ENTER2007 - 14th International Conference on Information Technology and Travel & Tourism
January 24th – 26th, 2007 - Ljubljana, Slovenia

In Sigala, M., Mich, L, & Murphy, J. (Eds.), Information and Communication Technologies in Tourism 2007 - Proceedings of the International Conference in Ljubljana - Slovenia, Springer, Wien, 2006, pp. 279-288.

The websites of a tourism destination: a network analysis

Rodolfo Baggio[a,b],
Magda Antonioli Corigliano[a]
Valeria Tallinucci[c,d]

[a] Master in Economics and Tourism, Bocconi University, Italy
[b] School of Tourism, The University of Queensland, Australia
{rodolfo.baggio, magda.antonioli}@unibocconi.it

[c]EURAC research, European Academy of Bolzano, Italy
valeria.tallinucci@eurac.edu
[d]Centro Studi della Banca dell'Elba, Italy

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.

Keywords: tourism destination; internet; web; cooperation; economic and social integration; complex networks.

1 Introduction

The destination management approach has emerged, in the last years, as an important strategic and operational approach to foster the development of areas where tourism is a relatively important activity and where the economy may be significantly influenced by tourism revenues (Ritchie & Crouch, 2003).

The spectrum of definitions describing Tourism Destination (TD) is quite extensive and there are many difficulties in setting clear borders to it (Jafari, 2000). Beside the definition issues, a TD is the archetype of a complex adaptive system (CAS). It shares many (if not all) of the characteristics usually associated with a CAS: non-linear relationships among the system components, self-organization of the structures, resilience to external shocks (Holland, 1995). As for many such systems, it can be represented by a network graph whose vertices are the organisations (public or private) belonging to the local tourism system.

Information technologies, and the Internet in particular, have shown to be a crucial tool for the operations and the development of this "information intensive" sector (Poon, 1993). A website looks to be a major (and, probably, it will be the only one in the future) tool to conduct business in the tourism field (Law et al., 2004). Furthermore, the network system based on the Internet may represent, for a destination, the tool to bind together the local tourism supply operators allowing them to integrate the destination product/service chain and to develop common competences and know-how (Laesser, 2003; Pollock, 1995) Integrating the offer of a TD requires an adaptation of the operators in overcoming the internal competitive reality. They are required to acknowledge the need to cooperate in order to recommend the destination to the international tourism market and to promote a unique selling proposition.

The network is a virtual space where the operators can meet, exchange information and experiences, develop common projects and build partnerships with the objective of reducing risk, distribution efforts, and financial costs. Moreover, promoting a common project through interlinked websites provides increased visibility. Even a single organisation can thus have a higher probability to be seen and selected among the huge offering of the World Wide Web. This is also due to the score system set up by the main search engines, which recognises and values the importance of the interlinking among websites (Brin & Page, 1998).

The Web, by its very nature, is a complex dynamically growing network and, in this respect, its global properties are closely related to its topological structure. In the last years, a vast catalogue of studies on complex networks has been compiled (Boccaletti et al., 2006; Pastor-Satorras & Vespignani, 2004). The tourism sector, however, has not yet been included in this collection. This paper aims at filling this gap by presenting an investigation on the websites network of a tourism destination. Moreover, it will be shown that some of the network topological metrics can be profitably used to characterise important properties of the network and of the underlying social and economic system.

2 Network models

Although almost 200 years old, the origins are ascribed to a famous paper by Leonhard Euler dated 1736, the so called “science of networks” has had an incredible thrust only in the last few years. The topology of complex systems, portrayed as networks, has been shown to be a fundamental feature of such systems (Boccaletti et al., 2006) Behaviours and processes can be described and explained by taking into account the system’s general connectivity properties. The spread of viruses over a computer network or of diseases in a population, the formation of opinions, the diffusion of information or knowledge, the resilience after external shocks, all exhibit a strong dependence on some basic topological features of the network representing a system.

Many metrics have been used to describe complex networks (da F. Costa et al., 2005). Most of them come from the long tradition of social network analysis and have been increased by the recent work of the community of statistical physicists. A handful of these, are the ones commonly employed to fully characterise topology and behaviours of a complex network:

- *degree distribution P(k)*: the statistical distribution of the number (and sometimes the type) of the linkages among the network elements;
- *average path length L*: the mean distance between any two nodes;
- *diameter D*: the maximal shortest path connecting any two nodes;
- *clustering coefficient C*: representing how concentrated is the connection of the node's neighbours in a graph and giving a measure of the inhomogeneity of local density of links;
- *efficiency* (at a local E_{loc} or global E_{glob} level): which can be interpreted as a measure of the capability of the system to exchange information over the network;
- *assortativity*: which gauges the correlation between the degrees of neighbour nodes. If positive, the networks is said assortative; in it well-connected elements tend to be linked to each other. This quantity, connected to the clustering coefficient, has been recently shown to influence directly the formation of strongly connected subnetworks or communities and to give an indication of their strength (Quayle et al., 2006).

The mathematical expressions for these quantities can be found in one of the recently published reviews of the research in this area (Albert & Barabási, 2002; Boccaletti et al., 2006). These parameters can be used to classify the networks into three broad classes (Amaral et al., 2000):

- *single-scale networks*: in which P(k) exhibits exponential (Poissonian or Gaussian tails). This class contains the random graphs (ER) and the small world (SW) networks described by Watts and Strogatz (1998) which are characterised by large clustering coefficients and short average path lengths;
- *scale-free networks*: P(k) has a power-law distribution: $P(k) \sim k^{-\gamma}$. The distribution is largely uneven, there is no characteristic mean nodal degree (the mean of a Poissonian ER distribution), but some (few) nodes act as very connected hubs, with very large number of ties, while the majority of nodes have a small number of links (Barabási & Albert, 1999);
- *broad-scale networks*: for which the degree distribution has a mixed behaviour, a power law regime followed by some sharp cut-off (exponential or Gaussian decay) of the tail.

The general topological properties of the World Wide Web have been studied by a number of authors. In particular, it has been possible to highlight a complex structure of its components (web pages or websites). According to Broder et al. (2000) the structure has a "bow-tie" shape (see Fig. 1), in which it is possible to recognise a number of components according to their connectivity characteristics. The model, widely accepted, calls for a strongly connected component (SCC), formed by all pages mutually connected by a directed link, an in-component (IN) and an out-component (OUT), formed by nodes connected to the SCC in a monodirectional way. In addition to these, there is a series of secondary structures called TENDRILS, containing pages that do not link the SCC and cannot be reached from it; TUBES, directly linking the IN and OUT parts without crossing the SCC and some disconnected elements (DCC), similar to isolated islands, with no connection at all to the other components. This structure has also been identified in several subnetworks of the whole Web, strengthening the hypothesis of a self-similar arrangement, i.e. the fact that subnetworks of the Web show similar structures (Dill et al., 2002).

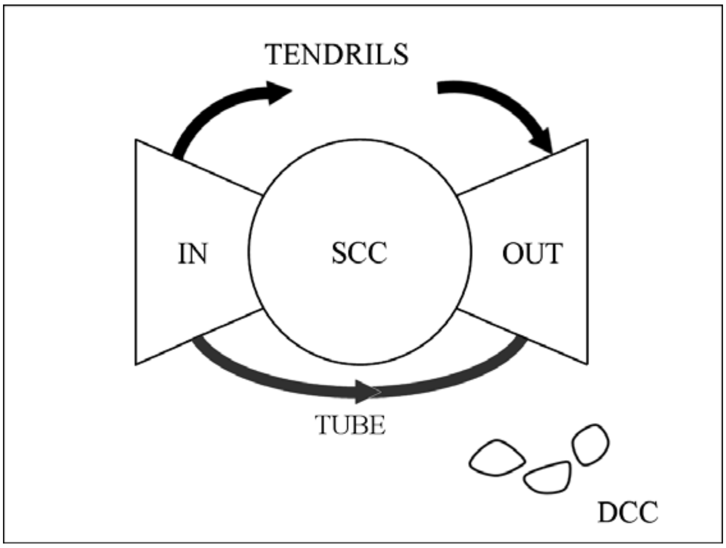

Fig. 1 A bow-tie model for the Web (after Baggio, 2006)

3 Methods

The websites of a tourism destination have been analyzed. The destination is the island of Elba, the main of the Tuscany Archipelago National Park and the third Italian island. It is an important environmental resource. Its geographic position, temperate climate, and the variety and beauty of its landscapes, coast and sea, make it a sea, sport and culture destination famous all over the world with almost 500 000 arrivals and 3 million overnights per year and several hundreds accommodation establishments (Tallinucci & Testa, 2006: 15-45). As for many other destinations, the Web has become, in the last years, an important means of promotion and commercialization for the whole community of local tourism operators.

A previous study (Pechlaner et al., 2003) had been conducted in order to reveal the interest to develop a business model based on the Internet as a local network platform able to manage a small island tourism destination. The empirical enquiry on the accommodation entrepreneurs showed a good *Internet attitude* (88.8% of the responders had their own homepage and 68,5% a site within a local tourism portal). The main motivations for a presence on the network were identified in the possibility to achieve visibility on a wider market and to acquire new customers, the speed of communication with their customers, but also the passive need to be present on the web. The independence of the operators emerged from the surprisingly high proportion of self-made websites (more than 30%). The websites were linked from some general destination portal (often an internet service provider), but the local suppliers tried to maintain independent and personal styles and direct connections with their guests.

The study presented here refocuses on the Elban network formed by websites belonging to “classical” tourism operators: accommodation (hotels, residences, camping sites etc.), intermediaries (travel agencies and tour operators), transport, regulation bodies, services etc in order to reveal their connectivity. The whole size of the network is of 468 elements (websites). It is not huge, but of a size which can be considered sufficient to show meaningfully the graph’s statistical properties (Angeloudis & Fisk, 2006; Dunne et al., 2002).

The websites have been analyzed considering them as the nodes of a complex network. Links among the websites have been listed by using a simple crawler and complementing the data obtained with a “manual” count of the hyperlinks to overcome the limitations of the program used (such as, for example, the impossibility to find hyperlinks embedded in Flash applications or Java applets). All links are considered of directed nature. The statistical analysis of the network thus obtained has been carried on by using standard software packages such as Pajek (Batagelj & Mrvar, 2004) and programs written by the one of the authors using the Matlab (MATLAB, 2004) development environment.

4 Results: the Elba tourism web graph

The graphical representation of the Elba websites’ network obtained is given in Fig. 2.

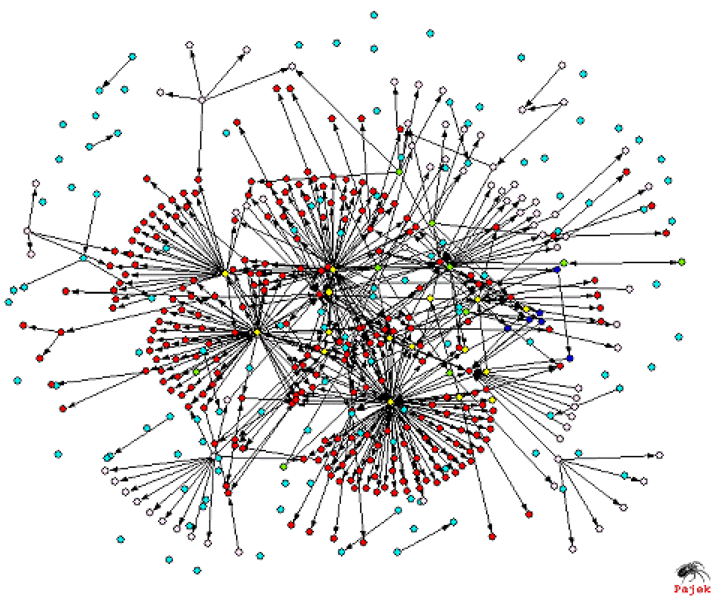

Fig. 2 The network of the Elban tourism websites

As it can be seen, the network exhibits a well identifiable structure. The network is rather sparse, its link density is d = 0.002 and almost 21% of the websites have no connection whatsoever to other sites. The diameter is D = 11, the average distance L = 4.5. These values are almost in line with those exhibited by similar networks (Albert & Barabási, 2002; Boccaletti et al., 2006; Pastor-Satorras & Vespignani, 2004).

The global clustering coefficient is: C = 0.003. Local and global efficiency (Latora & Marchiori, 2001) are: E_{loc} = 0.0145 and E_{glob} = 0.16981. These values are sensibly lower than those found for similar systems. The local efficiency value confirms the poor clusterisation of the network.

Key parameters characterising the structure of a directed network are the in-degree (k_{in}) and out-degree (k_{out}) distributions (the distributions of the number of links incoming to a node or outgoing from it). The cumulative degree distributions are shown in Fig. 3. Both display an almost perfect power law decay P(k) ~ $k^{-\gamma}$. The exponents calculated are: out-degree: γ_{out} = 1.89 and in-degree: γ_{in} = 2.96.

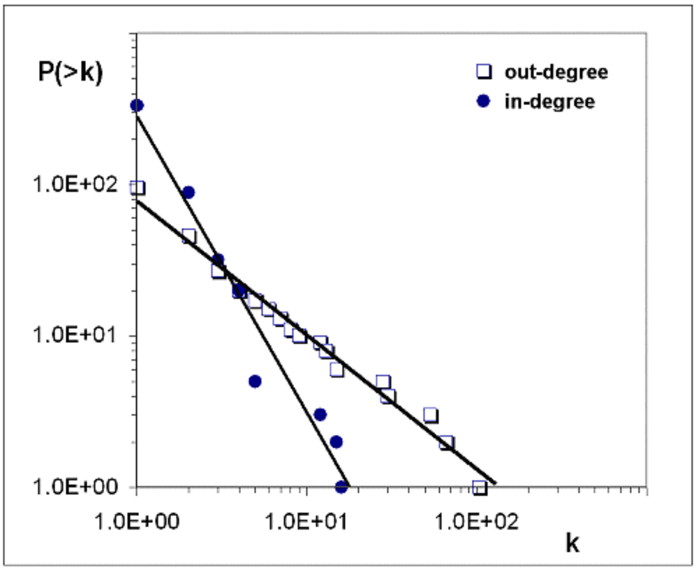

Fig. 3 In-degree and out-degree cumulative distributions for the Elba tourism websites' network

The in-degree exponent is higher and the out-degree exponent is lower than those typically measured for the Web: γ_{in} ~ 2.1 and γ_{out} ~ 2.7 (Broder et al., 2000). A spectral analysis confirms the main topological characteristics of the network, also acknowledging the scarce connectedness of the network.. Similar results were previously obtained, by using different techniques, in an assessment of the institutional websites of the island (Tallinucci & Testa, 2006: 126-135). The assortativity coefficient, calculated as the Pearson correlation coefficient between the degrees of adjacent vertices in the network (Newman, 2002) is: -0.101 ± 0.094.

The Elban network, besides the general low connectivity among its websites, still shows evidence of a bow-tie structure. Table 2 displays the estimated proportions for the components along with the values accepted for the whole Web (Broder et al., 2000). To be noted is the very low size of a central strongly connected component.

Table 1 Relative size of the components for the Elba network and the Web in the hypothesis of a bow-tie structure

	Elba Network (%)	Web (%)
SCC	3.4	28.0
IN	2.1	21.0
OUT	52.4	21.0
TENDRILS	15.6	21.0
TUBE	1.3	9.0
DCC	25.2	

Summarising the results of the statistical analysis of the Elba tourism websites' network, it is possible to say that:

- the results show a general agreement with similar results obtained by studying the Web and websites configurations (Broder et al., 2000; Dill et al., 2002). This may reinforce the idea of a substantial self-similarity in the structure of the Web space;
- some of the values, though, show quite different characteristics: basically, a much lower connectivity and a higher sparseness of the network (Pastor-Satorras & Vespignani, 2004);
- moreover, a very small and negative correlation among the degrees of the nodes has been found (assortativity coefficient), i.e.: nodes with high degree tend to connect with nodes with low degrees. This behaviour is opposite to what is typically found for social networks (Newman, 2002, 2003).

5 Discussion

The main consideration, which can be drawn from the results shown in the previous section, concerns the network in itself and the usage of the technological connections among the organisations acting at the destination.

The web network of Elba tourism stakeholders shows a very low connectivity and very low modularity (i.e.: low and sparse number of connected communities). This represents, essentially, a waste of resources both from a technical and an organisational point of view. It can be seen as a lost occasion in capturing the benefits and the advantages of cooperation mechanisms on the Internet. Mechanisms that could greatly ease the management of the destination and the efficiency in dealing with a globalised and highly competitive market (Barua et al., 2000; McLaren et al., 2002). Moreover, on a strategic perspective, it should be taken into account that the future search engines and recommendation systems will be based on dynamic agents whose task will consist in identifying connected communities on the web. Not forming a "community" will inevitably result in very poor reachability for a casual user, with detrimental consequences for the effectiveness of the marketing and communication activities and their economic outcomes (Baggio, 2006; Lawrence, 2000).

In this scenario, the quantities measured above assume also one more important meaning. First of all, the low connectivity of the Elba tourism network gives us an indication of a poor degree of "technological" collaboration or cooperation among the stakeholders. More importantly, two measurements, the clustering coefficient and the assortativity index, can be used as quantitative assessments of this specific characteristics. For their nature, the first one (clustering coefficient) can be thought as a "static" measurement of cooperation, i.e.: formation of cohesive communities inside the destination, while the second one (assortativity) can be interpreted as representing the tendency to form such communities.

In the case of Elba, these conclusions have been already shown elsewhere by performing more "traditional" qualitative analyses (Pechlaner et al., 2003; Tallinucci & Testa, 2006). It has been argued that the low propensity to reference the external world could find a foundation on the strong independent way to conduct the family-run enterprises and on the high rate of loyalty guests which produce a good *word-of-mouth* promotion.

6 Conclusive remarks

The network formed by the websites of a tourism destination has been analysed. The statistical mechanics tools developed in the last years for this purpose have been used to derive the main topological characteristics of this network.

The results show a general agreement with similar results, obtained by studying the Web and websites configurations. This reinforces the idea of a substantial self-similarity in the structure of the Web space. Some of the values, though, show different characteristics: basically a lower connectivity and a higher sparseness.

Four years after the first survey (Pechlaner et al., 2003), the *Internet attitude* of Elba's tourism operators has evolved, but less than their international competitors. In 2002, the web was seen as an opportunity to reach new guests and to open a window over the world, which potentially allowed to promote the destination on a global scale. Today, the advanced international destinations are able to sell integrated products and the individual tourist can satisfy their informational needs navigating through a well interconnected subnetwork. Elba, instead, still shows a very a scarce connectivity and a high rate of disconnected elements. This means, for example, that a guest has to identify a single supplier's website address in order to compose a holiday package. More importantly, this also means that the entrepreneurs are not cooperating to offer integrated products and services.

The analysis of the Elba’s website network allows to complement, with a quantitative assessment, analyses conducted with indirect or with qualitative methods, thus reinforcing the validity of those conclusions and allowing to overcome some of the limitations of those, such as the representativity of the samples chosen.

The authors believe also that by using structural network analysis methods, future work, already under way, may be able to better investigate the relationships between the technological network of websites and the real network of linkages among the stakeholders of a tourism destination. In this case some more tools may become available to address collaboration and cooperation issues, considered, by many, to be an important element for the success of a destination (Bramwell & Lane, 2000).

References

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Review of Modern Physics, 74*, 47-91.

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.

Angeloudis, P., & Fisk, D. (2006). Large subway systems as complex networks. *Physica A, 367*, 553-558.

Baggio, R. (2006). Complex systems, information technologies and tourism: a network point of view. *Information Technology and Tourism, 8*(1), 15-29.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286*, 509-512.

Barua, A., Whinston, A. B., & Yin, F. (2000). Value and Productivity in the Internet Economy. *Computer, 33*(5), 102-105.

Batagelj, V., & Mrvar, A. (2004). *Pajek - Program for Large Network Analysis*. Retrieved December, 2005, from http://vlado.fmf.uni-lj.si/pub/networks/pajek/.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics Reports, 424*(4-5), 175-308.

Bramwell, B., & Lane, B. (2000). *Tourism Collaboration and Partnerships: Politics Practice and Sustainability*. Clevedon, UK: Channel View Publications.

Brin, S., & Page, L. (1998). The Anatomy of a Large-Scale Hypertextual (Web) Search Engine. *Computer Networks and ISDN Systems, 30*(1-7), 107-117.

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.

da F. Costa, L., Rodrigues, A., Travieso, G., & Villas Boas, P. R. (2005). *Characterization of complex networks: A survey of measurements* (Preprint arxiv/cond-mat/0505185): ArXiv e-prints archive. Retrieved June, 2006, from http://arxiv.org/abs/cond-mat/0505185.

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.

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.

Holland, J. H. (1995). *Hidden Order: How Adaptation Builds Complexity*. Reading, MA: Helix Books.

Jafari, J. (Ed.). (2000). *Encyclopedia of Tourism*. London: Routledge.

Laesser, C. (2003). Destination management: sfide per la politica turistica del futuro. In H. Pechlaner, K. Weiermair & C. Laesser (Eds.), *Politica del turismo e destination management* (pp. 43-71). Milano: TCI.

Latora, V., & Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical Review Letters, 87*, 198701.

Law, R., Leung, K., & Wong, J. (2004). The impact of the Internet on travel agencies. *International Journal of Contemporary Hospitality Management, 16*(2), 100-107.

Lawrence, S. (2000). Context in Web Search. *IEEE Data Engineering Bulletin, 23*(3), 25-32.

MATLAB. (2004). MATLAB (Version 7, R14 - 2004). Natick, MA: The Matworks Inc.

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.

Newman, M. E. J. (2002). Mixing patterns in networks. *Physical Review E, 67*, 026126.

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review, 45*, 167-256.

Pastor-Satorras, R., & Vespignani, A. (2004). *Evolution and structure of the Internet - A Statistical Physics Approach*. Cambridge, UK: Cambridge University Press.

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.

Pollock, A. (1995). Occasional Studies: the impact of information technology on destination marketing. *EIU Travel and Tourism Analyst, 3:66-83.*

Poon, A. (1993). *Tourism, technology and competitive strategies.* Oxon, UK: CAB International.

Quayle, A. P., Siddiqui, A. S., & Jones, S. J. M. (2006). Modeling network growth with assortative mixing. *The European Physical Journal B, 50*, 617-630.

Ritchie, J. R. B., & Crouch, G. I. (2003). *The Competitive Destination: A Sustainable Tourism Perspective*. Oxon, UK: CABI Publishing.

Tallinucci, V., & Testa, M. (2006). *Marketing per le isole*. Milano: Franco Angeli.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small world' networks. *Nature, 393*, 440-442.