcb.ub.es/ieb/serie/doc2004-6.pdf.

Majorana, E. (1942). Il valore delle leggi statistiche nella fisica e nelle scienze sociali. *Scientia 71*, 58-66.

Mandelbrot, B. B. (1963). The variation of certain speculative prices. *The Journal of Business of the University of Chicago, 36*, 394-419.

Manna, S. S., & Kabakçıoğlu, A. (2003). Scale-free Network on Euclidean Space Optimized by Rewiring of Links. *Journal of Physics A, 36*, 279-285.

Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B, 11*, 193-197.

Mantegna, R. N. (2005). Presentation of the English translation of Ettore Majorana's paper: The value of statistical laws in physics and social sciences. *Quantitative Finance, 5*(2), 133-140.

Mantegna, R. N., & Stanley, H. E. (2000). *Introduction to Econophysics: Correlations & Complexity in Finance*. Cambridge: Cambridge University Press.

Marceau, J. (1994). Clusters, Chains and Complexes: Three Approaches to Innovation with a Public Policy Perspective. In M. Dodgson & R. Rothwell (Eds.), *The Handbook of Industrial Innovation*. Aldershot: Edward Elgar.

Marcussen, C. H. (2006). *Trends in European Internet Distribution of Travel and Tourism Services - 7th revision*. Centre for Regional and Tourism Research, Denmark. Retrieved April, 2006, from http://www.crt.dk/uk/staff/chm/trends.htm.

Markusen, A. (1996). Sticky Places in Slippery Space: A Typology of Industrial Districts. *Economic Geography, 72*, 293-313.

Marsden, P. V. (1990). Network Data and Measurement. *Annual Review of Sociology, 16*, 435-463.

Marshall, A. (1920). *Principles of Economics* (8th ed.). London: Macmillan and Co. Retrieved March, 2006, from http://www.econlib.org/LIBRARY/Marshall/marP24.html.

Martin, P., & Ottaviano, G. (2001). Growth and Agglomeration. *International Economic Review, 42*(4), 947-968.

Martin, R., & Sunley, P. (2002). Deconstructing Clusters: Chaotic Concept or policy panacea? *Journal of Economic Geography, 3*(1), 5-35.

Martinez-Fernandez, C. M. (2001). *Networks for Regional Development: Case Studies from Australia and Spain* (PhD Thesis). Sydney: The University of New South Wales Retrieved February 2006, from http://aegis.uws.edu.au/innovationedges/bio_publicat.html.

Matei, S. A., & Ball-Rokeach, S. J. (2001). Real and Virtual Social Ties: Connections in the Everyday Lives of Seven Ethnic Neighborhoods. *American Behavioral Scientist, 45*(3), 550-563.

Mathias, N., & Gopal, V. (2001). Small-worlds: How and why. *Physical Review E, 63*, 1.

MATLAB. (2004). MATLAB (Version 7, R14 - 2004). Natick, MA: The Matworks Inc.

Maxwell, J. C. (1870). Address to the Mathematical and Physical Sections of the British Association for the Advancement of Science. In *Five of Maxwell's Papers* (reprint 2004 ed., pp. 5-21). Whitefish, MT: Kessinger Publishing.

McIntosh, R. W., & Goeldner, C. R. (1990). *Tourism: Principles, Practices, Philosophies* (6th ed.). New York: John Wiley.

McKercher, B. (1999). A Chaos Approach to Tourism. *Tourism Management, 20*, 425-434.

McKercher, B. (2005). Destinations as Products? A Reflection on Butler's Life cycle. *Tourism Recreation Research, 30*(3), 97-102.

McLaren, T., Head, M., & Yuan, Y. (2002). Supply chain collaboration alternatives: understanding the expected costs and benefits. *Internet Research: Electronic Networking Applications and Policy, 12*(4), 348-364.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 415-444.

Merriam-Webster. (2005). *The Merriam-Webster Online Dictionary*. Springfield, MA: Merriam-Webster Inc. Retrieved March, 2006, from http://www.m-w.com/.

Metcalfe, R. M. (1995). Metcalfe's Law. *InfoWorld, October 2.*

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of State Calculation by Fast Computing Machines. *Journal of Chemical Physics, 21*, 1087-1091.

Michael, E. J. (2003). Tourism micro-clusters. *Tourism Economics, 9*(2), 133-145.

Milgram, S. (1967). The small world problem. *Psychology Today, 1*, 60-67.

Mill, R. C., & Morrison, A. M. (1992). *The tourism system: an introductory text*. Englewood Cliffs, NJ: Prentice-Hall.

Minerba, L., Chessa, A., Coppola, R. C., Mula, G., & Cappellini, G. (2007). *A Complex Network Analysis of a Health Organization* (working paper, personal communication).

Mirowski, P. (1989). *More Heat than Light: Economics as Social Physics, Physics as Nature's Economics*. Cambridge, UK: Cambridge University Press.

Mishan, E. J. (1971). The postwar literature on externalities: An interpretative essay. *Journal of Economic Literature, 9*, 1-28.

Mitzenmacher, M. (2004). A Brief History of Generative Models for Power Law and Lognormal Distributions. *Internet Mathematics, 1*(2), 226-251.

Moghaddam, A. I. (2007). Web metasearch engines: A comparative study on search capabilities using an evaluation check-list. *Online Information Review, 31*(3), 300-309.

Moody, J., McFarland, D., & Bender-deMoll, S. (2005). Dynamic Network Visualization. *The American Journal of Sociology, 110*(4), 1206-1241.

Moreno, J. L. (1934). *Who shall survive?* Washington, DC: Nervous and Mental Disorders Publishing Co.

Morrison, A. (2004). *"Gatekeepers of knowledge" within industrial districts: who they are, how they interact.* Paper presented at the The 4th Congress on Proximity Economics: Proximity, Networks and Co-ordination, Marseille, France, Retrieved June 17 - 18, 2004,

Morrison, A., Lynch, P., & Johns, N. (2004). International tourism networks. *International Journal of Contemporary Hospitality Management, 16*(3), 197-202.

Mossa, S., Barthélémy, M., Stanley, H. E., & Amaral, L. A. N. (2002). Truncation of Power Law Behavior in "Scale-Free" Network Models due to Information Filtering. *Physical Review Letters, 88*(13), 138701.

Motter, A. E., & Lai, Y.-C. (2002). Cascade-based attacks on complex networks. *Physical Review E, 66*, 065102(R).

Nagel, E. (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt, Brace and World.

Nagel, K., & Schreckenberg, M. (1992). A cellular automaton model for freeway traffic. *Journal of Physics I, 2*, 2221-2230.

Nakano, T. (2004). *Industries in a Large-Scale Industrial District: A Component Analysis* (Corporate Strategy Working paper No: 2004-03). Flint, MI: University of Michigan-Flint, School of Management. Retrieved April 2006, from http://som.umflint.edu/yener/2004-03.pdf.

Newman, M. E. J. (2002a). Assortative mixing in networks. *Physical Review Letters, 89*(20), 208701.

Newman, M. E. J. (2002b). Mixing patterns in networks. *Physical Review E, 67*, 026126.

Newman, M. E. J. (2003a). Random graphs as models of networks. In S. Bornholdt & H. G. Schuster (Eds.), *Handbook of Graphs and Networks* (pp. 35–68). Berlin: Wiley-VCH.

Newman, M. E. J. (2003b). The structure and function of complex networks. *SIAM Review, 45*, 167-256.

Newman, M. E. J. (2004). Detecting community structure in networks. *The European Physical Journal B, 38*, 321-330.

Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics, 46*(5), 323-351.

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E, 69*, 26113.

Newman, M. E. J., Strogatz, S. H., & Watts, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Physical Review E, 64*, 026118.

Nordin, S. (2003). *Tourism clustering and innovations* (Scientific Report No. 2003:14). Ostersund, Sweden: ETOUR - The European Tourism Research Institute. Retrieved November 2006, from http://www.etour.se/download/18.54c4adf7b427c3137fff155/ETOURU2003_14.pdf

Norton, R. D. (1992). Agglomeration and Competitiveness : From Marshall to Chinitz. *Urban Studies, 29*(2), 155-170.

Novelli, M., Schmitz, B., & Spencer, T. (2006). Networks, clusters and innovation in tourism: A UK experience. *Tourism Management, 27*(6), 1141-1152.

OECD. (1996). *Networks of Enterprises and Local Development, Territorial Development*. Paris: OECD.

OECD. (2000a). *Guidelines on Tourism Satellite Accounts*. Paris: OECD.

OECD. (2000b). *Measuring the role of tourism in OECD economies*. Paris: OECD.

OECD. (2001). *Innovative clusters: Drivers of National Innovation Systems*. Paris: OECD.

OECD. (2004). *Glossary of Statistical Terms*. Paris: OECD.

Olsen, W. (2004). Triangulation in Social Research, qualitative and quantitative methods can really be mixed. In M. Holborn (Ed.), *Developments in Sociology: An Annual Review*. Ormskirk, UK: Causeway Press.

Olson, H. F. (1943). *Dynamical Analogies*. New York: Van Nostrand.

Ottino, J. M. (2004). Engineering complex systems. *Nature, 427*, 399.

Park, H. W. (2003). Hyperlink Network Analysis: A New Method for the Study of Social Structure on the Web. *Connections 25*(1), 49-61.

Park, H. W., & Thelwall, M. (2003). Hyperlink Analyses of the World Wide Web: A Review. *Journal of Computer Mediated Communication [On-line], 8*(4). Retrieved March 2006, from http://jcmc.indiana.edu/vol8/issue4/park.html.

Parkhe, A., Wasserman, S., & Ralston, D. (2006). New frontiers in network theory development. *The Academy of Management Review, 31*(3), 560-568.

Pask, C. (2003). Mathematics and the science of analogies. *American Journal of Physics, 71*(6), 526-534.

Pastor-Satorras, R., & Vespignani, A. (2001). Epidemic spreading in scalefree networks. *Physical Review Letters, 86*(14), 3200-3203.

Pastor-Satorras, R., & Vespignani, A. (2002). Immunization of complex networks. *Physical Review E, 65*, 036104

Pastor-Satorras, R., & Vespignani, A. (2003). Epidemics and immunization in scale-free networks. In S. Bornholdt & H. G. Schuster (Eds.), *Handbook of Graphs and Networks*. Berlin: Wiley-VCH.

Pastor-Satorras, R., & Vespignani, A. (2004). *Evolution and structure of the Internet - A Statistical Physics Approach*. Cambridge, UK: Cambridge University Press.

Pavard, B., & Dugdale, J. (2000). *The contribution of complexity theory to the study of socio-technical cooperative systems*. Third International Conference on Complex Systems, Nashua, NH, May 21-26. Retrieved October, 2005, from http://www-svcict.fr/cotcos/pjs/.

Pavlovich, K. (2001). The Twin Landscapes of Waitomo: Tourism Network and Sustainability through the Landcare Group. *Journal of Sustainable Tourism, 9*(6), 491-504.

Pavlovich, K. (2003a). The evolution and transformation of a tourism destination network: The Waitomo Caves, New Zealand. *Tourism Management, 24*(2), 203-216.

Pavlovich, K. (2003b). Pyramids, Pubs, and Pizzas: An Interpretation of Tourism Network Structures. *Tourism Culture & Communication, 4*(1), 41-48.

Pavlovich, K., & Kearins, K. (2004). Structural Embeddedness and Community-Building through Collaborative Network Relationships. *M@n@gement, 7*(3), 195-214.

Pearce, D. G. (1989). *Tourist Development* (2nd ed.). Harlow, UK: Longman.

Pearce, D. G. (1992). *Tourist Organizations*. Harlow, UK: Longman.

Pechlaner, H., Abfalter, D., & Raich, F. (2002). Cross-Border Destination Management Systems in the Alpine Region-The Role of Knowledge Networks on the Example of AlpNet. *Journal of Quality Assurance in Hospitality & Tourism, 3*(3/4), 89-107.

Pechlaner, H., Tallinucci, V., Abfalter, D., & Rienzner, H. (2003). Networking for Small Island Destinations – The Case of Elba. In A. J. Frew, M. Hitz & P. O'Connor (Eds.), *Information and Communication Technologies in Tourism* (pp. 105-114). Wien: Springer.

Pennock, D. M., Flake, G. W., Lawrence, S., Glover, E. J., & Giles, C. L. (2002). Winners don't take all: Characterizing the competition for links on the web. *Proceedings of the National Academy of the Sciences of the USA, 99*(8), 5207-5211.

Pesämaa, O., & Skurla, J. A. (2003). *Secret Ties as a Way to Succeed: Describing and Exploring Relations among Successful Tourism Managers.* Paper presented at the Perspectives on Tourism in Nordic and Other Peripheral Areas International Conference, Umeå, Sweden, Retrieved 21-24 August 2003, from http://www.hfon.org/fon/engelska/Publikationer,%20rapporter-eng.htm.

Pforr, C. (2001). Tourism Policy in Australia's Northern Territory: A Policy Process Analysis of its Tourism Development Masterplan. *Current Issues in Tourism, 4*(2), 275-307.

Pforr, C. (2002). The 'makers and shapers' of tourism policy in the Northern Territory of Australia: A policy network analysis of actors and their relational constellations. *Journal of Hospitality and Tourism Research, 9*(2), 134-151.

Pforr, C. (2006). Tourism Policy in the Making: An Australian Network Study. *Annals of Tourism Research, 33*(1), 87-108.

PhoCusWright. (2005). *Online Travel Overview: Market Size and Forecasts 2004-2006*. Sherman, CT: PhoCusWright.

Poincaré, H. (1883). Sur certaines solutions particulieres du probleme des trois corps *Comptes-Rendus de l'Académie des Sciences, Paris, 97*(251-252).

Poincaré, H. (1884). Sur certaines solutions particulieres du probleme des trois corps *Bulletin Astronomique, 1*(63-74).

Pontes, J. P. (2000). *A game theoretic-model of multinational firm location* (Working Papers 2000/02): Department of Economics, Institute for Economics and Business Administration (ISEG), Technical University of Lisbon.

Poon, A. (1993). *Tourism, technology and competitive strategies*. Oxon, UK: CAB International.

Popper, K. (1959). *The Logic of Scientific Discovery*. London: Hutchinson.

Porter, M. E. (1990). *The Competitive Advantage of Nations*. London: MacMillan & Co.

Porter, M. E. (1998). Clusters and the New Economics of Competition. *Harvard Business Review, 76* (6), 77-90.

Pothen, A., Simon, H. D., & Liou, K.-P. (1990). Partitioning sparse matrices with eigenvectors of graphs. *SIAM Journal On Matrix Analysis and Applications, 11*(3), 430-452.

Priestley, J. L., & Samaddar, S. (2007). Multi-Organizational Networks: Three Antecedents of Knowledge Transfer. *International Journal of Knowledge Management, 3*(1), 86-99.

Prosser, G. M. (1995). *Tourist Destination Life Cycles: Progress, Problems and Prospects.* Paper presented at the CAUTHE National Tourism Research Conference, Melbourne, VIC,

Provan, K. G., Fish, A., & Sydow, J. (2007). Interorganizational Networks at the Network Level: A Review of the Empirical Literature on Whole Networks. *Journal of Management, 33*(3), 479-516.

Provan, K. G., & Kenis, P. (2007). Modes of Network Governance: Structure, Management, and Effectiveness. *Journal of Public Administration Research and Theory*, doi: 10.1093/jopart/mum1015.

Provan, K. G., Veazie, M. A., Staten, L. K., & Teufel-Shone, N. I. (2005). The Use of Network Analysis to Strengthen Community Partnerships. *Public Administration Review, 65*(5), 603-613.

Punch, K. F. (2005). *Introduction to Social Research - Quantitative & Qualitative Approaches*. London: Sage.

Pyke, F., & Sengenberger, W. (Eds.). (1992). *Industrial Districts and Local Economic Regeneration*. Geneva: International Institute for Labour Studies.

Quévit, M., & van Doren, P. (1997). *The Problem of Innovative Milieux and Territorial Innovative Regions. The GREMI Approach*. Aldershot, UK: Ashgate.

Radcliffe-Brown, A. R. (1940). On Social Structure. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland, 70*(1), 1-12.

Radicchi, F., Castellano, C., Cecconi, F., Loreto, V., & Parisi, D. (2004). Defining and identifying communities in networks. *Proceedings of the National Academy of the Sciences of the USA, 101* (9), 2658-2663.

Rafiei, D., & Curial, S. (2005). Effectively Visualizing Large Networks Through Sampling. *Proceedings of the IEEE Visualization Conference (VIS'05), Minneapolis (October)*, 48-56.

Ravasz, E., & Barabási, A.-L. (2003). Hierarchical organization in complex networks. *Physical Review E, 67*, 026112.

Rayport, J. F., & Sviokla, J. J. (1995). Exploiting the Virtual Value Chain. *Harvard Business Review, 73*(6), 75-86.

Reagans, R., & McEvily, B. (2003). Network structure and knowledge transfer: The effects of cohesion and range. *Administrative Science Quarterly, 48*, 240-267.

Reynolds, C. (1987). Flocks, herds, and schools: a distributed behavioral model. *Computer Graphics, 21*, 25-34.

Rheingold, H. (1993). *The Virtual Community: Homesteading on the Electronic Frontier*. Reading, MA: Addison-Wesley.

Rhodes, R. A. W. (2002). Putting People Back into Networks. *Australian Journal of Political Science, 37*(3), 399-416.

Richardson, G. (2005). Medieval Guilds. *EH.Net Encyclopedia of Economic and Business History (edited by R. Whaples)*. Retrieved March, 2006, from http://eh.net/encyclopedia/article/richardson.guilds.

Ritchie, J. R. B., & Crouch, G. I. (2003). *The Competitive Destination: A Sustainable Tourism Perspective*. Oxon, UK: CABI Publishing.

Ritchie, R. J. B., & Ritchie, J. R. B. (2002). A framework for an industry supported destination marketing information system. *Tourism Management, 23* 439-454.

Ritter, T., & Gemünden, H. G. (2003). Interorganizational relationships and networks: An overview. *Journal of Business Research, 56* 691-697.

Ritter, T., Wilkinson, I. F., & Johnston, W. J. (2004). Managing in complex business networks. *Industrial Marketing Management, 33* 175-183.

Roberts, L., & Simpson, F. (1999). Developing partnership approaches to tourism in Central and Eastern Europe. *Journal of Sustainable Tourism, 7*(3&4), 314-330.

Rogers, E. M., Medina, U. E., A., R. M., & Wiley, C. J. (2005). Complex Adaptive Systems and the Diffusion of Innovations. *The Innovation Journal 10*(3), 30.

Romano, A., Passiante, G., & Elia, V. (2001). New sources of clustering in the digital economy. *Journal of Small Business and Enterprise Development, 8*(1), 19-27.

Rosenhead, J. (1998). *Complexity Theory and Management Practice. Science as Culture*. Retrieved December 2005, from http://www.human-nature.com/science-as-culture/rosenhead.html.

Rothenberg, R. B. (1995). Sampling in Social Networks. *Connections 18*(1), 104-110.

Russell, R. (2006). Chaos theory and managerial approaches. In D. Buhalis & C. Costa (Eds.), *Tourism Dynamics, Challenges and Tools: Present and Future Issues* (pp. 108-115). Oxford Butterworth-Heinemann.

Russell, R., & Faulkner, B. (2004). Entrepreneurship, Chaos and the Tourism Area Lifecycle. *Annals of Tourism Research, 31*(3), 556-579.

Ryan, C. (1991). Tourism and marketing - A symbiotic relationship. *Tourism Management, 12*(2), 101-111.

Saari, D. G. (1995). Mathematical Complexity of Simple Economics. *Notices of the American Mathematical Society, 42*(2), 222-230.

Sääskilahti, P. (2005). *Essays on the Economics of Networks and Social Relations* (Acta Universitatis Oeconomicae Helsingiensis A-263). Helsinki: Helsinki School of Economics.

Salganik, M. J., & Heckathorn, D. D. (2004). Sampling and Estimation in Hidden Populations Using Respondent-Driven Sampling. *Sociological Methodology, 34*, 193-239.

Saxenian, A. (1994). *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press.

Schelling, T. C. (1971). Dynamic Models of Segregation. *Journal of Mathematical Sociology, 1*(2), 143-186.

Schmitz, H. (1992). On the Clustering of Small Firms. *IDS Bulletin - Institute of Development Studies, Brighton, UK 23*(3), 64-69.

Schmitz, H. (1998). Collective Efficiency and Increasing Returns. *Cambridge Journal of Economics, 23* (4), 465-483.

Schmitz, H., & Nadvi, K. (1999). Clustering and industrialization: Introduction. *World Development, 27*(9), 1503-1514.

Schucan, C. (1998). The role of an information concept in relation to destination management. In D. Buhalis, A. M. Tjoa & J. Jafari (Eds.), *Information and Communication Technologies in Tourism 1998: Proceedings of the International Conference in Istanbul, Turkey* (pp. 84-94). Wien: Springer.

Scott, A. (1995). The Geographic Foundations of Industrial Performance. *Competition and Change, 1*, 51-66.

Scott, A. (1998). *New Industrial Spaces: Flexible Production Organisations and Regional Development in North America and Western Europe*. London: Pion.

Scott, J. (1988). Social Network Analysis. *Sociology, 22*(1), 109-127.

Scott, J. (2000). *Social network analysis: a handbook* (2nd ed.). London: Sage Publications.

Scott, N., Cooper, C., & Baggio, R. (2008a). Destination Networks - Theory and practice in four Australian cases. *Annals of Tourism Research, 35*(1), 169-188.

Scott, N., Cooper, C., & Baggio, R. (2008b). *Network Analysis and Tourism: From Theory to Practice*. Clevedon, UK: Channel View.

Seary, A. J., & Richards, W. D. (2003). Spectral methods for analyzing and visualizing networks: an introduction. In R. Breiger, K. Carley & P. Pattison (Eds.), *Dynamic Social Network Modeling and Analysis* (pp. 209-228). Washington, DC: National Academy Press.

Sforzi, F. (2000). Local development in the experience of Italian industrial districts. In B. Cori, G. Corna Pellegrini, G. Cortesi, G. Dematteis, M. T. Di Maggio Alleruzzo, C. Minca, F. Sforzi & G. Zanetto (Eds.), *Geographies of Diversity. Italian Perspectives* (pp. 133-156). Roma: Società Geografica Italiana.

Sforzi, F., & Lorenzini, F. (2002). I distretti industriali. In IPI (Ed.), *L'esperienza italiana dei distretti industriali* (pp. 20-33). Roma: Ministero delle Attività Produttive.

Shao, J. (1999). *Mathematical Statistics*. Heidelberg: Springer.

Sheehan, L., & Ritchie, J. R. B. (2005). Destination stakeholders: Exploring Identity and Salience. *Annals of Tourism Research, 32*(3), 711–734.

Sheskin, D. J. (2000). *Handbook of Parametric and Nonparametric Statistical Procedures*. Boca Raton, FL: Chapman&Hall/CRC.

Shih, H.-Y. (2006). Network characteristics of drive tourism destinations: An application of network analysis in tourism. *Tourism Management, 27*(5), 1029-1039.

Shulman, L. S. (1999). The core, the canon, and the development of research skills. In L. S. Shulman & E. C. Lagemann (Eds.), *Issues in education research: Problems and possibilities* (pp. 166-223). San Francisco, CA: Jossey-Bass Publishers.

Shy, O. (2001). *The Economics of Network Industries*. Cambridge Cambridge University Press.

Siegel, S., & Castellan, N. J. J. (1988). *Nonparametric Statistics for the Behavioral Sciences* (2nd ed.). New York: McGraw-Hill.

Simon, H. A. (1955a). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics, 69*(1), 99-118.

Simon, H. A. (1955b). On a class of skew distribution functions. *Biometrika, 42*, 425-440.

Simon, H. A. (1997). *Models of Bounded Rationality: Empirically Grounded Economic Reason*. Cambridge, MA: MIT Press.

Sinclair, M. T., & Stabler, M. (1997). *The Economics of Tourism*. London: Routledge.

Skopal, T., Snášel, V., Svátek, V., & Krátký, M. (2003). Searching the Internet Using Topological Analysis of Web Pages. *Proceedings of the 2003 International Conference on Communications in Computing (CIC'03), Las Vegas, NV*, 271-277.

Smeral, E. (1998). The Impact of globalization on small and medium enterprises: New challenges for tourism policies in European countries. *Tourism Management, 19*(4), 371-380.

Smith, D. M. D., & Johnson, N. F. (2004). *Evolution Management in a Complex Adaptive System: Engineering the Future* (arXiv/cond-mat/0409036). Retrieved January 2006, from http://arxiv.org/abs/cond-mat/0409036.

Smith, S. L. J. (1988). Defining Tourism, A Supply-Side View. *Annals of Tourism Research, 15*, 179-190.

Smith, V. A., Jarvis, E. D., & Hartemink, A. J. (2003). Influence of Network Topology and Data Collection on Network Inference. *Proceedings of the Pacific Symposium on Biocomputing 2003, Kauai, Hawaii, USA (January 3-7)*, 164-175.

Sobkowiz, P. (2003). *Opinion formation in networked societies with strong leaders* (arXiv/cond-mat/0311521). Retrieved January 2006, from http://arxiv.org/abs/cond-mat/0311521.

Solé, R. V., Ferrer i Cancho, R., Valverde, S., & Montoya, J. M. (2000). Selection, Tinkering, and Emergence in Complex Networks. *Complexity, 8*(1), 20-33.

Sorenson, O. (2003). Social networks and industrial geography. *Journal of Evolutionary Economics, 13*, 513-527.

Sorenson, O., Rivkin, J., & Fleming, L. (2005). *Complexity, networks and knowledge flow.* Paper presented at the DRUID Tenth Anniversary Summer Conference - Dynamics of Industry and Innovation: Organizations, Networks and Systems Copenhagen, Denmark, Retrieved June 27-29,

Sornette, D. (2003). *Why Stock Markets Crash*. Princeton: Princeton University Press.

SPSS. (2004). SPSS (Statistical Package for the Social Sciences) (Version 13). Chicago, IL: SPSS Inc. (www.spss.com).

SSSR. (2006). *L'informatizzazione delle strutture ricettive (aprile 2006)* (Supplemento n. 4, Informazioni Statistiche - Mensile della Giunta Regionale Toscana). Retrieved March 2007, from http://www.regione.toscana.it.

Staber, U. (2001). The Structure of Networks in Industrial Districts. *International Journal of Urban and Regional Research, 25*, 537-552.

Stacey, R. D. (1993). *Strategic Management and Organisational Dynamics*. London: Pitman.

Stacey, R. D. (1996). *Complexity and Creativity in Organizations*. San Francisco: Berrett-Koehler.

Stake, R. (1995). *The art of case research*. London: Sage.

Stamboulis, Y., & Skayannis, P. (2003). Innovation strategies and technology for experience-based tourism. *Tourism Management, 25*, 35-43.

Stauffer, D., & Aharony, A. (1992). *Introduction to Percolation Theory* (2nd ed.). London: Taylor & Francis.

Štefančić, H., & Zlatić, V. (2005). Preferential attachment with information filtering - node degree probability distribution properties. *Physica A, 350*, 657-670.

Stopford, J., & Strange, S. (1991). *Rival states, rival firms, competition for World market shares*. Cambridge Cambridge University Press.

Storper, M. (1995). The resurgence of regional economies ten years later: the region as a nexus of untraded interdependencies. *European Urban and Regional Studies, 2*(3), 191-221.

Storper, M. (1997). *The Regional World, territorial development in a global economy*. New York: The Guilford Press.

Stumpf, M. P. H., Ingram, P. J., Nouvel, I., & Wiuf, C. (2005a). Statistical model selection methods applied to biological networks. *Transactions in Computational Systems Biology, 3*, 65-77.

Stumpf, M. P. H., & Wiuf, C. (2005). Sampling properties of random graphs: The degree distribution. *Physical Review E, 72*, 036118

Stumpf, M. P. H., Wiuf, C., & May, R. M. (2005b). Subnets of scale-free networks are not scale-free: Sampling properties of networks. *Proceedings of the National Academy of the Sciences of the USA, 102* (12), 4221-4224.

Sudman, S., & Kalton, G. (1986). New Developments in the Sampling of Special Populations. *Annual Review of Sociology, 12*, 401-429.

Sznajd-Weron, K., & Sznajd, J. (2000). Opinion Evolution in Closed Community. *International Journal of Modern Physics C, 11*(6), 1157-1165.

Tallinucci, V., & Testa, M. (2006). *Marketing per le isole*. Milano Franco Angeli.

Tellis, W. (1997a). Application of a Case Study Methodology. *The Qualitative Report, 3*(3). Retrieved April, 2005, from http://www.nova.edu/ssss/QR/QR3-3/tellis2.html.

Tellis, W. (1997b). Introduction to Case Study. *The Qualitative Report, 3*(2). Retrieved April, 2005, from http://www.nova.edu/ssss/QR/QR3-2/tellis1.html

Thelwall, M. (2006). Interpreting Social Science Link Analysis Research: a Theoretical Framework. *Journal of the American Society for Information Science and Technology, 57*(1), 60-68.

Tichy, N. M., Tushman, M. L., & Fombrum, C. (1979a). Social network analysis for organizations. *Academy of Management Review, 4*(4), 507-519.

Tichy, N. M., Tushman, M. L., & Fombrun, C. (1979b). Social Network Analysis for Organizations. *The Academy of Management Review, 4*(4), 507-519.

Timothy, D. J. (1998). Co-operative tourism planning in a developing destination. *Journal of Sustainable Tourism, 6*(1), 52-68.

Torraco, R. J. (2002). Research Methods for Theory Building in Applied Disciplines: A Comparative Analysis *Advances in Developing Human Resources, 4*, 355-376.

Tremblay, P. (1998). The economic organization of tourism. *Annals of Tourism Research 25*(4), 837-859.

Tremblay, P. (1999). An empirical investigation of tourism business relationships in Australia's Top End. *Proceedings of the Tourism & Hospitality: Delighting the Senses, Canberra, Australia*, 194-206.

Tribe, J. (1997). The Indiscipline of Tourism. *Annals of Tourism Research, 24*(3), 638-657.

Tribe, J. (2005). New Tourism Research. *Tourism Recreation Research, 30*(2), 5-8.

Tsai, W. P. (2001). Knowledge transfer in intraorganizational networks: Effects of network position and absorptive capacity on business-unit innovation and performance. *Academy of Management Journal, 44*, 996-1004.

Turner, J. (1955). Maxwell on the Method of Physical Analogy. *The British Journal for the Philosphy of Science, 6*(23), 226-238.

Tyler, J. R., Wilkinson, D. M., & Huberman, B. A. (2005). E-Mail as Spectroscopy: Automated Discovery of Community Structure within Organizations. *The Information Society, 21*(2), 143-153

UFI. (2005). *The Trade Fair sector: background and fundamentals* UFI, The Global Association of the Exhibition Industry Retrieved December, 2005, from http://www.ufinet.org/pages/thetradefairsector/basicknowledge.asp.

UIA. (2006). *Integrative Concepts Online Database*. Brussels, BE: Union of International Associations. Retrieved March, 2006, from http://www.uia.org/db/.

UNCTAD. (1998). *Promoting and Sustaining SMEs Clusters and Networks for Development*. Geneva: United Nations (UN) Conference on Trade and Development.

UNWTO. (1995). *Concepts, Definitions, and Classifications for Tourism Statistics. Technical Manual No. 1*. Madrid: World Tourism Organization.

UNWTO. (1999). *Tourism Satellite Account (TSA): The Conceptual Framework*. Madrid: World Tourism Organization.

UNWTO. (2000). *Basic References on Tourism Statistics*. Madrid: World Tourism Organization.

UNWTO. (2002a). *News release: WTO Think Tank enthusiastically reaches consensus on frameworks for tourism destination success*. Retrieved November, 2005, from http://www.world-tourism.org/newsroom/Releases/more_releases/november2002/thinktank.htm.

UNWTO. (2002b). *Tourism proves as a resilient and stable economic sector*. Retrieved July, 2004, from http://www.world-tourism.org/newsroom/Releases/more_releases/june2002/data.htm.

UQ HUPP. (2002). *The Handbook of University Policies and Procedures*. Retrieved April, 2006, from http://www.uq.edu.au/hupp/index.html.

Urry, J. (2002). *The Tourist Gaze* (2nd ed.). London: Sage.

Uzzi, B., & Spiro, J. (2005). Collaboration and Creativity: The Small World Problem. *American Journal of Sociology, 111*(2), 447-504.

van der Linde, C. (2003). The Demography of Clusters - Findings from the Cluster Meta-Study. In J. Bröcker, D. Dohse & R. Soltwedel (Eds.), *Innovation Clusters and Interregional Competition* (pp. 130-149). Berlin: Springer.

van Gigch, J. P. (2002a). Comparing the epistemologies of scientific disciplines in two distinct domains: modern physics versus social sciences. I: The epistemology and knowledge characteristics of the physical sciences. *Systems Research and Behavioral Science, 19*(3), 199-209.

van Gigch, J. P. (2002b). Comparing the epistemologies of scientific disciplines in two distinct domains: modern physics versus social sciences. II: Epistemology and knowledge characteristics of the 'new' social sciences. *Systems Research and Behavioral Science, 19*(6), 551-562.

Vanhove, N. (2005). *Economics of Tourism Destinations*. London: Elsevier Butterworth-Heinemann.

Vaughan, L., Gao, Y., & Kipp, M. (2006). Why are hyperlinks to business Websites created? A content analysis. *Scientometrics, 67*(2), 291-300.

Veal, A. J. (2006). *Research Methods for Leisure and Tourism: a Practical Guide* (3rd ed.). Harlow, UK: Financial Times - Prentice Hall /Pearson Education.

Vessuri, H. (2002). Ethical Challenges for the Social Sciences on the Threshold of the 21st Century. *Current Sociology, 50*(1), 135-150.

Viganò, G. (2004). La programmazione della destinazione enogastronomica. In M. Antonioli Corigliano & G. Viganò (Eds.), *Turisti per gusto. Enogastronomia, territorio, sostenibilità* (pp. 59-90). Novara: De Agostini.

Vogt, C. A., Fesenmaier, D. R., & MacKay, K. (1993). Functional and aesthetic information needs underlying the pleasure travel experience. *Journal of Travel and Tourism Marketing, 2*(2), 133-146.

von Friedrichs Grängsjö, Y. (2003). Destination networking. Co-opetition in peripheral surroundings. *International Journal of Physical Distribution & Logistics Management, 33*(5), 427-448.

Wahab, S., & Cooper, C. (Eds.). (2001). *Tourism in the age of globalisation*. New York: Routledge.

Waldrop, M. (1992). *Complexity: The Emerging Science and the Edge of Order and Chaos*. London: Simon and Schuster.

Walker, J. (2002). Links and power: the political economy of linking on the Web. *Proceedings of the 2002 ACM Hypertext Conference, Baltimore, MD*, 72-73.

Walker, P. A., Greiner, R., McDonald, D., & Lyne, V. (1998). The Tourism Futures Simulator: a systems thinking approach. *Environmental Modelling and Software, 14*(1), 59-67.

Wasserman, S., & Faust, K. (1994). *Social Network Analysis. Methods and Applications*. Cambridge, MA: Cambridge University Press.

Watts, D. J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of the Sciences of the USA, 99*(9), 5766–5771.

Watts, D. J. (2003). *Six Degrees: The Science of a Connected Age*. New York: Norton.

Watts, D. J. (2004). The "New" Science of Networks. *Annual Review of Sociology, 30*, 243-270.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small world' networks. *Nature, 393*, 440-442.

Weinberg, S. (1994). *Dreams of a Final Theory: The Scientist's Search for the Ultimate Laws of Nature*. New York: Vintage/Random House.

Wellman, B. (1988). Structural Analysis: From method and methaphor to theory and substance. In B. Wellman & S. D. Berkowitz (Eds.), *Social structures: A network approach* (pp. 19-61). Cambridge: Cambridge University Press.

Wellman, B. (2001). Computer Networks as Social Networks. *Science 293*, 2031-2034.

Wellman, B. (2002). Structural analysis: From metaphor to theory and substance. In J. Scott (Ed.), *Social Networks: Critical Concepts in Sociology* (Vol. 1, pp. 81-122). London: Routledge.

Werthner, H., & Klein, S. (1999). *Information technology and tourism - a challenging relationship*. Wien: Springer.

Whitney, D. E., & Alderson, D. (2006). *Are technological and social networks really different?* Paper presented at the Sixth International Conference on Complex Systems (ICCS2006), Boston, MA, USA, June, 25-30, from http://necsi.org/events/iccs6/viewpaper.php?id=268.

Wigner, E. P. (1955). Characteristic vectors of bordered matrices with infinite dimensions. *Annals of Mathematics, 62*, 548-564.

Wigner, E. P. (1958). On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics, 67*, 325-328.

Wigner, E. P. (1960). The Unreasonable Effectiveness of Mathematics in the Natural Sciences. *Communications in Pure and Applied Mathematics, 13*(1), 1-14.

Wilensky, U. (1999). NetLogo. http://ccl.northwestern.edu/netlogo. Evanston, IL: Center for Connected Learning and Computer-Based Modeling. Northwestern University.

Williamson, O. E. (1981). The Economics of Organization: The Transaction Cost Approach. *The American Journal of Sociology, 87*(3), 548-577.

Wolfram, S. (2002). *A New Kind of Science*. Champaign, IL: Wolfram Media.

WTTC. (2007). *The 2007 Travel & Tourism Economic Research*. London: World Travel & Tourism Council.

Wu, F., Huberman, B. A., Adamic, L. A., & Tyler, J. R. (2004). Information flow in social groups. *Physica A, 337*, 327-335.

Xu, X.-J., Wu, Z.-X., & Chen, G. (2007). Epidemic spreading in lattice-embedded scale-free networks. *Physica A, 377*(1), 125-130.

Yin, R. K. (1994). *Case study research, design and methods* (2nd ed.). London: Sage.

Zanette, D. H. (2002). Dynamics of rumor propagation on small-world networks. *Physical Review A, 65*(4), 041908.

Appendix 1: Main definitions

This appendix contains, for the convenience of the reader, the definitions of the most important concepts used in this thesis along with the references to the sections in which these concepts are described.

Case study (section 5.3)

> "An intensive study of a single unit for the purpose of understanding a larger class of (similar) units" (Gerring, 1994).

Complex system (sections 2.1 and 2.2)

> A system of many parts which are coupled in a nonlinear fashion. Natural complex systems are described using techniques of the mathematical and physical sciences such as: nonlinear dynamics, statistical mechanics and network theory. Because of their intrinsic nonlinearity, complex systems are more than the sum of their parts because a linear system is subject to the principle of superposition, and hence is literally the sum of its parts, while a nonlinear system is not. A nonlinear connection means that change on one side is not proportional to change on the other. When there are many nonlinearities in a system (many components), behaviour can be as unpredictable as it is interesting. Complex systems research studies such behaviour.

Complex adaptive system (section 2.1)

> A complex system in which the interactions among the elements are of a dynamic nature. They are influenced by and influence the external environment. In this type of system, the parts (Stacey, 1996: 10): "interact with each other according to sets of rules that require them to examine and respond to each other's behaviour in order to improve their behaviour and thus the behaviour of the system they comprise". Complex adaptive systems have the capacity to learn from experience and change.

Graph (section 2.3)

A diagram consisting of a number of dots (nodes or vertices) and a number of lines (arcs or edges) connecting certain pairs of dots. Formally, a *graph G* is an ordered pair of disjointed sets (V, E), where $V = \{v_1, \ldots v_n\}$ is the set of vertices and $E = \{(v_1,u_1), \ldots, (v_i,u_j)\}$ is the set of arcs. E is a subset of the Cartesian product $V \times V$. In other words, E is a binary relation on V and its elements are pairs of elements belonging to V. A graph comprising n nodes is also represented by an adjacency matrix A: a square $n \times n$ matrix whose elements a_{ij} are null if no link exists between node i and node j, and have a value representing the weight of the link otherwise.

Industrial district (section 4.1)

A spatially contiguous group of interconnected companies, linked by commonalities and complementarities, involving a localised support infrastructure, with a collective vision for economic and social growth and based on competition and cooperation.

Network (section 2.4)

A complex, interconnected group of elements. Basic elements are:

- nodes: objects of interests (e.g. people, companies, computers, brain cells etc.). Synonyms frequently used are: vertex, actor;
- links: relationships between nodes (e.g. friendship, communication, information exchange, chemical reaction, transaction etc.). Synonyms frequently used are: edge, arc, tie, relation.

Relations are the characterising feature of a network, so that different networks may be defined with the same elements if they share different types of relations. Network analysis takes into account the structure of relations among the actors and their ‘location’ in the network, deriving behavioural consequences for both the single actors and the whole system. The mathematical representation of a network is a graph. Given the biunivocal relationships between a system, the network representing it, the associated graph and its adjacency matrix, all these terms are considered to be equivalent and are used as synonyms in this thesis.

Structure [of system, network etc.] (Chapter 1 and section 5.1)

A combination of the components of the system (individual elements, the relationships among them) and how they are assembled (the number and the distribution of both elements and relationships). In this definition, structure refers both to the static elements and the dynamic processes that govern the system.

System (section 2.1)

Assembly of elements in which each one is related to some other. Any element which has no relationship with any other element of the system cannot be a part of that system. Many systems have a variable relation with the environment in which they are embedded and this relation affects both the composition and the behaviour of the whole system: they are called dynamic systems.

Tourism destination (section 4.2)

"A physical space in which a visitor spends at least one overnight. It includes tourism products such as support services and attractions, and tourism resources within one day's return travel time. It has physical and administrative boundaries defining its management, and images and perceptions defining its market competitiveness. Local destinations incorporate various stakeholders often including a host community, and can nest and network to form larger destinations" (UNWTO, 2002a). More generally a system composed of:

- a large endowment of resources: natural, cultural, artistic, but also artificially built, such as museums, theme parks or sport complexes;
- a network of groups of actors: economic, non economic and institutional, whose prevalent activity is providing tourism related services to visitors and travellers.

Appendix 2: Ethical clearance

The following pages contain copies of the Ethical clearance application form and of the authorisation.

THE UNIVERSITY OF QUEENSLAND AUSTRALIA

School of Tourism and Leisure Management

CLEARANCE NO°
(Office use only)

Application Form for Ethical Clearance for Research Involving Human Participants

Please be advised that National Health and Medical Research Council (NHMRC) Statements and guidelines referenced in this document may be sourced from the NHMRC website, available at: *http://www.health.gov.au/nhmrc.*"

- **Departmental Committee** *(Honours, Postgraduate Diploma, Coursework Masters Research Higher Degrees and Supervised Projects)*

Project Title:
Network analysis of a tourism destination

Student name:	Rodolfo Baggio
Student No°:	41033159

Supervisor/s:	Prof. Chris Cooper Dr. Noel Scott

Contact Details of student:	***Phone***	***E-mail***
	+39 335 6168425	rodolfo.baggio@unibocconi.it r.baggio@uq.edu.au

Project Location:	Italy	**Project Duration:**	2 to 3 years

Does This Submission Relate To A Previous Protocol? (similar or an amendment) *If YES, please provide clearance no°*________________	☐YES ☒NO

Does This Submission Hold Other Ethical Clearance?: **Note:** *Attach copies from Other **AHEC Fully** Registered Ethics Committees*	☐YES ☒NO

PLEASE ANSWER ALL OF THE FOLLOWING QUESTIONS:

1) Who Are The Participants?: *Children, University Students, or other persons.* **Note:** *Details of approximate number, age range, and male/female ratios are required.*
Managers of tourism organizations and companies. estimated approximate number: 50 to 100 estimated age range: 25 to 65 estimated male/female ratio: 50/50

2) Participant Recruitment Details:
Please exact provide details of contact

The contact with participants will be by letter, fax or email.
The objectives and details of the project are explained and a formal request for an interview will be made.
The message (letter) will provide an explanation of the reasons why the potential participant has been contacted and a justification for the research.
The same letter presents the questions that will be asked during the interview (or the type of issues to be discussed in non-structured interviews) and it will offer the option of choosing different settings for conducting the interview. The letter will include the contact details (telephone, email, home address) of the researcher.
The potential participants will be asked to manifest their willingness to participate in the study by formally answering the contact letter.
If no answer is received, the researcher will assume that the potential participant is not interested in the study and there will be no further contact.
The contact letter will also thanks in advance the potential participant stating clearly that she or he is totally free not to participate in the study.

3) In 'Every-day/Lay Language' Please Provide A Summary Of The Project:

The nature of development and evolution of a tourism destination (TD) has profound implications not only for public authorities, as an aid in control and planning, but also for all the destination stakeholders, based on the hypothesis that one on the main determinants of a TD competitiveness is a harmonic progress of all its components. Tourism destinations are dynamic evolving complex systems. Network analysis is one of the most promising series of methods that can be used to model a complex system. Today, we know that the topology of a network is not just a curiosity, but a predictable property of some types of them, that it may greatly affect the overall dynamical behaviour and explain and control a number of processes from the diffusion of ideas to the robustness to external attacks to the optimization of the relationships among the network components. These analysis techniques can be together considered a diagnostic method for collecting and analysing data about the patterns of relationships among people in groups or among organizations. They provide a view into the network of relationships that may give tourism organizations managers a strong leverage to improve the flow of information and to target opportunities where this flow may have the most impact on regulatory or business activities.
By applying methods and techniques of the so called "network science" this work aims at studying the evolution of the system and at simulating dynamical processes such as information and knowledge diffusion, efficiency optimization.
The main outcome will be a better understanding of the general behaviour and dynamical evolution of a TD with important implications for the general management of a tourism destination.
Understanding the nature of the development and its evolution will support not only tourism managers in control and planning, but also the public authorities, responsible for developing the infrastructure and maintaining environmental quality and sustainability.
So far none of these methods have been applied to the study of tourism destinations.

The destination chosen is the island of Elba (Italy).

4)What Is/Are The Specific Aim(s) Of The Project:

Analysis of the structure of network of relationships connecting the stakeholders of a tourism destination
Study of the historical evolution of such a network
Development of tools and methods to perform numerical simulations of the network characteristics and behaviours

5) Give Full Details Of The Research Plan:

The network will be reconstructed on the basis of data collected from publicly available sources (directories, internet, etc.).
These will be integrated and completed with a series of semi-structured interviews conducted at the destination chosen.

Participants will be managers of tourism organizations chosen on the basis of personal knowledge and following indications that may be given by other participants.
Data collected will be information on the type and the number of relationships connecting the different organizations.
Data will be used to build the network and to simulate possible evolutionary paths.
Data will also be used to test theoretical optimization models with respect to processes such as information and knowledge diffusion, economical relationships, adaptivity to external shocks, etc.

6.) Give Details Of The Ethical Considerations Attached To The Proposed Project:

The following issues have been addressed:

Issue	*Implication for analysis*
Worthiness of the project	A study that is only opportunistic is likely to be pursued with less care devoted to design and data collection
Competence boundaries	Unacknowledged incompetence is responsible for analytical weakness, accumulation of useless data and superficial conclusions
Informed consent	Weak consent leads to poorer data because respondents try to protect themselves in a mistrusted relationship
Benefits, costs and reciprocity	Study participants' concern about inequity of benefits and costs serves to jeopardize access and thin out data
Harm and risk	Participants' expectation of any harm impacts and the access to the information and on the quality of data
Honesty and trust	If people feel betrayed when reading the results of the research, it becomes almost impossible for them to accept it as a reasonable interpretation of what happened.
Privacy, confidentiality, anonymity	Breaching one of them impact on data quality because it reduces trust
Intervention and advocacy	A situation in which the researcher have "dirty hands" skews the understanding of the results
Research integrity and quality	Shaky intellectual grounds for the research impact on its confirmability, dependability, credibility and potential transferability to other settings
Ownership of data and conclusions	Use of audits, even done informally, in order to improve the goodness of conclusions
Use and misuse of results	Clarity on the use of findings encourages a strong technical attention to utilization effort

7) How Will Informed Consent Be Obtained From Participants?

Participants will be asked to sign a consent letter provided by the researcher.
The letter will be written also according to the Italian regulations for privacy and personal data.
(Italian laws No. 196/2003 and No. 675/1996)

8) Provide Details Of Procedures For Establishing Confidentiality And Protecting Privacy Of Participants:

Any direct information that could lead to the identification of any participant or any particular organization will be removed from the data that will be used for the project activities and may be published.
All data and information collected will be considered of confidential nature.
Italian laws No. 196 (2003) and No. 675 (1996) regulate strictly all the matters concerning the collection and the storing of personal and sensible data. For the purpose of this research no such data are required, however, the same recommendations will be used. (see next point)

9) Provide Details Of Data Security and Storage:

Italian laws No. 196 (2003) and No. 675 (1996) regulate strictly all the matters concerning the collection

and the storing of personal and sensible data. For the purpose of this research no such data are required. However, the same recommendations will be used. All the electronic and paper documentation for this research will be stored on researcher's personal spaces. Paper materials will be locked. Electronic materials will be password secured. Routine backups for data files and software applications used will be performed weekly. Backup copies are kept on physically separated media (CDs/external disk drives) and on a researcher's private password-protected space on the Internet.

10) In What Form Will The Data Be Collected:		
(i) Identified ☒	**(ii) Potentially Identifiable** ☐	**(iii) De-Identified** ☐

11) In What Form Will The Data Be Stored and/or Accessed:		
(i) Identified ☒	**(ii) Potentially Identifiable** ☐	**(iii) De-Identified** ☐

12) Does The Project Involve Any Of The Following Procedures? If YES, Give Details

a) The use of drugs

NO

b) Any invasive procedures (eg., blood sampling)

NO

c) The possibility of physical stress/distress, discomfort

NO

d) The possibility of psychological/mental stress/distress, discomfort

NO

e) Deception of/or withholding information from, participant at ANY stage of the project

NO

f) Access to data held by a Commonwealth Department or Agency

NO

g) Access to data by bodies or people other than the investigators *(eg., Medical Records)*

NO

h) Participant involvement by any "Vulnerable Groups"

NO

13) How Has The Possibility Of Participants Withdrawal From The Project Been Addressed?: **Note:** *Ensure that details and effects of withdrawal without prejudice AT ANY TIME have been considered and explained.*
Participants may withdraw at any stage if they wish or feel uncomfortable.

14) Please Provide Details Of Any Participant Reimbursement For Their Involvement In The Project. *Note: This could be by cash payment, food vouchers, free services or movie passes, etc.*
The participation in the project is completely voluntary and no reimbursement of any form will be offered.

15) In Undertaking This Research Do Any 'Conflict Of Interest' Issues Arise? **If YES, Please Provide Details.**
NO

16) Some Projects May Involve Permits From National Parks & Wildlife In Relation To Collection Of Data And Native Title Issues. How Have You Addressed This Issue?
Not applicable in this case

17) The Project May Relate To , Or Involve, Aboriginal & Torres Strait Islander People, What Additional Measures Have Been Used To Address This Aspect Of The Project?
Not applicable in this case

ATTACHMENTS:

1) Consent Form (where relevant) ☒Yes ☐No

2) Information Sheet (where relevant) ☒Yes ☐No

Note: *for External Use - forms should be released on letterhead and contain University Ethical Paragraph.*
Refer to Guidelines at:- http://www.uq.edu.au/research/services/human/paragraphs.html

3) Questionnaire *(if applicable)* ☒Yes ☐No

Note: please attach ONLY those developed or adapted specifically for this project

5) Gatekeepers ☐Yes ☒No

Note: *A 'Gatekeeper' is a letter of Authority and Recognition from an Organisation of ANY type Involved with the research on the project*

6) References ☐Yes ☒No

8) Other - Please Specify____

I, the undersigned researcher(s) have considered the ethical issues in relation to this project and agree to abide by *The University of Queenslands' Guidelines for Ethical Review of Research Involving Humans - 2000*
It is understood that this includes the reporting and monitoring roles associated with the approval by the University of Queensland.

Signature of student:______________________________

Date: 18 / September / 2006 .

Signature of Supervisor *(if applicable)*:______________________________

Date: / / .

An <u>Original</u> should be submitted to:

Associate Professor Ian Patterson
School of Tourism and Leisure Management
The University of Queensland
Ipswich Queensland 4305
ian.patterson@uq.edu.au

21st September, 2006

Mr. Rodolfo Baggio
40857125
rodolfo.baggio@uq.edu.au

Dear Rodolfo,
Subject: Re: Ethical Clearance Number TALM 25
I have examined your Application Form for Ethical Clearance for your study entitled "Network analysis of a tourist destination."

As you stated in your application form, you intend to:

- Approach approximately between 50-100 managers of tourist organizations and companies. You will make contact with them by letter, fax or email.
- If they agree to the interview, you will contact them to arrange a time to interview respondents. At the interview you will provide them the an information sheet and written consent form, ask them to read and sign it after explaining the purposes of your research, and before you start the interview process.
- I am pleased that you have written an information sheet and consent form and that you have also emphasised that they are not obliged to be involved in your study and that no names will be required.
- You also have stressed the confidential nature of the study and the right of participants to withdraw at any stage if they feel uncomfortable with any of the questions.
- You need to **change a sentence** in the consent form to state that your Research Study has been approved by the School of Tourism and tell participants that if they have any queries to contact me at ian.patterson@uq.edu.au You will also need to delete the sentence stating that they should contact the University Ethics Officer, and the telephone number and email address should also be deleted for the University.

I have examined the questions that you will ask the respondents and am happy for inform you that there are no other ethical considerations that warrant further attention and that I give you permission to proceed to collect your data.
Good luck with your study,

Ian Patterson PhD
Associate Professor
Ethics Officer
School of Tourism and Leisure Management
The University of Queensland
Ipswich Queensland 4305

Note: clearance letter received by e-mail on September 22, 2006

Interview protocol pro-forma

Introduction

My name is Rodolfo Baggio. I am a doctoral student at the School of Tourism and Leisure Management at The University of Queensland, Australia.
I am researching the network characteristics of a tourism destination.
During my research I am interviewing people that are involved or have a stake in a tourism destination.

Taking part is voluntary. If you don't want to take part, you do not have to give a reason and no pressure will be out on you to try and change your mind. You can pull out of the discussion at any time. Please note, if you choose not to participate, or pull out during the discussion this will not affect you. <u>You can withdraw from this interview at any time. At any time you can ask to remove part or all of the information provided from my recordings.</u>

All the information you give me will be recorded. It will then be confidential and used for academic purposes only. If you agree to take part, please sign the consent form attached.

Questions

Could you give me a little information about yourself and how long you have been involved in tourism in Elba?

With what companies do you have ongoing relationships?
What kind of relationships do you have with these companies?
How often do you exchange/have transactions?
Since when did you have relationships with these companies?
What is the importance of these relationships (low, medium, high)?

With what associations do you have ongoing relationships?
What kind of relationships do you have with these associations?
Since when did you have relationships with these associations?
How often do you have contacts?
What is the importance of these contacts (low, medium, high)?

With what public organisations do you have ongoing relationships?
How often do you have contacts?
What kind of relationships do you have with these associations?
Since when did you have relationships with these organisations?
What is the importance of these contacts (low, medium, high)?

INFORMATION SHEET

My name is Rodolfo Baggio. I am a doctoral student at the School of Tourism and Leisure Management at The University of Queensland, Australia. I am researching the network characteristics of a tourism destination. During my research I am interviewing people that are involved or have a stake in a tourism destination. This is the reason why I am asking you to take part in this study.

What will you have to do if you take part?

If you agree to take part, I shall ask you to answer some questions. There aren't any right or wrong answers – I just want to hear about your opinions. The discussion should take about an hour at the longest. Your participation is voluntary and there are no reimbursements or compensations for the time you will be using in answering my questions.

Do you have to take part?

No, taking part is voluntary. If you don't want to take part, you do not have to give a reason and no pressure will be out on you to try and change your mind. You can pull out of the discussion at any time. Please note, if you choose not to participate, or pull out during the discussion this will not affect you. <u>You can withdraw from this interview at any time. At any time you can ask to remove part or all of the information provided from my recordings.</u>

If you agree to take part what happens to what you say?

All the information you give me will be recorded. It will then be confidential and used for academic purposes only. The data will be collected and stored and will be disposed of in a secure manner. The information will be used in a way that will not allow you to be identified individually. The treatment and the storage of all the data collected will be in accordance with the official laws and regulations (in Italy: laws No. 196/2003 and No. 675/1996).

What do you do now?

Think about the information on this sheet, and ask me if you are not sure about anything. If you agree to take part, sign the consent form attached. The consent form will not be used to identify you. It will be filed separately from all other information. If, after the discussion, you want any more information about the study you can contact me by telephone: (+39) 335 6168425 or by e-mail: rodolfo.baggio@unibocconi.it

This study adheres to the Guidelines of the ethical review process of The University of Queensland and has been approved by the UQ School of Tourism. Should you have any queries or want to speak to an officer of the University not involved in the study, you may contact Prof. Ian Patterson at: ian.patterson@uq.edu.au.

THANK YOU VERY MUCH FOR YOUR HELP!

Participant Consent Form

Research Project:

Network analysis of a tourism destination

1. I have read the Information Sheet for this study and have had details of the study explained to me.

2. My questions about the study have been answered to my satisfaction, and I understand that I may ask further questions at any time.

3. I also understand that I am free to withdraw from the study at any time, or to decline to answer any particular questions in the study.

4. I agree to provide information to the researchers under the conditions of confidentiality set out on the information sheet.

5. I wish to participate in this study under the conditions set out in the Information Sheet.

6. I would like my information: (circle your option)
   a) returned to me
   b) other(please specify) ……………………………………………….

7. I consent/do not consent to the information collected for the purposes of this research study to be used for any other research purposes. (Delete what does not apply)

Participant's Name: ____________________________________

Participant's Signature: ____________________________________

Date: ____________________________________

Contact details: ____________________________________

Researcher's Name: Rodolfo Baggio
Contact details: rodolfo.baggio@unibocconi.it – tel. : (+39) 335 6168425

Researcher's Signature: ____________________________________