Chapter 2
Social and Communication Networks

For all practical purposes, our behavior is random. Unpredictable. Episodic. Indeterminable. Unforeseeable. Irregular. There's only one problem with this assumption. It's simply wrong.

—Albert-Lázló Barabási.

As any progress in science, also the current description of social networks has been accomplished one step at a time. Traditionally, the study of social and complex networks has been the territory of graph theory, which allows to define a network for any system by means of the simple representation of a graph. Since the 1950's real networks have been described as completely random graphs (Bollobás 1985), proposed as the simplest description of connections between entities and which mathematical formulation inheres in the work of Paul Erdös and Alfréd Rény (Erdös and Rényi 1959, 1960). According to this formulation, any member of the network has the same probability to be connected to any one else, thus all members have approximately the same number of connections. This vision radically changed in the late '90's with the empirical evidence that actually the structure of real complex systems is far from random (Barabási and Albert 1999; Watts and Strogatz 1998). In fact, in many real social networks, individuals exhibit preferential connectivity (higher probability that one of them will be linked to another one that already has a large number of connections), some connections are much more strong (or important) than others, social interactions are organized in tight groups or communities (Barabási and Albert 1999; Granovetter 1973; Watts and Strogatz 1998). Due to the increasing evidences of these structural heterogeneities, a big effort has been done from the scientific community to measure the topological properties and to understand, model and predict the mechanisms which regulate the formation of such structure and their impact on real phenomena. The standard way to address this has been focused on determining the contact network (who interacts with whom) in given time window, then characterize

"Bursts: The Hidden Pattern Behind Everything We Do" (Barabási 2010).

G. Miritello, *Temporal Patterns of Communication in Social Networks*,
Springer Theses, DOI: 10.1007/978-3-319-00110-4_2,

its elements by the aggregated properties measured in that time window and try to model the network dynamics to explain the observed structure. According to this frame, a time-aggregated and static picture of the network is given, where any inherent dynamics characterizing human behavior is neglected: the aggregated volume of interaction or the flux passing through a given link is the main quantity to assess the importance of that connection thus it fully determines its strength (or importance), interactions can happen at any time, there is no causality between events, communication is homogeneous in time, etc. This traditional approach, to which we will refer as *static approach*, was motivated by the expectancy that the characterization of the network structure would lead to a better knowledge of its dynamical and functional behavior. However, real social networks are dynamical objects, whose interactions between their members happen at a given time, may have a given duration and a causal relation. In this respect, the assumption of projecting out the temporal dimension may discard important information about the dynamical properties of real networks, their correlations with the topological ones and the dynamics of real phenomena. Only in the very last years, the large availability of massive databases of human behavior and interaction patterns such as e-mail, phone calls or online interactions databases led to the observation of non-trivial temporal properties of real social networks, with important implications on the way in which social networks and real phenomena have been traditionally understood and modeled. It has been observed, for example, that contrary to the predictions of statics approaches, individual actions do not happen at any time nor with the same probability to any other member of the network, that social interactions are not everlasting but, in contrast, they appear to be very unstable and volatile (Barabási 2005; Eckmann et al. 2004; Hidalgo and Rodriguez-Sickert 2008; Kossinets and Watts 2006; Vázquez et al. 2006). These results indicate that, as well as the (static) connections between members of a network, not even the temporal patterns of human actions and interactions can be modeled as random. Once again, therefore, the vision of how to model real networks is experiencing a radical change.

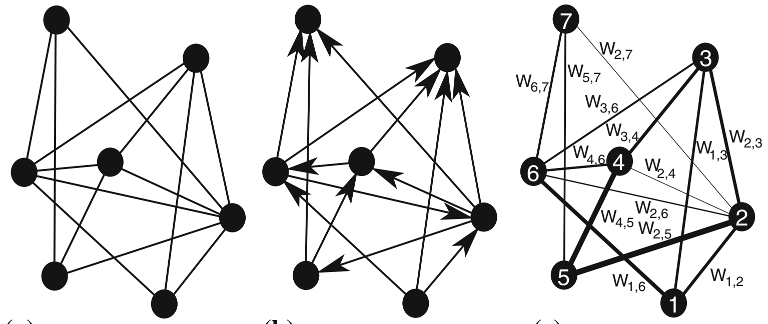
Without claiming to be complete, this chapter is intended to serve as a brief introduction of what social networks are, how they can be characterized and what are their main properties and features known in the literature. The first part of the chapter is dedicated to those topological properties of social networks that will be referenced in the rest of the thesis, what they represent and how they have been traditionally measured. Then we get more into the particular case of social networks that constitutes the main subject of this thesis: communication networks. We concentrate mainly on those aspects of communication networks that make them different from other social networks. After presenting the basic guideline of how the topology of social networks have been historically modeled, we discuss the main limitation of such traditional approaches on the basis of those temporal aspects of social networks that have been observed and measured only in the very last years. The latter part offers a first overview of the crucial role played by temporal aspects of human interaction into the characterization of social networks. At the same time it constitutes the starting point of all the work done in this thesis, presented in the following chapters.

2.1 Topology of Social Networks

Quantifying the topology is the first step towards detecting patterns in the network. The understanding of how a graph is wired, whether its member are equally connected or not or if every zone of the graph is reachable from any other one, gives very important insights on the network and the dynamical system it represents. For example, the way in which a disease or a piece of information would spread across society or how opinions form depends, above all, on the topology of the underlying graph. Nevertheless in many cases what we observe of a network is not the structure itself, but several instantaneous interactions between agents. Recover the very structure of how individuals are wired can be, therefore, a hard task. In most network studies the solution has been to aggregate all the interactions observed in the whole observation time window between any two members into a static edge between them, and possibly use the total number of interactions to asses the intensity or importance of the connection. Although, as we will see in the rest of this thesis, it represents only the first approximation of a network, this approach allowed to define an abundance of useful quantities to measure the main properties of real networks (Costa et al. 2007) and led to the discovery of several features which make them peculiar with respect to other types of networks, such as technological and biological networks (Newman and Park 2003).

2.1.1 Definitions and Notations

A network is a set of items, called vertices (or nodes), with connections between them, called edges (or links or ties) (Fig. 2.1). In mathematical terms a network is represented by a *graph* (West 1995). A undirected (directed) graph is a pair of sets $\mathcal{G} = \{\mathcal{P}, \mathcal{E}\}$, where $\mathcal{P}$ is a set of $\mathcal{N}$ nodes (or vertices) $p_1, p_2, ..p_N$ and $\mathcal{E}$ is a set of $\mathcal{M}$ edges (or links) that connect two elements of $\mathcal{P}$. A vertex typically represents an object (or an individual), while an edge represents a relation between two objects (or the same object). In a undirected graph, as the one depicted in Fig. 2.1a, each of the links is defined by a couple of nodes i and j, and is denoted as (i, j), e_{ij} or simply ij. In a directed graph, the order of the two nodes is important: e_{ij} stands for a link from i to j, and $e_{ij} \neq e_{ji}$. The property that two nodes in a directed network point to each other is called *reciprocity*. For a graph $\mathcal{G}$ with $\mathcal{N}$ nodes, the number of edges $\mathcal{M}$ is at least 0 and at most $\mathcal{N}(\mathcal{N} - 1)/2$ (when all the nodes are pairwise adjacent). Graphs are usually represented as a set of dots, corresponding to the nodes, which are joined together by a segment if the corresponding nodes are connected by a link (see Fig. 2.1). Usually, it may be useful to consider a matricial representation of a graph. Given a graph $\mathcal{G} = \{\mathcal{P}, \mathcal{E}\}$, it can be in fact completely described by giving the adjacency matrix $\mathcal{A}$, which is a $\mathcal{N} \times \mathcal{N}$ matrix whose element a_{ij} $(i, j = 1, ..., n)$ is equal to 1 if an edge e_{ij} exists and 0 otherwise. For undirected graphs $\mathcal{A}$ is thus



(a) (b) (c)

Fig. 2.1 Schematic representation of **a** a undirected, **b** a directed and **c** a weighted (and undirected) graph with N = 7 nodes and M = 14 edges. Adapted from "Complex networks: Structure and dynamics", Boccaletti et al. (2006)

a symmetric matrix. A *subgraph* $\mathcal{G}' = \{\mathcal{P}', \mathcal{E}'\}$ of the graph $\mathcal{G} = \{\mathcal{P}, \mathcal{E}\}$ is a graph with $\mathcal{P}' \subseteq \mathcal{P}$ and $\mathcal{E}' \subseteq \mathcal{E}$.
Although two nodes are not adjacent, they may however be reachable from one to the other. An important concept in graph theory is in fact the concept of *walk* or *path* from node i to node j, defined as a sequence of nodes and edges that begins with i and ends with j. The length of the walk is defined as the number of edges in the sequence.

In the case of social networks nodes represent a set of individuals or social entities linked through some kind of social interactions among them such as friendship, kinship, status, sexual, business or political, which define the links among them (Scott 2000; Wasserman and Faust 1994). Due to the recent development of communication systems, such as Internet or mobile phones, many other examples of social networks can be actually defined (Wellman et al. 1996; Wellman 2001), such as networks of human interaction through e-mail, Web forms, mobile phone or online social networks and services as Facebook, LinkedIn, Twitter. We refer to the latter as *communication* or *interaction* networks.

Node Degree and Assortative Mixing

The *degree* or *connectivity* k_i of a node i is the number of edges it has to other nodes and can be defined in terms of the adjacency matrix $\mathcal{A}$ as:

$$k_i = \sum_{j \in \mathcal{N}} a_{ij}. \tag{2.1}$$

In the case of a directed graph, the degree of a node is defined as the sum of the *out-degree* $k_i^{out} = \sum_j a_{ij}$ and the *in-degree* $k_i^{in} = \sum_j a_{ji}$, which measure respectively the number of outgoing and ingoing edges. Connectivity is a fundamental concept for networks. In real networks, not all nodes have the same number of edges and a first characterization of the network can be indeed obtained by dividing it in groups of nodes according to their connectivity. Especially for large networks, a more convenient characterization is obtained in terms of a distribution function (*degree distribution*) $P(k)$, which gives the probability that a randomly selected node has k edges. For directed networks, both $P(k^{in})$ and $P(k^{out})$ are defined. The n-moment of $P(k)$ is given by:

$$\langle k^n \rangle = \sum_k k^n P(k). \tag{2.2}$$

The first moment $\langle k \rangle$ defines the mean degree of the graph $\mathcal{G}$ and the second moment $\langle k^2 \rangle$ measures the fluctuations of the degree distribution. As mentioned above, until the late '90's many real networks have been modeled as random graphs (Erdös and Rényi 1959). A key prediction of random network theory is that most nodes have approximately the same degree, close to the average $\langle k \rangle$ of the network. In this case the degree distribution is a bell-shaped Poisson distribution with a peak at $P(\langle k \rangle)$, as the one depicted in Fig. 2.2a. Finding nodes that have a significantly greater or smaller number of links than a randomly chosen node is therefore rare. One also refers to random networks as exponential networks since the probability that a node is connected to other k nodes decreases exponentially (Haight 1967). For many real networks however, it has been found that $P(k)$ displays a power law shaped degree distribution $P(k) \sim k^{-\gamma}$ (see Fig. 2.2b), with exponent varying in the range $2 < \gamma < 3$ (Albert et al. 1999; Faloutsos et al. 1999; Jeong et al. 2000). In these networks, the average degree $\langle k \rangle$ is therefore well defined and bounded, while the variance $\sigma^2 = \langle k^2 \rangle - \langle k \rangle^2$ is dominated by the second moment of the distribution, which is highly fluctuating. Contrary to random networks, the average degree is not anymore a meaningful characterization of the network properties. Due to the property of power-laws of having the same functional form at all scales, such networks are referred as *scale-free networks* (Barabási and Albert 1999) and have been the focus of a great deal of attention in the literature (Albert and Barabási 2002; Dorogovtsev and Mendes 2002; Strogatz 2001). For many real networks, actually, $P(k)$ displays an exponential cutoff. However, its functional form still deviates significantly from the Poisson distribution expected for a random graph. Indeed, contrary to random networks, scale-free networks have a highly heterogeneous degree distribution, which results in the simultaneous presence of a few nodes (also called *hubs*) linked to many other nodes, and a large number of poorly connected elements.

The most known model to explain the origin of this scale invariance is the *Albert-Barabási* model (Barabási and Albert 1999) which is based on two basic ingredients: growth and preferential attachment, two key features of real networks. According to this model, a node with more links increases its connectivity faster than nodes with fewer links, since incoming nodes tend to connect to it with higher probability, a mechanisms that actually leads to the appearance of an hub hierarchy that exemplifies the scale-free structure (Bollobás et al. 2001; Dorogovtsev et al. 2000).

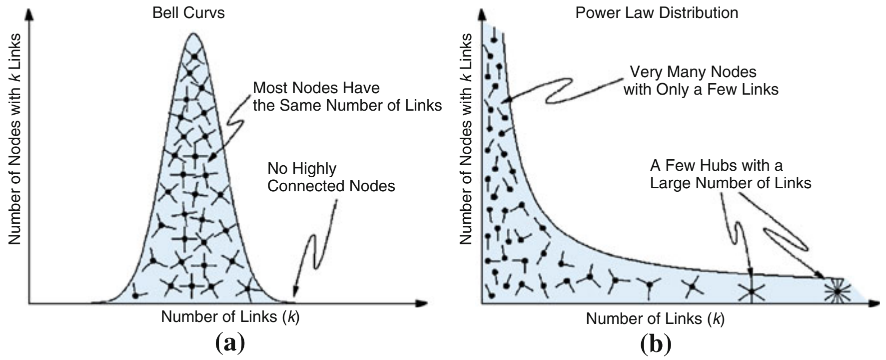


Fig. 2.2 Degree distribution $P(k)$ for a **a** random and **b** scale-free network. For random networks $P(k)$ follows a *bell-shaped* Poisson distribution where most nodes have the same number of connections. In contrast, in scale free-networks the degree distribution is power-law shaped and indicates that most nodes have only a few connections, while a few nodes are very highly connected (hubs). Adapted from "The architecture of complexity", Barabási (2007)

The *Albert-Barabási* model has attracted an exceptional amount of attention in the literature. In addition to analytic and numerical studies of the model itself, many authors have proposed modifications and generalizations to make the model a more realistic representation of real networks such as models with nonlinear preferential attachment, dynamic edge rewiring, fitness models and hierarchically and deterministically growing models (Albert and Barabási 2002; Dorogovtesev and Mendes 2001; Huberman and Adamic 1999; Goh et al. 2002; Gómez-Gardeñes and Moreno 2004).

Another class of models is based on three mechanisms: duplication (a randomly selected node and all its connections are duplicated); divergence (connections of a duplicated node are re-moved with a give probability) and mutation (connections are added from the duplicated node to a fraction of nodes which are not neighbors of the original node). The latter model also produces power-law degree distribution, although it fails to predict other properties of real networks that we will introduce in the rest of the chapter, such as degree correlations and the clustering coefficient (Solé et al. 2002). A better prediction for clustering coefficient is instead given by models that only includes duplication and divergence (Vázquez et al. 2003).

The fat tail of the degree distribution and the divergence of the second moment can affect the properties of a network, such as the clustering coefficient (Newman and Park 2003), which in many real social networks is much higher than the one expected for the corresponding random model (Amaral et al. 2000; Newman et al. 2002; Watts and Strogatz 1998). The heterogeneity of scale-free connectivity patterns also affects the behavior of dynamical processes that take place over the graph such as spreading processes. It has been shown, for example, that the presence of large degrees nodes favors epidemic spreading not only by suppressing the epidemic threshold, but also by accelerating the virus propagation in the population (Barthélemy et al. 2004; Moreno et al. 2002; Pastor-Satorras and Vespignani 2001b), a topic that will be analyzed in Chap. 5. The implications of this result can be very important in the set-up of dynamic control strategies, such as targeted immunization strategies, in populations with heterogeneous connectivity patterns (Pastor-Satorras and Vespignani 2002).

The degree distribution $P(k)$ describes all the statistical properties of uncorrelated networks. However, in many real networks the probability that a node with degree k is connected to a node with degree k', depends on k (correlated networks). In these cases, it is convenient to define the conditional probability $P(k'|k)$ which represents the probability that a node with degree k is connected to a node with degree k'. Another measure of degree-degree correlation is the *average nearest neighbors degree* of a node i:

$$k_{nn,i} = \frac{1}{k_i} \sum_{j \in N_i} k_j = \frac{1}{k_i} \sum_{j=1}^{N} a_{ij} k_j, \tag{2.3}$$

where N_i is the set of neighbors of node i. By means of this definition one can calculate the average degree of the nearest neighbors of nodes with degree k. The latter quantity, denoted as $k_{nn}(k)$, is related to the conditional probability as $k_{nn}(k) = \sum_{k'} k' P(k'|k)$ and implicitly incorporates the dependence on k. In absence of degree correlations one gets $k_{nn} = \langle k^2 \rangle / \langle k \rangle$, which shows the independence of k_{nn} on k. For correlated graphs, instead, k_{nn} depends on k. Depending on whether k_{nn} is an increasing or decreasing function of k, this properties is known as *assortative* or *disassortative* mixing (Newman 2002a). While in assortative networks nodes tend to connect with nodes with similar degree, in disassortative networks nodes with high degree are more likely connected with lowly connected ones.

Degree correlations and assortative mixing are very important properties in social networks since they have implications for questions as diffusion of information, network resilience or vaccination strategies (Callaway et al. 2000; Pastor-Satorras and Vespignani 2002). In disassortative networks, for example, a path between pairs of vertices can be destroyed by the removal of just a few of the highest degree nodes. Attacks on the highest degree vertices are therefore much more effective since the removal of few of them leads to a fast collapse of the whole network. On the contrary, in assortative networks, the removal of high-degree nodes is a relatively inefficient strategy for destroying network connectivity, since these vertices tend to be clustered together, thus their removal would result to be ineffective (Newman 2002a). In Sect. 2.2.1 we will see that in many social networks assortativity emerges as a natural phenomenon not only in the degree, but also with respect to a variety of psychological, sociodemographic and behavioral attributes (Christakis and Fowler 2007; Lewis et al. 2008; McPherson et al. 2001).

Shortest Path Length, Diameter and Betweenness

The shortest path d_{ij} between two nodes i and j measures the geodesic or optimal path way that go from i to j. A measure of the typical separation between two nodes in a graph $\mathcal{G}$ is given by the *average shortest path length* l, defined as the mean geodesic distance between nodes pairs (Watts 1999):

$$l = \frac{1}{N(N-1)\sum_{i,j\in N, i\neq j}} d_{ij}. \tag{2.4}$$

The maximum value of d_{ij} is called the *diameter* of the graph. In networks with more than one component (maximally connected induced subgraph), the definition in (2.4) can be problematic since there exist nodes pairs that have no connecting path. One can assign infinite geodesic distance to such pairs, but then the value of l also becomes infinite. For this reason, on such networks one usually defines l as the mean geodesic distance between all pairs of nodes belonging to the largest connected component (Latora and Marchiori 2001). Despite their large size, most of the real networks usually show a relatively short path between any two nodes. This feature is known as the *small-world effect* and is mathematically characterized by l, that depends at most logarithmically on the network size N (Watts and Strogatz 1998; Watts 1999).

In the social context, the small-world effect was first investigated by Milgram in the 1960s, in a series of experiments to estimate the actual number of steps in a chain of acquaintances (Milgram 1967). In its first experiment, Milgram asked randomly selected people in Nebraska to send letters to a distant target individual in Boston, identified only by his name, occupation and rough location. The letters could only be sent to someone whom the current holder knew by first name, and who was presumably closer to the final recipient. Milgram kept track of the paths followed by the letters and of the demographic characteristics of their handlers. Although the common guess was that it might take hundreds of these steps for the letters to reach their final destination, for those letters which arrived at destination, Milgram found that it had only taken an average of six steps for a letter to get from Nebraska to Boston. He labeled this situation “six degrees of separation” (Guare 1990), a phrase which since then has passed into popular folklore. Although the experiment certainly contained many possible sources of error, the general result that two randomly chosen persons can be connected by a short chain of intermediate acquaintances has been subsequently verified, and it is now widely accepted (Dodds et al. 2003; Korte and Milgram 1970).

Two nodes i and k that are not directly connected by an edge in a graph, can be linked through the nodes belonging to all the paths connecting i and k. In this regard, a measure of the relevance of a given node is given by its *betweenness* b_i which measures the number of geodesics going through it and is defined as:

$$b_i = \sum_{j,k\in N, j\neq k} \frac{n_{ij}(i)}{n_{jk}}, \tag{2.5}$$

where n_{jk} is the number of shortest paths connecting j and k, while $n_{jk}(i)$ is the number of shortest paths connecting j and k passing through i. Together with the degree, the betweenness of a node is one of the standard measures of the centrality of a node in a network (Scott 2000). Betweenness centrality can also be viewed as a measure of network resilience, indicating how much effect on path length the removal of a vertex will have (Holme et al. 2002; Newman 2003b).

Clustering or Transitivity

In many networks it is found that if node i is connected to node j and node j to node k, then there is a high probability that node i will also be connected to node k. In the language of social relationships, this translate in: the friend of your friend is likely also to be your friend (Wasserman and Faust 1994). In the study of networks, this property is known as *transitivity* or *clustering* and, in terms of a generic graph $\mathcal{G}$ it means the presence of a high number of triangles (sets of three vertices each of which is connected to each of the others). This can be quantified by defining the clustering coefficient $\mathcal{C}$ thus:

$$\mathcal{C} = \frac{3 \times number\ of\ triangles\ in\ the\ graph}{number\ of\ connected\ triples\ of\ vertices\ in\ the\ graph}, \tag{2.6}$$

where a triple consists of three nodes connected by two (open triple) or three (close triple) and the factor 3 in the numerator accounts for the fact that each triangle contributes to three triples and ensures that $\mathcal{C}$ lies in the range $0 \leq \mathcal{C} \leq \infty$. An alternative definition of the clustering coefficient has been given by Watts and Strogatz (Watts and Strogatz 1998). It is obtained by defining a *local clustering coefficient* of a node i as

$$\mathcal{C}_i = \frac{number\ of\ triangles\ connected\ to\ node\ i}{number\ of\ connected\ triples\ centered\ on\ node\ i}. \tag{2.7}$$

One assumes $\mathcal{C}_i = 0$ for nodes with degree 0 or 1, for which both numerator and denominator are zero. The clustering coefficient for the whole network is given by the average

$$\mathcal{C} = \frac{1}{N} \sum_i \mathcal{C}_i, \tag{2.8}$$

and, by definition, $0 \leq \mathcal{C} \leq 1$. In sociologic literature, the local clustering coefficient $\mathcal{C}_i$ has been widely used as a measure of the "network density" (Scott 2000). In general for social networks, regardless of which definition of the clustering coefficient is used, the values of $\mathcal{C}$ tend to be considerably higher than for a random graph with a similar number of vertices and edges. This indicates that nodes tend to create tightly connected groups characterized by a relatively high density of ties (Watts and Strogatz 1998). According to Newman and Park, together with the degree correlations, this property is what makes social networks different from other networks (Newman and Park 2003).

Communities

Given a graph $\mathcal{G}$, a *community* is a cohesive subgraph $\mathcal{G}'$ whose nodes are tightly connected. Since the structural cohesion of the nodes of a graph can be quantified in several different ways, there are many formal definitions of community structures (Ahuja et al. 1993; Everitt 1974; Girvan and Newman 2002; Guimerá et al. 2003; Holme 2002; Newman and Girvan 2004; Wilkinson and Huberman 2004). The most known definition is based on the concept of a *clique* and requires that all pairs of community members have relationships with each other. A clique is a maximal complete subgraph of three or more nodes, that is a subset of nodes all of which are adjacent to each other, and such that no other nodes exist adjacent to all of them. By extending this definition, a *n-clique* is a maximal subgraph in which the largest geodesic distance between any two nodes is no greater than n. Community structures are a typical feature of social networks; it is a matter of common experience that people divide into groups along lines of interest, occupation, location, age or family ties (Newman and Girvan 2004). The different way in which an individual is embedded in the structure within the network is also related to the behavior or function he is likely to practice. Individuals belonging to tightly connected group may be crucial in providing emotional and material support to each other (Granovetter 1973; Wellman 2007), while individuals who act as bridges between groups may provide access to a greater variety of information (Eagle et al. 2010; Onnela et al. 2007b).

Social analysts were the first to formalize the idea of communities and develop mathematical measures and methods to define the cohesion of communities and identify subgroups (Wasserman and Faust 1994; Scott 2000). In the last few years there has been an increasing interest and research in this area, which has become one of the most prominent areas of network science (Arenas et al. 2010; Blondel et al. 2008; Danon et al. 2008; Kumpula et al. 2009; Lancichinetti et al. 2010a,b; McDaid 2010; Newman and Girvan 2004; Raghavan et al. 2007; Rosvall and Bergstrom 2008; Toivonen et al. 2006, 2007). Finding the communities within a network is in fact a powerful tool for understanding not only the structure and the growth mechanisms of a network, but also its functioning: a community in a social network might indicate a circle of friends, a community in the World Wide Web might indicate a group of pages on closely related topics, and a community in a cellular or genetic network might be related to a functional module.

Motifs

A *motif* $\mathcal{M}$ in a network is a pattern of interconnections occurring in a graph $\mathcal{G}$ at a number significantly higher than in randomized versions of the graph, i.e. in graphs with the same number of nodes, links and degree distribution as the original one, but where the links are distributed at random. $\mathcal{M}$ can be considered as a sub-graph of $\mathcal{G}$. The concept of motif was first studied for biological networks (Milo et al. 2002; Mangan and Alon 2003) and then extended to other networks from neurobiology and ecology to social networks (Zhao and Oliver 2010). The reasons of the high frequency of different subgraphs in a specific network are not totally understood. There are at least two possible explanations. On the one hand, it is possible that certain constraints on the growth mechanism of a network as a whole determine which motifs become abundant. On the other hand, it is well known that the structure has important consequences on the network dynamics and functional robustness (Valverde and Solé 2005; Vázquez et al. 2004). So that a particular sub-graph can become overrepresented because, due to its structure, it possesses some relevant functional properties (Milo et al. 2002). As we will see in Sect. 2.4.3 the presence of motifs can also reveal important correlations between the agents, which may have a causal explanation.

2.1.2 Weighted Networks

Up to now, we have focused on networks in which edges between nodes have a binary nature, in the sense that they are either present or not. This networks are known as unweighted networks and each connection is assumed to be equivalent to any other. Nevertheless, in many real networks not all edges have the same importance of role and display instead a large heterogeneity in the capacity and the intensity of the connections. For example, social relationships with family members are usually different from friends or acquaintances (Granovetter 1973; Roberts 2010; Wellman and Wortley 1990). In all these cases it may be useful to assign to edges attributes that somehow should allow to distinguish connections of different type. These systems are better described in terms of weighted networks, i.e. networks in which each link carries a numerical value w_{ij} measuring the *strength* or *weight* of the connection (see Fig. 2.1c). Depending on their weight, ties can be distinguished between "strong" (large weight) and "weak" ties (small weight). Examples of strong and weak ties are found in social networks (Barabási et al. 2002; Granovetter 1973; Latora and Marchiori 2001; Newman 2001a) as well as other types of networks such as neural networks (Latora and Marchiori 2001, 2003; Sporns 2003), airline networks (Barrat et al. 2004; Guimerá et al. 2005), network of pages on the Internet (Pastor-Satorras and Vespignani 2004), biological systems (Csermely 2004). In most of these networks, tie strength quantifies the attention or the flow of information through that connection. As we will see in the next section, in the case of social networks such as mobile phone, e-mail or online social networks, the weight of a tie is usually assigned depending on the volume of interaction between the two involved individuals in a given time window, which has been found to correlate with the intensity of the social relationship (Baym et al. 2004; Wellman and Haythornthwaite 2003). Exploring the strength of the ties in social networks helps in the understanding of the structure of the network and also of the dynamics of many phenomena that involve human behavior, such as communities formation, information spreading and social influence (Grabowicz et al. 2012; Onnela et al. 2007b; Toivonen et al. 2007; Watts 2004). To measure the level of edge reciprocity in weighted (directed) networks, one can also define the bias $b_{ij} = w_{ij}/(w_{ij} + w_{ji})$ (which takes values from 0 to 1), where w_{ij} and w_{ji} are respectively the number of calls from i to j and from j to i. Several generalizations of the quantities defined in the previous section for unweighted networks (average nearest neighbors degree, the shortest path length, the clustering coefficient or motifs) have been proposed to characterize the complex statistical properties and heterogeneity of weighted social networks (Barrat et al. 2004; Boccaletti et al. 2006; Onnela et al. 2005; Saramäki et al. 2007).

Node Strength and Strength Distributions

In a weighted graph each edge has a weight w_{ij}, which is equal to 0 if the nodes i and j are not connected. It has been found that the weights characterizing the various connections exhibit generally complex statistical features with highly varying distributions and power-law behaviors (Almaas et al. 2004; Colizza et al. 2006; Goh et al. 2001; Onnela et al. 2007b). For a node i, the sum of w_{ij} over all his connections defines the node strength s_i:

$$s_i = \sum_j w_{ij}, \tag{2.9}$$

which is a measure of node strength in terms of the total weight of its connections (Barrat et al. 2004; Yook et al. 2001; Onnela et al. 2003). When the weights are independent on the topology, the strength of nodes of degree k is $s(k) \sim \langle w \rangle\, k$, where $\langle w \rangle$ is the average tie weight across the whole network. In the presence of correlations, instead, one has $s(k) \simeq A\, k^{\beta}$ with $\beta \neq 1$ and $A \neq \langle w \rangle$ (Barrat et al. 2004; Miritello et al. 2012b). Together with the degree distribution $P(k)$, the strength distribution $P(s)$, which measures the probability that a node have strength s, gives useful information about the network. In many real network the node strength is related to the node degree, thus $P(s)$ is also heavy-tailed. The same is observed for the distribution of tie weights. The local coupling between node and tie strength and network topology has important consequences for the network's global stability if ties are removed, as well as for the spread of news and ideas within a network (Onnela et al. 2007b).

Node Disparity

For a node i with a given connectivity k_i and a given strength s_i, there are different combinations of w_{ij}. In the two limit cases, all weights can be of the same order s_i/k_i or, in contrast, only one or few weights can dominate over the others. To measure this level of diversity, a quantity which is widely used in network literature is the *disparity* Y_i (Almaas et al. 2004; Barthélemy et al. 2003, 2005; Boccaletti et al. 2006; Miritello et al. 2012b), given by:

$$Y_i = \sum_{j=1}^{k_i} \left(\frac{w_{ij}}{s_i} \right)^2. \tag{2.10}$$

The disparity is a measure of local heterogeneity and it has an implicit dependence on k_i. In the homogeneous case, in which all edges have comparable weights, $Y_i(k) \sim 1/k$ since $w_{ij} = s_i/k_i$. In contrast, if the weight of a single edge dominates, then $Y(k) \simeq 1$ and it is independent of k. Other measures to quantify the topological diversity in a network have also been used, as the Shannon Entropy H_i or the Rényi Disparity $D_i(\gamma)$, where γ is a tunable constant (Eagle et al. 2009; Lee et al. 2010). These quantities however are strongly related to each other: in fact H_i behaves like $1/Y_i$ and D_i reduces to $1/Y_i$ in the case $\gamma = 2$, while for $\gamma = 1$ it reduces to the Shannon disparity, which is the exponential of the Shannon entropy.

2.2 Communication Networks

Communication (or interaction) networks, as the same word indicates, are a particular case of social networks where interactions between individuals refer to sequences of communication events. Examples of communication networks include e-mail networks, text messages or phone-calls, communication via blogs, online friendship networks as Facebook or micro-blogging services as Twitter. The increasing understanding and modeling of many real social networks can actually be attributed to the analysis of communication networks derived from massive electronic data set generated by millions of people through all these communication channels (Lazer et al. 2009). These records, which are routinely collected on websites, communication companies or electronic providers, represent a very rich laboratory to study how humans act and interact and to understand phenomena such as computer viruses or disease spreading (Anderson and May 1992; Newman 2002b; Newman et al. 2002), diffusion of information, innovations and products (Aral and Van Alsyne 2007; Dodds and Watts 2003; Szabó and Barabási 2006), opinion and influence dynamics (Friedkin and Johnsen 1990; Quattrocchi et al. 2010), teams formation (Guimerá et al. 2005) and so on. In contrast to others social networks, edges in communication networks typically arise from instant communication events and capture relationships as they happen. At any given instant, in fact, the network consists of the collection of ties connecting the people who are currently having a conversation. Together with the large dimension of these databases, this is certainly one of the main advantages of the study of communication networks. However, while in networks such as co-authorship networks a relationship between two scientists can be easily inferred whenever they coauthored at least one paper together (Newman 2001a), in many communication networks the problem of tie definition is usually not so trivial. For example, in mobile phone networks not all the observed events necessarily correspond to a social relationship between the individuals involved, since they may refer to calls to the operator or to wrong numbers. At the same way, declared “friends” in

Facebook may not correspond to the real individual social circle, since users actually interact only with few of them (Baym et al. 2004; Feld 1991). In this respect, to infer unobserved social relationships from the observed communication events is usually a hard task (De Choudhury et al. 2010; Wuchty 2009). Another important thing to bear in mind when dealing with networks of human communication is the different nature they may have. Although it is not always relevant, sometimes it can be important to distinguish between *offline* and *online* networks, where *offline* usually refer to phone calls or short messages (SMS) networks, to distinguish them from *online* communication such as blog interaction, Facebook, LinkedIn or Twitter networks. Another possible classification is related to the number of individuals involved at each interaction event. In this respect, while mobile phone or text services networks typically represent *one-to-one* communication networks, in the case of e-mail, Twitter or Facebook one may talk about *one-to-many* communication networks, since interaction can involve many recipients. Finally, it is often important to account for the directional/undirectional nature of the communication. While a phone call allows a bidirectional communication and the party that initiates the call may be only partially relevant, when dealing with SMS, e-mails or direct messages, to distinguish between the sender and the receiver may be instead crucial to characterize the underlying social relationship.

Each of these communication channels represents of course only one of the several possible ways in which two individuals may be connected in the real life. However, the use of electronic data as a proxy for social interactions has already proved successful in several recent investigations. For example, in the case of mobile phone networks, it has been observed that communication ties constitute an accurate representation of face-to-face interaction and self-reported friendships as measured using traditional sociometric methods (Eagle et al. 2009). This allows the quantification of previously rather elusive quantities such as tie strength that serve as signatures of work, family or acquaintances relationships (Onnela et al. 2007a,b), led to analyze how social groups evolve and change over time (Palla et al. 2007a,b). Other studies of e-mail communication networks have also shown that the use of e-mail in local social circles is strongly correlated with face-to-face and telephone interactions (Baym et al. 2004; Wellman and Haythornthwaite 2003; Wuchty and Uzzi 2011) and that the patterns of e-mail communication are related to the underlying social structure, shared activities, and personal attributes (Kossinets and Watts 2006; Wuchty and Uzzi 2011). Questions like how well electronic communication represent real social relationships or until what extent social media can predict the intimacy of a social relationships have been addressed also in online settings (Adamic and Adar 2003; Golder et al. 2007; Kivran-Swaine et al. 2011; Ugander et al. 2011). For example, a comparison between a network of Facebook interactions and self-reported data revealed that it is possible to infer the existence and the importance of a *offline* social contact by looking at the *online* network of the involved individuals and its properties such as the number of exchanged messages, number of common friends, etc. (Gilbert and Karahalios 2009). Furthermore, location-based networks such as Foursquare or GPS records of mobile phone networks provide important information on the physical places that people visit and how they move (Scellato et al. 2011; Volkovich et al. 2012), by providing more interesting insights about human mobility patterns (Candia et al. 2008; González et al. 2008) or the relations between friendships and mobility (Cho et al. 2011; Cranshaw et al. 2010).

However, despite the success of the use of electronic communication to represent and study social relationships, in social network analysis the measurement and characterization of what constitutes a social link remains still an unanswered issue. As mentioned above, this is mainly due to the fact that networks of electronic communication significantly differ from the physical one in one substantial thing: by quoting Tang et al. (2011) “physical social networks are colorful (‘family members’, ‘colleagues’, and ‘classmates’)” while when looking at electronic networks they are usually “black-and-white”, in the sense that no information is given by the activity data on the type of the underlying social tie. To give a representation of social ties close to the real one, one usually keeps only ties that are reciprocated and assign them a weight that should reflect its importance and nature in real life (see Sect. 2.1.2). Of course there are several ways to assess tie reciprocity and as much in which tie weights can be defined and assigned. In line with aggregated and static approaches traditionally used to model social networks, one usually assesses the reciprocal character of a tie between two individuals depending on whether there has been at least one reciprocated pair of communication events between them during the whole observation time window under investigation. Within the same picture, communication tie weights are usually taken as the total volume or intensity of communication between any pair of individuals in the whole time period (e.g. number or duration of calls in mobile-phone networks or number of directed messages in Twitter) (Huberman et al. 2009; Onnela et al. 2007b). Indeed, it has been shown that the volume of communication correlates with the importance of the face-to-face relationship (Eagle et al. 2009). As we will see in this section, the interpretation of tie strength leads to observation and validation of many structural properties that were already known from smaller and/or face-to-face networks (Baym et al. 2004; Wellman and Haythornthwaite 2003). As a particular case of social networks, communication networks present many of the topological properties described in Sect. 2.1: large heterogeneity in both the social connectivity and the strength of nodes and ties; non-trivial clustering or network transitivity; positive correlations or assortative mixing between the degrees of adjacent vertices; they are often divided into groups or communities that, as it has recently been suggested by Newman (2003a), may account for the observed clustering. The aim of this section is to present some of the most outstanding results about communication networks known in literature, with a particular focus on all those topological aspects which make them different from other social networks. All the results presented have been obtained by considering a static snapshot of the network, resulting from the aggregation of all communication events over a given time window. There are however several limitations to this description, starting from the definition of tie weight as the mere volume of communication, limitations that will be discussed in more details in Sect. 2.3.

2.2.1 Topological Properties

As discussed in Sect. 2.1, a basic network characteristic is the degree distribution. We have seen that many social networks show in general a skewed degree distribution which follows a power-law behavior $P(k) \sim k^{-\gamma}$ with exponent γ between 2 and 3, indicating the existence of hubs or people with a very large number of connections. In this aspect, communication networks slightly differ from other social networks. For e-mail communication networks, for example, it has been observed that the distribution of the number k of a node's next neighbors obeys an exponential behavior $P(k) \propto \exp(-k)$ (Guimerá et al. 2003). According to other studies (Amaral et al. 2000; Newman et al. 2002), the truncation of the scale-free behavior in real world networks is due to the physical costs of adding ties and the limited capacity of an individual (Bonney 1956). Other results indicate instead a heavily skewed degree distributions for e-mail where a power-law $P(k) \sim k^{\gamma}$ with an exponent $\gamma \sim -1.8$ (Ebel et al. 2002; Ferrara 2012). Small deviations from a power-law behavior with $\gamma \in [2, 3]$ have been observed also for online social networks, where in general two different regimes are observed for $P(k)$: a rapid decay ($\gamma \sim 4 - 5$) for small k and a heavy tailed ($\gamma \sim 1 - 2$) for large values of k (Ahn et al. 2007; Ferrara 2012; Kwak et al. 2010; Mislove et al. 2007). The range of large k for this type of networks is usually associated to atypical users, i.e. those individuals with a large audience whose messages get broadcasted. A rapid decay of $P(k)$ has also been observed in mobile phone communication networks where, although the tail of the degree distribution is better approximated by a power-law that an exponential, the obtained exponent is significantly higher. Onnela et al. (2007a) for example, found an exponent $\gamma = 8.4$, in which cases the power-law distribution can be easily confused with exponential (Clauset et al. 2009). Despite the differences, most results for communication networks show however a decay in the degree distribution which is faster than a power-law, indicating that the hubs are few. In phone communication networks this decay is probably due to the fact that business numbers are usually filtered out from the analysis and that each event usually represents a one-to-one communication (Miritello et al. 2012b; Onnela et al. 2007b), in contrast with e-mail or instant messaging networks, in which the recipients of the message can be many and well-connected hubs are observed (Ebel et al. 2002). Other times the observed differences may depend on the way in which interaction ties are defined. In networks as Twitter, for example, the social graph resulting from the following/follower relations can significantly change from the interaction one, where a tie between two persons is considered only if there has been direct communication between them (Huberman et al. 2009).

Consistently with general results about social networks, also communication networks show a well organized structure typically characterized by the existence of communities of individuals. Since the existence of communities between individuals is intuitively clear in human society, it is not surprising to find communities also in all those networks that represent interactions between individuals. Generally, the distribution of community sizes shows a slow decay, indicating that there is no characteristic group size (Ferrara 2011; Grabowicz et al. 2012; Guimerá et al. 2003). In large online social networks such as Facebook however, it has been found that there is a high probability of finding a high number of communities that contain few individuals and a lower probability of finding communities constituted by a large number of members (Ferrara 2012; Leskovec et al. 2009). The latter result suggests that individuals are more likely to aggregate in small communities, such as those representing family, friends or colleagues, rather than in large communities. On the other hand, for some networks such as e-mail or mobile phone networks, results show that there exists an important number of communities with a large amount of individuals, constituting the heavy long tail of the observed power law distribution (Blondel et al. 2008). As well as for other social networks, also in communication networks the existence of a community structure can be associated to the assortative mixing of some attribute of the vertices. As mentioned in Sect. 2.1, social networks are assortative: people with many friends are connected to others who also have many friends. This has been observed for online communication networks (Kwak et al. 2010) as well as for mobile-phone networks (Onnela et al. 2007a). This human tendency to interact with individuals similar to themselves is also know as *affinity* or *homophily* and emerges not only in the degree, but also with respect to a variety of attributes from psychological states such as loneliness or happiness (Bliss et al. 2012; McPherson et al. 2001) or health attributes and habits (Christakis and Fowler 2007, 2008), to tastes and interests (Lazarsfeld and Merton 1954; Lewis et al. 2008) and sociodemographic features such as age or race (Ibarra 2002; Mollica et al. 2003). As we will show in Chap. 3, the inherent homophilous nature of humans not only emerges in topological or psychological and exogenous factors, but also in dynamical processes of human interaction which have not been thoroughly investigated so far (Miritello et al. 2012a).

Since the nodes of a network may have positions in space, in many cases, it is reasonable to assume that geographical proximity plays a role in deciding how to connect the nodes, something that has indeed been observed for many real networks (Barthélemy et al. 2003; Boguña et al. 2004; Kleinberg 2000). For example, in mobile phone and online social networks it has been observed that the probability for two individuals to be connected decays with their geographical distance (Lambiotte et al. 2008; Onnela et al. 2011). On the other hand, most of these networks also appear to be geographically disperse (Barthélemy 2003; Lambiotte et al. 2008; Kwak et al. 2010). This result can be easily understood since, although the geographical proximity is an essential condition for face-to-face interactions, this is not true for the majority of communication networks whose function is indeed to enable the interaction with people who do not necessarily share the same physical place. Despite their geographical dispersion, however, communication networks are highly connected in terms of graph-distance and they appear to be even smaller than other social networks. A very recent study of Facebook interaction network show for example that the average number of intermediate ties between two randomly chosen humans (shortest path length) is almost 4 (Backstrom et al. 2012). This value, which is significantly smaller than the 6-degrees found in the original experiment by Milgram (Milgram 1967), indicates that, when considering another person in the world, "a friend of your friend knows a friend of their friend, on average" (Backstrom et al. 2012).

2.2.2 Correlation Between Topological Structure and Tie Weights

Topological Overlap and the Strength of Weak Ties

As mentioned in the previous section, people tend to form groups with other people similar to themselves. This suggests that in general ties within communities have different properties than ties connecting the communities (bridges). One of the most known result in this respect, hypothesized by the american sociologist Mark Granovetter (Granovetter 1973), is that ties within communities tend to be stronger that the ones between them (Onnela et al. 2007b; Lewis et al. 2008). This hypothesis is known as the *weak ties hypothesis* and implies the existence of important correlations between local network structure at the level of communities, and interactions strengths. One way to measure this correlation is by looking at the relation between the tie strength w_{ij} between two nodes i and j in a network and the relative topological overlap of their common neighbors, defined as:

$$O_{ij} = \frac{n_{ij}}{(k_i - 1) + (k_j - 1) - n_{ij}}, \tag{2.11}$$

where k_i and k_j are respectively the degrees of the two nodes and n_{ij} the number of neighbors common to both of them (Onnela et al. 2007b). If i and j have no common acquaintances, then $O_{ij} = 0$ and the tie between the two nodes represents a potential bridge between two different communities. On the contrary, if i and j are part of exactly the same circle of friends, then $O_{ij} = 1$. A positive correlation between w_{ij} and O_{ij} has been observed, for example, in mobile communication networks (Onnela et al. 2007b). By defining w_{ij} as the total volume of communication between i and j over a given time window, Onnela et al. found that the more the time two individuals spend talking together is, the more their friends overlap or the other way around, which is in line with the strength of weak ties hypothesis. As a consequence, the network structure in the vicinity of a randomly selected individual is similar to the one depicted in Fig. 2.3: consistent to the weak ties hypothesis, the majority of strong ties are found within the clusters, while most links connecting different communities are much weaker. An alternative measure of the topological overlap is given by the Adamic-Adar index s^{AA} which refines the simple counting of common neighbors by giving the lower-connected neighbors more weights (Adamic and Adar 2003). It is defined as $s_{ij}^{AA} = \sum_{z \in N(i) \cap N(j)} 1/\log(k(z))$, where $N(i)$ and $N(j)$ are, respectively, the neighbors of i and j.

The structural configuration of a network can have global implications on its stability and functionality since links may have a different role or function in the system depending on their strength and/or their location with respect to the groups.

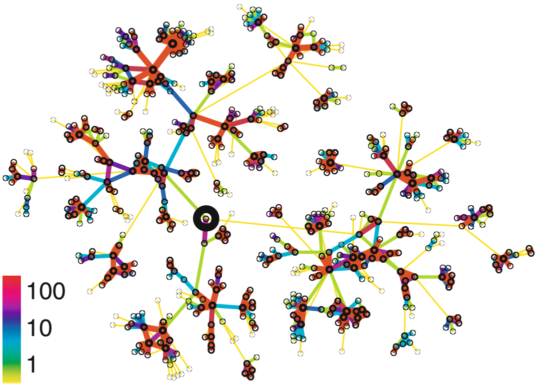


Fig. 2.3 Structure of a mobile phone network around a randomly selected individual (marked by the *black circle*), where only nodes which are at distance less that six from the selected one are shown. Each tie represents a reciprocated tie (mutual calls) between the involved users and tie weight is defined as the aggregated call duration in minutes (see *color bar*). Adapted from "Structure and tie strengths in mobile communication networks", Onnela et al. (2007b)

It has been observed, for example, that weaker ties are crucial for maintaining the network's structural integrity and that a removal of few of them from the whole network drives the system into a rapid disintegration. On the contrary, given that strong ties are predominantly within the communities, their removal only disintegrate a community but does not affect the overall integrity of the network (Onnela et al. 2007b).

This finding shows a significant difference between social networks and biological or technological ones, where exactly the opposite is observed and immediate network's collapse is caused by the removal of strong links (Barthélemy et al. 2004). Weak ties play an important role also in the dissemination of information within a network, since they help to link together different parts of a system while strong ties significantly slow information flow, trapping it in communities (Onnela et al. 2007b). Recently it has been proven that the weak ties hypothesis also applies to online networks (Grabowicz et al. 2012; Ferrara 2011). For example, by defining the strength of a tie ij as the total number of personal messages exchanged between i and j over the period of observation, Grabowicz et al. (2012) found that weak links are typically connections between persons not sharing neighbors and also in this case they contribute to more efficient information flow. In fact, while personal messages tend to concentrate inside the communities, retweets, which are associated to information propagation events, appear with higher probability in links between groups.

Dunbar's Number

Despite the large size of many real-world communication networks and their small diameter, they appear to organize in relatively small size communities of 50–200 individuals (Leskovec et al. 2009; Ferrara 2011). This observation agrees well with the so-called *Dunbar's theory* which predicts that the cognitive limit to the number of people with whom each person can have a close relationship is roughly 150. This number takes its name from the anthropologist Robin Dunbar who in 1992 measured the correlation between neocortical volume and typically social group size in a wide range of primates and human communities (Dunbar 1992). In real-world social interactions, beside the biological constraints, also other physical limitation may play their role, first of all the fact that time and attention are scarce resources and people have a finite amount of it to dedicate to social relations. Recently, Dunbar's theory has been tested with regard to Twitter users, and it was found that the new mode of communication did not have a significant impact on the human biological and cognitive limits to social interactions, with Twitter users maintaining a maximum of 100–200 stable relationships (Gonçalves et al. 2011). The Dunbar's theory also asserts that numbers of social relationships larger than 100–200 have generally a higher cost and require more effort to maintain a stable connection. This is in line with that found in Grabowicz et al. (2012) with regard to Twitter users: despite the existence of communities with size significantly larger than Dunbar's number, links with direct messages are much more abundant within groups of size up to 150 users. A similar result has been observed in Facebook communication where, despite the large number of declared friends, users only poke and message a small number of people (Golder et al. 2007).

As we will show in Chap. 3, we obtain similar results for mobile phone communication networks. We observe however a slightly smaller limit of the social capacity, probably because beside the cognitive limit, also temporal and monetary constraints play their role in phone communication (Miritello et al. 2012b).

2.3 Traditional Network Modeling

Up to now we have seen how to characterize the structure of a network, focusing on how the main properties of its elements have been traditionally defined and measured. We have also seen that in some cases the simple modeling framework can be made more powerful by including additional levels of details, as for example ties weights, whose correlations with the network topology give also important insights on how people behave, interact and organize. In the last years, the understanding of the structure of many communication networks has received huge interest among the scientific community and a lot of literature has been produced based on the analysis of phone call networks (Akoglu and Dalvi 2010; Hidalgo and Rodriguez-Sickert 2008; Nanavati et al. 2008; Onnela et al. 2007b), e-mail networks (Ebel et al. 2002; Eckmann et al. 2004; Guimerá et al. 2006; Kossinets and Watts 2006) and online social networks (Ahn et al. 2007; Ferrara 2012; Mislove et al. 2007; Ugander et al. 2011; Kwak et al. 2010; Huberman et al. 2009). The majority of studies on social networks structure have focused mainly on *(i)* the way real networks deviate from completely random networks, *(ii)* how the observed structure can emerge from individual behavior and *(iii)* how the observed network topology can be modeled. This type of analysis constitutes the first step to characterize network topology, understand which nodes play a similar role in a given system and the way in which they are connected to each other. Characterizing network topology not only helps in the understanding of how a part of the network differs from others, but also in the description and modeling of dynamical processes such as information spreading (Iribarren and Moro 2011a; Ugander et al. 2011) or influence (Aral and Walker 2012). All these processes, in fact, are constrained by the way in which whom and how each individual is connected to and located within the network. The fact that in social networks the flux of information that passes through each tie is unevenly distributed, some individuals are much more connected than others and social relationships are organized in communities, must indeed reflect in the way in which information, opinions or influence spread.

Paradoxically, the majority of these studies are based on aggregated statistics and ignore one inherent property of real social networks: the fact that they tend to change dynamically. Most social and communication networks are in fact *temporal networks*, not only in the sense that they are subject to a continuous evolution over time, but also because interactions between agents happen at a given time and may have a given duration (Holme and Saramäki 2012). Traditional network models are instead essentially static and all information about the time at which social interactions take place is discarded. The contact network is in fact obtained by aggregating over time all the interaction events observed in a given time window. This representation results therefore in a static snapshot of social interactions where the temporal dimension is completely projected out. In this representation all nodes and ties are considered as appearing at the same time and are assumed continuously active. In addition, the nature of social relationships reduces to a static strength which, although it incorporates the volume of communication between two individuals, it does not include any information about the way in which the involved individuals interact in time. The implications of these assumptions are very strong since they imply that people can interact with the same probability with any other individual in their social circle and that interaction can happen at any time and that there is no causal correlation between interaction events.

In contrast, everyday life experience suggests that none of these assumptions actually holds: people do not communicate everyday with any other people they know; interaction with one of our friends or colleagues may trigger the short-term interaction with other people; while some social relationship can last for years, others are very short in time or more occasional. Therefore, although the volume of communication may be an indicator of the importance or the role that a particular person has in our life, it does not capture any information on the temporal duration of a social tie or the way in which such a volume of interaction events is distributed within a given time window. In Chap. 4 we will show that, actually, the same tie strength can correspond to ties with very different temporal features and that static approaches thus radically misrepresent the real patterns of interactions and miss important aspects and tendencies of dynamical networks. As we will see in Chap. 3 and 5, the ongoing change of the networks may affect the instantaneous structure of the network, playing a fundamental role in the evolution of communities

(Palla et al. 2007a; Tantipathananand et al. 2007) and individual's communication strategies (Miritello et al. 2012a) and the way in which communication events are distributed in time is crucial in spreading phenomena (Iribarren and Moro 2009; Karsai et al. 2011; Miritello et al. 2011; Rocha et al. 2010; Vázquez et al. 2007).

One of the standard approaches to take into account temporal properties of interaction when modeling temporal networks has been to divide the time period under consideration into smaller time windows, then aggregate the social network in each window separately (Sarkar and Moore 2005; Snijders 2001). These approaches are, however, still static since they do not account for the fact that social relationships are mostly instantiated intermittently over time. An alternative method is given by the *time-respecting* graph where a tie (i, j) is defined as a time-labelled tie (i, j, t), where paths need to obey the time order of the appearance of ties (Kempe et al. 2002). According to this model however, temporally disconnected nodes are not considered and the frequency of contacts between nodes is not taken into account. A similar approach has been proposed where nodes, instead of ties, are labeled at each time instant they appear and, whenever a connection between two nodes is observed, a link between them is established, with weight equal to the time difference between the nodes' time appearances (Kostakos 2009). The main problem of these approaches is however the fact that, as we will show in Chap. 3, due to the strong heterogeneities of human interactions, nodes and/or ties activity can be confused with their appearance and, although no activity between i and j is observed at time t, a connection between them can, instead, exist. Other approaches use the concept of *reachability* to define temporal distance metrics where a directed tie from i to j is considered if there is a time-respecting path from i to j (Moody 2002; Tang et al. 2009, 2010). In contrast to others, the latter methods are able to capture the duration and time order of contact and have been show to be useful to quantify information diffusion processes (Holme 2005; Tang et al. 2009). To take into account the time between two consecutive contact on a path, generalizations of time-respecting path approaches which set a limit to the maximum allowed waiting time at a node have been also proposed (Pan and Saramäki 2011). All these methods certainly represent an improvement over the static aggregated approaches. However, a general frame of how to model dynamical social networks by taking into account both topological and temporal aspects of human behavior, as well as the correlations between them, is still lacking. The main reason why the temporal dimension has been neglected in traditional network modeling is certainly due to the fact that it is usually much easier to analyze static graphs. However, there are at least two other reasons why the underlying static network and the dynamical system usually appear separated and little is known about how to model temporal networks and their dynamical processes. The first reason, addressed in Chap. 3, is based on the belief that temporal processes happen very slowly such that the structure of the network is not remarkably affected. Although in many cases the temporal dimension can indeed be too insignificant over the periods of study used to be included in the network analysis, it can be of paramount importance in other cases. The second reason is associated to the lack of longitudinal data until few years ago. In fact, as mentioned above, the interest in modeling dynamically temporal networks has enormously increased only in the last few years with the rapid appearance of fine grained electronic longitudinal data, such as phone-communication records, emails, web, blogs and online social networks. The availability of this data has sparked numerous investigations into not only the topological, but also into the temporal properties of human interactions (Barabási 2005; Gaito et al. 2012; Goh and Barabási 2008; Kleinberg 2008; Kossinets et al. 2008; Kovanen et al. 2011; Jo et al. 2012; Rybski et al. 2010). What emerges is that temporal patterns of human interaction are actually very complex and articulated to be neglected in the description and characterization of social networks. Actually, as we will see in the next section, the inherent properties of temporal activity patterns also play a crucial role in all those processes where the temporal ordering of events is important, such as spreading phenomena, emergence of collective behavior, opinion formation or human synchronization, indicating that traditional models of social networks need to be revised.

2.4 Temporal Properties of Social Networks

As we have seen in the previous section, static descriptions of networks usually neglect the temporal dimension of human activity. Among others, some of the implicit assumption of this approach are that *(i)* nodes, edges, communities do not change in time, *(ii)* human actions are markovian and randomly distributed in time, therefore well approximated by Poisson processes, and *(iii)* there is no correlation or causality between interaction events. In recent years, however, there has been an increasing evidence that none of these hypothesis applies for real social networks. In this section we will see that these recent findings show that the dynamics of social networks is much more articulated and evolves as, contrary to static descriptions, social interactions are dynamical, tie decay/form and nodes enter/exit the social network. Since the understanding of how each of these properties affect the current way of modeling social networks is the main goal of this thesis, we will also discuss some of the implications that temporal patterns of communication have on the static description of the underlying contact network. This should serve to the reader as a preliminary baseline for all the results presented in rest of this work.

2.4.1 Nodes and Ties are Not Persistent

One of the basic assumption of traditional networks models is that nodes and ties within a network are continuously active. However, in the majority of real systems they are not since people may join and leave the network over time (Hidalgo and Rodriguez-Sickert 2008; Kossinets and Watts 2006). This makes the number of total nodes in the network not constant in time, thus changing the whole network structure and its properties (Ebel et al. 2002). For example, in co-authorship networks new authors can appear or abandon the network over time (Barabási et al. 2002), an effect

that has also be observed in other networks as internet dating (Holme et al. 2004) and mobile communication networks (Palla et al. 2007a). As well as individuals, also social relations are not always active in time and are characterized by very different lifetimes. As we will see in Chap. 4, while some relationship is observed only for the duration of an interaction, others may last beyond the interaction with which they began. In some cases this is due to the very definition of a link: since interactions usually have a given duration in time, ties activation and deactivation reflects the continuous changes in the activity and communication patterns of individuals. In all these cases, the instantaneous picture of the network can be significantly different from the aggregated one, where all the nodes and interactions observed in the entire observation period are taken into account. Moreover, the fact that individuals activate and deactivate social ties over time not only alters the structure of the networks in which they participate but also affect the dynamics of many phenomena that happen through the network, from communities formation (Palla et al. 2007a; Tantipathananand et al. 2007) to spreading processes (Iribarren and Moro 2009; Karsai et al. 2011; Miritello et al. 2011). There are many forces that govern the activation/deactivation of social relationships and the tendency for relations to weaken and disappear. The problem to identify under what conditions some ties are more likely to dissolve or persist, which will be the main focus of Chap. 4, is know in the literature as the *link prediction* problem and it has been the objective of many studies in recent years. The most exemplary work on edge decay in social networks is probably that of Burt who studied the social networks of the most important bankers over time and analyzed those factors that contribute to the disappearance of edges between them (Burt 2000, 2002) . Several other studies have focused on link characterization in other communication networks as phone and SMS networks (Akoglu and Dalvi 2010; Hidalgo and Rodriguez-Sickert 2008; Raeder et al. 2011), online networks (Aiello et al. 2010; Crandall et al. 2008; Gilbert and Karahalios 2009; Kivran-Swaine et al. 2011; Romero et al. 2011) and off-line settings (Burt 2000; Martin and Yeung 2006).
One of the main substantive conclusions that emerge from these studies is that links exhibit a memory, meaning that old links are more likely to persist in time than newly-formed ones (where the "age" and persistence of a link is defined in terms of observed communication events or when the link is registered or deleted in on-line social networks). Several other factors influence edge decay/persistence including *structural* properties of ties as reciprocation, neighborhood overlap or clustering coefficient, *temporal* properties as the time since the last communication (Raeder et al. 2011), homophily (Crandall et al. 2008) or *geographical* aspects as individual location and distance (Liben-Nowell et al. 2005; Lee et al. 2009). Tie creation or decay may also signal changes in the community structure and the condition for stability for large communities is strictly related to the continuous changes in their membership (Palla et al. 2007a).

Taking into account the ongoing appearance and disappearance of nodes and edges in the modeling of real social networks can lead to a characterization that significantly differs from the aggregated static one. In Chap. 3, for example, we will show that, since many social interactions are not always active over time, the instantaneous number of connections of an individual is actually much smaller than the time-aggregated one. This indicates that the standard social connectivity (defined in Sect. 2.1.1) usually overestimates the actual peoples social capacity of maintaining ties. We will show that according to the volatility of social relationships, it is possible to identify different individual communication strategies, from exploratory to stable. Depending on the adopted strategy, individuals also play a different role in spreading phenomena. Again, the dynamics of tie creation/removal is crucial since, in contrast, aggregated models assume that all connections are equally stable over time and that all users have the same communication strategy. Due to the importance that both social connectivity and its correlations with tie weights have on the characterization of network topology and on the modeling of many real phenomena as information spreading or network resilience, the consequence of accounting or not for the ongoing dynamics of human interactions may be therefore considerable.

2.4.2 Inter-Event Times and Bursty Behavior

As mentioned above, one of the implications of considering a static and aggregated snapshot of temporal interactions is that human interactions are randomly distributed in time, thus well approximated by homogeneous and memoryless Poisson processes (Greene 1997; Haight 1967; Reynolds 2003). Homogeneous Poisson processes have two main statistical properties: the number of events during a time interval of duration T follows a Poisson distribution with mean ρT and the time δt between consecutive events, called the *inter-event* or *waiting* times, follows an exponential distribution $p(\delta t) \sim \rho \exp(-\rho \delta t)$. As a consequence, individual actions happen at relatively regular time intervals δt and very short or very long inter-event times occur with small probability (see Fig. 2.4a). However, in real networks, this is usually not the case.

One of the main results that has emerged in the last few years from observing social interactions is that temporal patterns of human individuals are strongly inhomogeneous. This is reflected in the slowly decaying of the inter-event time distribution, which has been found to have a heavy tail with a power-law decay as $P(\delta t) \sim \delta t^{-\gamma}$, with $\gamma \simeq 1$ (Barabási 2005). As shown schematically in Fig. 2.4, the latter is in stark contrast with the prediction of a homogeneous Poisson process. This behavior seems to be a universal feature of human activity. It has in fact been observed in several systems driven by human activity sequences (Barabási 2010; Eckmann et al. 2004; Goh and Barabási 2008; Oliveira and Barabási 2005; Rybski et al. 2010) and is known in literature as *bursty behavior* since long periods of inactivity are separated by intense bursts of activity. According to other studies (Karsai et al. 2011; Rybski et al. 2010), in Chap. 5 we will show that bursty behavior is observed not only in the way an individual executes tasks, but also in the interaction between two individuals (Miritello et al. 2011). The heavy-tailed nature of the distribution of inter-event times generated a huge interest in the last years in the understanding what are the mechanisms responsible for its emergence. Two main classes of mechanisms have

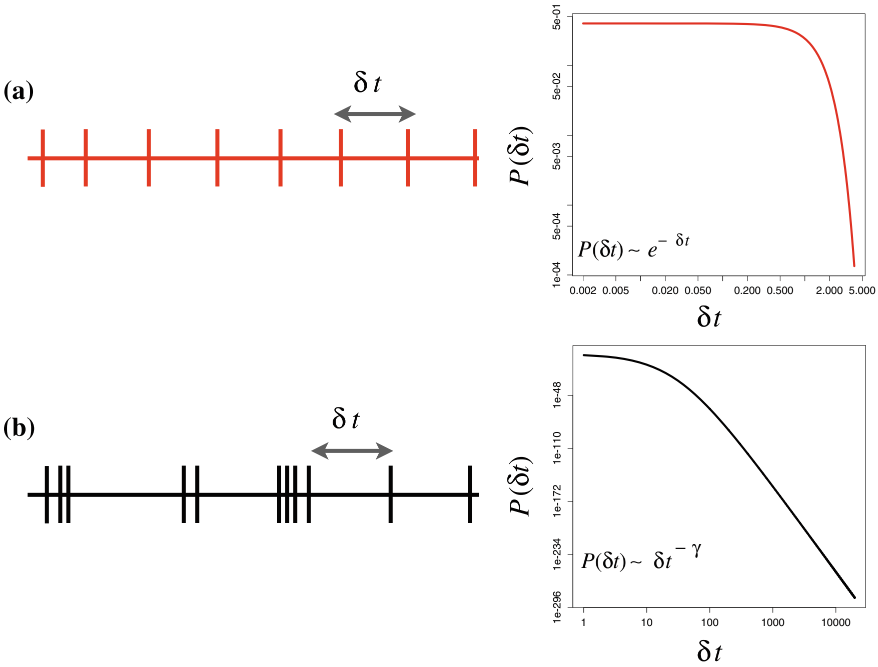


Fig. 2.4 Schematic representation of the difference between the homogeneous activity pattern predicted by a Poisson process and the heterogeneous behavior observed in human dynamics. **a** According to a Poissonian process the events are homogeneously distributed in time, thus the inter-event time δt follows an exponential distribution. **b** Bursty pattern of activity observed in real systems, in which bursts of events are followed by periods of inactivity giving rise to an heavy-tailed distribution for the inter-event time δt

been suggested: *(i)* human behavior is driven by external factors such as circadian and weekly cycles, which introduce a set of different time scales that give rise to the heavy tails (Malmgren et al. 2008); *(ii)* temporal inhomogeneities are caused from human task execution behavior, which is driven by a priority selection mechanism which introduce correlations in activity (Barabási 2005; Vázquez et al. 2006; Waleaevens et al. 2012). This latter mechanism is supported by a recent finding by Jo et al. (2012) who, after removing circadian and weekly patterns in the time-series by applying de-seasoning methods, observe the robustness of the inter-event times distribution of mobile phone communication events of individuals. The bursty dynamics of tie interactions has very important and at times drastic effects in the characterization of the underlying social network. For example, one of the implications of a Poissonian description of human dynamics is that social ties are characterized only by the number of communication events between the involved individuals, which regulates the rate of the process. This implies that all ties having the same weight are equivalent. In contrast, as we will show in Chap. 4, ties with exactly the same number of communication events can be instead characterized by a very different distribution of these events within their lifetime period. The bursty dynamics of human interactions, together with the temporal correlations between interactions, also have significant implications on many dynamical phenomena happening on the network in which the time ordering and delay is crucial, such as spreading processes. In recent experiments of electronic recommendation, indeed, it has been shown that the large heterogeneity found in the waiting times is responsible for the slow dynamics of information at the collective level (Iribarren and Moro 2009), which makes the observed dynamics significantly different from the Poissonian expectations (Iribarren and Moro 2011b; Karsai et al. 2011; Miritello et al. 2011; Vázquez et al. 2007). The latter topic will be discussed in details in Chap. 5.

2.4.3 Temporal Correlations: Motifs and Group Conversations

In Sect. 2.1.1 we have seen that the topology of many social networks is characterized by motifs, e.g. patterns of interconnections that appear with significantly high frequency. When dealing with temporal networks, however, these patterns can be affected by the temporal order at which interactions take place. Several ways have been proposed to extend the concept of motifs in order to take into account the changes in the network structure over time. Some of these consider snapshots of the network at different times, then look at the aggregated ties in each sub-window and count the different networks in these snapshots (Braha and Bar-Yam 2008), while others use the temporal information for defining the subgraphs of interest (Kossinets and Watts 2006; Kovanen et al. 2011; Zhao and Oliver 2010). The latter approach gives a more precise description of real temporal motifs since the sequences of events are based on temporal order of the events, instead of considering aggregated snapshot windows. The results show that the number of such paths is significantly larger when compared with reference networks, where the event times have been randomly reshuffled. Understanding recurrent temporally ordered patterns might yield a lot of insights on network analysis, especially in social and interaction networks, since it reveals the existence of correlations between patterns of communication, which may have a causal explanation. In communication networks, the presence of correlations between events has also been observed by looking at the distribution of the *relay time* τ_{ij}, also called *inter-contact time*, which measures the time it takes for the individual i to interact with j after having interacted with any other person $* \neq j$ (Cattuto et al. 2010; Eckmann et al. 2004; Isella et al. 2011; Miritello et al. 2011; Wu et al. 2010; Zhao and Oliver 2010). The relay time τ_{ij} depends not only on the inter-event times δt_{ij} in the $i \leftrightarrow j$ communication, but also on the possible correlations with the $* \leftrightarrow i$ events (Newman 2002b), thus on the way in which group conversations happen. In a first approximation, where the latter correlation is neglected, the relation between τ_{ij} and δt_{ij} is given by the waiting time density for δt_{ij}:

$$P(\tau_{ij}) = \frac{1}{\overline{\delta t}_{ij}} \int_{\tau_{ij}}^{\infty} P(\delta t_{ij})\, d\delta t_{ij}, \tag{2.12}$$

where $\overline{\delta t}_{ij}$ is the average inter-event time. The heavy-tail properties of the distribution $P(\delta t_{ij})$ of inter-event times are therefore inherited by $P(\tau_{ij})$, which appears to be skewed with a long-tail (Miritello et al. 2011; Karsai et al. 2011; Rybski et al. 2009). Interestingly, the results for $P(\tau)$ also show that not only large, but also very small relay times are much more probable when compared with the series of time-reshuffled events. This reveals indeed the existence of group conversations between individuals. As well as the other temporal inhomogeneities of human interactions described above, also correlations between events have crucial implications on the dynamics of real processes. Indeed, in Chap. 5 we will show that, together with the bursty behavior, group conversations are the main dynamical ingredient in the understanding of the spread of information in social networks and the principal responsible for the observed disagreement between Poissonian expectations and empirical results (Miritello et al. 2011).

2.5 Discussion

The purpose of this chapter has been to introduce the basic concepts and the main properties of social networks, with a particular interest on communication networks, which are the focus of this thesis. In communication networks each node usually represents an individual and the ties between individuals represent one or more communication events between them, such as phone-call, e-mail or online message communication.

We have seen that these networks, as well as many other social networks, are characterized by a very heterogeneous structure: some individual is much more connected than others, the flux of information which passes through social connection is not evenly distributed and social relationships are organized within communities. We have also presented how all these properties can be measured and analyzed in order to get insights on how people act and interact and have discussed the traditional way in which social networks have been modeled. We have seen that most of the analysis present in the literature has been restricted on characterizing the network topological structure in a given observation time window and on modeling its dynamics. In this analysis, real networks have been usually modeled by following what we call static approaches, where the temporal dimension is completely projected out: nodes and relationships between them are considered active during the whole period under investigation, interactions are basically characterized by the volume of interaction between the two end-nodes and can happen at any time and that the communication patterns are homogeneous in time, thus well approximated by Poissonian processes.

In contrast, many real systems are temporal networks, characterized by non trivial temporal patterns (Holme and Saramäki 2012). We have seen, in fact, that in many social networks relationships appear and disappear in time, leading to an instantaneous contact network which may be different from the one that emerges when they are considered everlasting (Hidalgo and Rodriguez-Sickert 2008; Kossinets and Watts 2006; Palla et al. 2007b). We have also seen that interaction events are correlated and human actions are bursty (Barabási 2005; Eckmann et al. 2004; Kossinets and Watts 2006; Vázquez et al. 2006) thus can not be modeled by homogeneous Poissonian processes. All these properties, which can not be captured by static network models, are emerging only in the very last years, thanks to the increasing availability of large electronic data sets of human communication that contain fine-grained temporal information of how people act and interact. We have also mentioned that, as well as topological features of human interaction, temporal properties are crucial to understand and characterize the underlying social network and the dynamical processes happening through it (Iribarren and Moro 2009; Vázquez et al. 2007), something that we will analyze in more details in the next chapters.

References

Adamic L, Adar E (2003) Friends and neighbors on the web. Soc Netw 25(3):211–230

Ahn Y-Y, Han SK, Moon S, Jeong H (2007) Analysis of topological characteristics of huge online social networking services. In: Proceedings of the 16th international conference on the, World Wide Web, pp 835–844.

Ahuja RK, Magnanti TL, Orlin JB (1993) Network flows: theory, algorithms, and applications. Upper Saddle River, New Jersey

Aiello L, Barrat A, Cattuto C, Ruffo G, Schifanella R (2010) Link creation and profile alignment in the aNobii social network. In: Proceedings of the sec- ond IEEE international conference on social computing socialCom 2010, Minneapolis, USA.

Akoglu L, Dalvi B (2010) Structure, tie persistence and event detection in large phone and SMS networks. In: MLG '10 Proceedings of the eighth workshop on mining and learning with graphs.

Albert R, Jeong H, Barabási A-L (1999) The diameter of the world wide web. Nature 491:130–131

Albert R, Barabási A-L (2002) Statistical mechanics of complex networks. Rev Mod Phys 74:47–97

Almaas E, Kovacs B, Viscek T, Oltvai Z, Barabási A-L (2004) Global organization of metabolic fluxes in the bacterium Escherichia coli. Nature 427:839

Amaral LAN, Scala A, Barthélemy M, Stanley H (2000) Classes of small-world networks. Proc Natl Acad Sci U S A 97:11149–11152

Anderson RM, May R (1992) Infectious diseases in humans. Oxford University Press, Oxford

Aral S, Walker D (2012) Identifying influential and susceptible members of social networks. Science 337:337–341

Aral S, Van Alsyne M (2007) Network structure and information advantage. Proceeding of the academy of management conference, Philadelphia, PA, In

Arenas A, Duch J, Gómez S, Danon L, Díaz-Guilera A (2010) Communities in complex networks: Identification at different levels. In: Caldarelli G (ed) Enciclopedia of life support system (EOLSS). EOLSS Publishers, Oxford, UK, Developed under the auspices of the Unesco

Backstrom L, Boldi P, Rosa M, Ugander J, Vigna S (2012) Four degrees of separation. arXiv:1111.4570v3.

Barabási A-L (2005) The origin of bursts and heavy tails in human dynamics. Nat Sci Rep 435:207–211

Barabási A-L (2007) The architecture of complexity. IEEE Control Syst Mag 27:33–42

Barabási A-L (2010) Bursts: the hidden pattern behind everything we do. Dutton Books, New York

Barabási A-L, Albert R (1999) Emergence of scaling in random networks. Science 286:509–512
Barabási A-L, Jeong H, Ravasz R, Neda Z, Vicsek T, Schubert A (2002) Evolution of the social network of scientific collaborations. Physica A 311:590–614
Barrat A, Barthélemy M, Pastor-Satorras R, Vespignani A (2004) The architecture of complex weighted networks. Proc Natl Acad Sci U S A 101:3747
Barthélemy M (2003) Crossover from scale-free to spatial networks. Europhys Lett 6:915
Barthélemy M, Gondran B, Guichard E (2003) Spatial structure of the internet traffic. Physica A 319:633642
Barthélemy M, Barrat A, Pastor-Satorras R, Vespignani A (2004) Velocity and hierarchical spread of epidemic outbreaks in scale-free networks. Phys Rev Lett 92:178701
Barthélemy M, Barrat A, Pastor-Satorras R, Vespignani A (2005) Characterization and modelling of weighted networks. Physica A 346:34–43
Baym NK, Zhang YB, Lin M (2004) Social interactions across media: interpersonal communication on the internet, face-to-face, and the telephone. New Media Soc 6:299
Bliss C, Kloumann I, Harris KD, Danforth C (2012) Twitter reciprocal reply networks exhibit assortativity with respect to happiness. J Comput Sci 3:388–397
Blondel V, Guillaume J-L, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. J Stat Mech, P10008
Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D-U (2006) Complex networks: structure and dynamics. Phys Rep 424:175
Boguña M, Pastor-Satorras R, Díaz-Guilera A, Arenas A (2004) Models of social networks based on social distance attachment. Phys Rev E 70:056122
Bollobás B (1985) Random graphs. Academic, New York
Bollobás B, Riordan O, Spencer J, Tusnady G (2001) The degree sequence of a scale-free random graph process. Random Struct Alg 18:279–290
Bonney M (1956) Sociometry and the science of man. Beacon House, New York
Braha D, Bar-Yam Y (2008) Time-dependent complex networks: dynamic centrality, dynamic motifs, and cycles of social interaction. In: Gross T, Sayama H (eds) Adaptative networks: Theory, models and applications. Springer, Dordrecht, pp 39–50
Burt R (2000) Decay functions. Soc Netw 22:1–28
Burt R (2002) Bridge decay. Soc Netw 24:333–363
Callaway D, Newman MEJ, Strogatz SH, Watts DJ (2000) Network robustness and fragility: percolation on random graphs. Phys Rev E 85:5468–5471
Candia J, González M, Wang P, Schoenharl T, Madey G, Barabási A-L (2008) Uncovering individual and collective human dynamics from mobile phone records. J Phys A: Math Theor 41:224015
Cattuto C, Van den Broeck W, Barrat A, Colizza V, Pinton J-F, Vespignani A (2010) Dynamics of person-to-person interactions from distributed RFID sensor networks. PLoS ONE 5:e11596
Cho E, Myers S, Leskovec J (2011) Friendship and mobility: user movement in location-based social networks. In: KDD '11 Proceedings of the 17th ACM SIGKDD international conference on knowledge discovery and data mining, pp 1082–1090.
Christakis N, Fowler J (2007) The spread of obesity in a large social network over 32 years. N Engl J Med 357:370–379
Christakis NA, Fowler J (2008) The collective dynamics of smoking in a large social network. N Engl J Med 358:2249–2258
Clauset A, Shalizi C, Newman MEJ (2009) Power-law distributions in empirical data. SIAM Review 51:661–703
Colizza V, Barrat A, Barthélemy M, Vespignani A (2006) The role of the airline transportation network in the prediction and predictability of global epidemics. Proc Natl Acad Sci U S A 103:2015–2020
Costa LDF, Rodrigues FA, Travieso G (2007) Characterization of complex networks: a survey of measurements. Adv Phys 56:167–242
Crandall D, Cosley D, Huttenlocher D, Kleinberg J, Suri S (2008) Feed- back effects between similarity and social influence in online communities. In KDD 08: proceeding of the 14th ACM

SIGKDD international conference on knowledge discovery and data mining New York, NY, USA, pp 160–168.
Cranshaw J, Toch E, Hong JI, Kitturr A, Sadeh N (2010) Bridging the gap between physical location and online social networks. In: 12th ACM international conference on ubiquitous, computing, pp 119–128.
Csermely P (2004) Strongs links are important, but weak links stabilize them. Trends Biochem Sci 29:331
Danon L, Arenas A, Díaz-Guilera A (2008) Impact of community structure on information transfer. Phys Rev E 77:036103
De Choudhury M, Mason WA, Hofman J, Watts DJ (2010) Inferring relevant social networks from interpersonal communication. In: WWW '10 Proceedings of the 19th international conference on, World Wide Web, pp 301–310.
Dodds P, Muhamad R, Watts DJ (2003) An experimental study of search in global social networks. Science 301:827–829
Dodds P, Watts DJ (2003) Information exchange and the robustness of organizational networks. Proc Natl Acad Sci U S A 100:12516–12521
Dorogovtesev S, Mendes JFF (2001) Effect of the accelerating growth of communication networks on their structure. Phys Rev E 63:025101
Dorogovtsev SN, Mendes JFF, Samukhin AN (2000) Structure of growing networks with preferential linking. Phys Rev Lett 85:4633–4636
Dorogovtsev SN, Mendes JFF (2002) Evolution of networks. Adv Phys 51:1079–1187
Dunbar R (1992) Neocortex size as a constraint on group size in primated. J Hum Evol 22:469
Eagle N, Pentland A, Lazer D (2009) Inferring friendship network structure by using mobile phone data. Proc Natl Acad Sci U S A 106(36):15274–15278
Eagle N, Macy M, Claxton R (2010) Network diversity and economic development. Science 328:5981
Ebel H, Mielsch L, Bornholdt S (2002) Scale-free topology of e-mail networks. Phys Rev E 66:035103
Ebel H, Davidsen J, Bornholdt S (2002) Dynamics of social networks. Complexity 8:24–27
Eckmann J-P, Moses E, Sergi D (2004) Entropy of dialogues creates coherent structures in e-mail traffic. Proc Natl Acad Sci U S A 40:14333–14337
Erdös P, Rényi A (1959) On random graphs I. Publ Math Debrecen 6:290–297
Erdös P (1960) On the evolution of random graph. Publ Math Inst Hungarian Acad 5:17–61
Everitt B (1974) Cluster analysis. John Wiley, New York
Faloutsos M, Faloutsos P (1999) On relationships of the internet topology. In SIGCOMM Comput Com 29:251–262
Feld S (1991) Why your friends have more friends than you do. Am J Sociol 95:1464–1477
Ferrara E (2011) A large-scale community structure analysis in Facebook. arXiv:1106.2503v3.
Ferrara E (2012) Topological features of online, social networks. arXiv:1202.0331v1.
Friedkin N, Johnsen E (1990) Social influence and opinions. J Math Sociol 15:193–206
Gaito S, Zignani M, Rossi G, Sala A, Wang X, Zheng H, Zhao B (2012) First ACM International Workshop on Hot Topics on Interdisciplinary Social Networks Research. Beijing, China
Gilbert E, Karahalios K (2009) Predicting tie strength with social media. In: CHI '09 Proceedings of the 27th international conference on human factors in, computing systems, pp 211–220.
Girvan M, Newman MEJ (2002) Community structure in social and biological networks. Proc Natl Acad Sci U S A 99:8271–8276
Goh K-I, Barabási A-L (2008) Burstiness and memory in complex systems. Europhys Lett 81:48002
Goh K-I, Kahng B, Kim D (2001) Universal behavior of load distribution in scale-free networks. Phys Rev Lett 87:278701
Goh K-I, Kahng B, Kim D (2002) Fluctuation-driven dynamics of the internet topology. Phys Rev E 88:108701

Golder SA, Wilkinson D, Huberman BA (2007) Rhythms of social interaction: Messaging within a massive online network. In , Steinfield C, Pentland B, Ackerman M, Contractor N (eds) Proceedings of third international conference on communities and technologies , pp 41–66
Gómez-Gardeñes J, Moreno Y (2004) Local versus global knowledge in the Barabasi-Albert scale-free network model. Phys Rev E 69:37103
Gonçalves B, Perra N, Vespignani A (2011) Modeling users' activity on Twitter networks: validation of Dunbar's number. PLoS ONE 6(8):e22656
González M, Hidalgo C, Barabási A-L (2008) Understanding individual human mobility patterns. Nature 453(7196):779–782
Grabowicz P, Ramasco J, Moro E, Pujol J, Eguiluz V (2012) Social features of online networks: the strength of intermediary ties in online social media. PLoS ONE 7(1):e29358
Granovetter M (1973) The strength of weak ties. Am J Sociol 78:1360–1380
Greene J (1997) Production and inventory control handbook, Tercera edn. McGraw-Hill, New York
Guare J (1990) Six degree of separation: a play. Random House, New York
Guimerá R, Danon L, Díaz-Guilera A, Giralt F, Arenas A (2003) Self-similar community structure in organisations. Phys Rev E 68:065103
Guimerá R, Mossa S, Turtschi A, Amaral L (2005) The worldwide air transportation network: anomalous centrality, community structure, and cities' global roles. Proc Natl Acad Sci U S A 102:7794
Guimerá R, Uzzi B, Spiro J, Amaral LA (2005) Team assembly mechanisms determine collaboration network structure and team performance. Science 308(5722):697–702
Guimerá R, Danon L, Díaz-Guilera A, Giralt F, Arenas A (2006) The real communication network behind the formal chart: community structure in organizations. J Econ Behav Organ 61:653–667
Haight FA (1967) Handbook of the poisson distribution. Wiley, New York
Hidalgo C, Rodriguez-Sickert C (2008) The dynamics of a mobile phone network. Phys A 387:3017
Holme P, Kim BJ, Yoon CN, Han SK (2002) Attack vulnerability of complex networks. Phys Rev E 65:056109
Holme P (2002) Edge overload breakdown in evolving networks. Phys Rev E 66:036119
Holme P, Edling C, Liljeros F (2004) Structure and time-evolution of an internet dating community. Soc Netw 26:155–174
Holme P (2005) Network reachability of real-world contact sequences. Phys Rev E 71:046119
Holme P, Saramäki J (2012) Temporal networks. Phys Rep 519:97–125
Huberman BA, Adamic L (1999) Internet-growth dynamics of the World-Wide Web. Nature 401:131
Huberman B, Romero D, Wu F (2009) Social networks that matter: Twitter under the microscope. First Monday 14(1):1–9
Ibarra H (2002) Homophily differential returns: sex differences in network structure and access in an advertising firm. Adm Sci Q 37:422
Iribarren J, Moro E (2009) Impact of human activity patterns on the dynamics of information diffusion. Phys Rev Lett 103:038702
Iribarren J, Moro E (2011a) Affinity paths and information diffusion in social networks. Soc Netw 33:134–142
Iribarren J, Moro E (2011b) Branching dynamics of viral information spreading. Phys Rev E 84:046116
Isella L, Stehlé J, Barrat A, Cattuto C, Pinton J-F, Van den Broeck W (2011) What's in a crowd? Analysis of face-to-face behavioral networks. J Theor Biol 166:166–180
Jeong H, Tombor B, Albert R, Oltvai Z-N, Barabási A-L (2000) The large-scale organization of metabolic networks. Nature 407:651–655
Jo H-H, Karsai M, Kertész J, Kaski K (2012) Circadian pattern and burstiness in mobile phone communication. New J Phys 14:013055
Karsai M, Kivelä M, Pan R, Kaski K, Kertész J, Barabási A-L (2011) Small but slow world: how network topology and burstiness slow down spreading. Phys Rev E 83:025102(R).
Kempe D, Kleinberg J, Kumar A (2002) Connectivity and inference problems for temporal networks. J Comput Syst Sci 64(4):820–842

Kivran-Swaine F, Govindan P, Naaman M (2011) The impact of network structure on breaking ties in online social networks: unfollowing on twitter. In: CHI '11 Proceedings of the 2011 annual conference on human factors in computing systems.

Kleinberg J (2000) Navigation in a small world. Nature 406:845

Kleinberg J (2008) The convergence of social and technological networks. Commun ACM 51(11):66–72

Korte C, Milgram S (1970) Acquaintance linking between white and negro populations: application of the small world problem. J Pers Soc Psychol 15:101–118

Kossinets G, Watts DJ (2006) Empirical analysis of an evolving social network. Science 311:5757

Kossinets G, Kleinberg J, Watts DJ (2008) The structure of information pathways in a social communication network. In: Proceedings of ACM SIGKDD '08, pp 435–443.

Kostakos V (2009) Temporal graphs. Physica A 388(6):1007–1023

Kovanen L, Karsai M, Kaski K, Kertész J, Saramäki J (2011) Temporal motifs in time-dependent networks. J Stat Mech: Theory Exp 11:11

Kumpula J, Onnela J-P, Saramäki J, Kertész J, Kaski K (2009) Model of community emergence in weighted networks. Comp Phys Comm 180:517

Kwak H, Lee C, Park H, Moon S (2010) What is Twitter, a social network or a news media? In: Proceedings of the 19th international conference on, World wide web, pp 591–600.

Lambiotte R, Blondel V, de Kerchove C, Huens E, Prieur C, Smoreda Z, Van Dooren P (2008) Geographical dispersal of mobile communication networks. Phys A 387:5317–5325

Lancichinetti A, Kivelä M, Saramäki J, Fortunato S (2010a) Characterizing the community structure of complex networks. PLoS ONE 5(8):e11976

Lancichinetti A, Radicchi F, Ramasco J, Fortunato S (2010b) Finding statistically significant communities in networks. PLoS ONE 6:e18961

Latora V, Marchiori M (2001) Efficient behavior of small-world networks. Phys Rev Lett 87:198701

Latora V, Marchiori M (2003) Economic small-world behavior in weighted networks. Eur Phys J B 32:249

Lazarsfeld P, Merton R (1954) Friendship as a Social Process: A Substantive and Methodological Analysis. In: Berger M, Abel T, Page CH (eds) Freedom and control in modern society. Van Nostrand, New York, pp 18–66

Lazer D, Pentland A, Adamic L, Aral S, Barabási A-L, Brewer N, Christakis NA, Contractor N, Fowler J, Gutmann M, Jebara T, King G, Macy M, Roy D, Van Alsyne M (2009) Computational social science. Science 323:721–723

Lee C, Scherngell T, Barber MJ (2009) Real-world separation effects in an online social network. Soc Netw 33:2

Lee S-H, Kim P-J, Ahn Y-Y, Jeong H (2010) Googling social interactions: web search engine based social network construction. PLoS ONE 5(7):e11233

Leskovec J, Lang KJ, Dasgupta A, Mahoney M (2009) Community structure in large networks: natural cluster sizes and the absence of large well-defined clusters. Internet Math 61(1):29–123

Lewis K, Kaufman J, Gonzales M, Wimmer A, Christakis NA (2008) Tastes, ties, and time: a new social network dataset using Facebook.com. Soc Netw 30:330–342

Liben-Nowell D, Novak J, Kumar R, Raghavan P, Tomkins A (2005) Geographic routing in social networks. Proc Natl Acad Sci U S A 102:623–628

Malmgren RD, Stouffer DB, Motter AE, Amaral LAN (2008) A poissonian explanation for heavy tails in e-mail communication. Proc Natl Acad Sci U S A 105:18153–18158

Mangan S (2003) Structure and function of the feed-forward loop network motif. Proc Natl Acad Sci U S A 100:11980–11985

Martin JL, Yeung K-T (2006) Persistence of close personal ties over a 12-year period. Soc Netw 28:331–362

McDaid A, Hurley N (2010) Detecting highly overlapping communities with model-based overlapping seed expansion. In: Proceedings of the ASONAM10.

McPherson J, Smith-Lovin L, Cook J (2001) Birds of a feather: homophily in social networks. Ann Rev Sociol 27:415–444

Milgram S (1967) The small world problem. Psychol Today 1:60–67
Milo R, Shen-Orr S, Itzkovitz S, Kashan N, Chklovskii D, Alon U (2002) Network motifs: simple building blocks of complex networks. Science 298:824
Miritello G, Lara R, Cebrián M, Moro E (2012a) Limited communication capacity unveils strategies for human interaction. arXiv:1304.1979.
Miritello G, Moro E, Lara R (2011) Dynamical strength of social ties in information spreading. Phys Rev E 83:045102(R).
Miritello G, Moro E, Lara R, Martinez R, Belchamber J, Roberts S, Dunbar R (2012b) Time as a limited resource: communication strategy in mobile phone networks. Social Networks 35:89–95.
Mislove A, Marcon M, Gummadi K, Druschel P, Bhattacharjee B (2007) Measurement and analysis of online social networks. In: Proceedings of the 7th ACM conference on internet, measurement, pp 29–42.
Mollica K, Gray B, Trevino L (2003) Racial homophily and its persistence in newcomers' social networks. Organ Sci 14:123–136
Moody J (2002) The importance of relationship timing for diffusion. Soc Forces 81:25–56
Moreno Y, Pastor-Satorras R, Vespignani A (2002) Epidemic outbreaks in complex heterogeneous networks. Eur Phys J B 26:521–529
Nanavati AA, Singh R, Chakraborty D, Dasgupta K, Mukherjea S, Das G, Gurumurthy S, Joshi A (2008) Analyzing the structure and evolution of massive telecom graphs. IEEE Trans Knowl Data Eng 20:5
Newman MEJ (2001a) Scientific collaboration networks: I. Network construction and fundamental results. Phys Rev E 64:016131
Newman MEJ (2002b) The spread of epidemic disease on networks. Phys Rev E 66:016128
Newman MEJ, Forrest S, Balthrop J (2002) Email networks and the spread of computer viruses. Phys Rev E 66:035101
Newman MEJ (2002a) Assortative mixing in networks. Phys Rev E 89:208701
Newman MEJ (2003a) Properties of highly clustered networks. Phys Rev E 68:026126
Newman MEJ, Park J (2003) Why social networks are different from other types of networks. Phys Rev E 68:036122
Newman MEJ (2003b) The structure and function of complex networks. SIAM Review 45:167–256
Newman MEJ, Girvan M (2004) Finding and evaluating community structure in networks. Phys Rev E 69:026113
Oliveira J, Barabási A-L (2005) Human dynamics: Darwin and Einstein correspondence patterns. Nature 437:1251
Onnela J, Chakraborti A, Kaski K, Kertész J, Kanto A (2003) Dynamics of market correlations: taxonomy and portfolio analysis. Phys Rev E 68:056110
Onnela J, Saramäki J, Kertész J, Kaski K (2005) Intensity and coherence of motifs in weighted complex networks. Phys Rev E 71:065103
Onnela J-P, Saramäki J, Hyvönen J, Szabó Z, Argollo de Menezes M, Kaski K, Barabási A-L, Kertész J (2007a) Analysis of a large-scale weighted network of one-to-one human communication. New J Phys 9:179
Onnela J-P, Saramäki J, Hyvönen J, Szabó Z, Lazer D, Kaski K, Kertész J, Barabási A-L (2007b) Structure and tie strengths in mobile communication networks. Proc Natl Acad Sci U S A 104:7332
Onnela J-P, Arbesman S, González M, Barabási A-L, Christakis NA (2011) Geographic constraints on social network groups. PLoS ONE 6(4):e16939
Palla G, Barabási A-L, Tamás V (2007a) Community dynamics in social networks. Noise Stochast Complex Syst Finan 6601:660106
Palla G, Barabási A-L, Vicsek T (2007b) Quantifying social group evolution. Nature 446:664–667
Pan R, Saramäki J (2011) Path lengths, correlations, and centrality in temporal networks. Phys Rev E 84:016105
Pastor-Satorras R, Vespignani A (2001b) Epidemic spreading in scale-free networks. Phys Rev Lett 86:3200–3203
Pastor-Satorras R, Vespignani A (2002) Immunization of complex networks. Phys Rev E 65:036104

Pastor-Satorras R, Vespignani A (2004) Evolution and structure of the internet: a statistical physics approach. Cambridge University Press, Cambridge
Quattrocchi W, Conte R, Lodi E (2010) Simulating opinion dynamics in heterogeneous communication systems. In: Proceedings of 7th european conference on complex systems (ECCS), pp 70–85.
Raeder T, Lizardo O, Chawla N, Hachen D (2011) Predictors of short-term decay of cell phone contacts in a large scale communication network. Soc Netw 33:245–257
Raghavan U, Albert R, Kumara S (2007) Near linear time algorithm to detect community structures in large-scale networks. Phys Rev E 76:036106
Reynolds P (2003) Call center staffing. The Call Center School Press, London
Roberts S (2010) Constraints on social networks. Proc Br Acad 158:115–134
Rocha L, Liljeros F, Holme P (2010) Information dynamics shape the sexual networks of internet-mediated prostitution. Proc Natl Acad Sci U S A 107:5706–5711
Romero D, Meeder B, Barash V, Kleinberg J (2011) Maintaining ties on social media sites: the competing effect of balance, exchange and betweenness. In: Proceedings of the ICWSM'11, fifth international AAAI conference on weblogs and social media.
Rosvall M, Bergstrom C (2008) Maps of information flow reveal community structure in complex networks. Proc Natl Acad Sci U S A 105:1118–1123
Rybski D., Buldyrev SV, Havlin S, Liljeros F, Makse HA (2010) Communication activity: temporal correlations, clustering, and, growth. arXiv:1002.0216v1.
Rybski D, Buldyrev SV, Havlin S, Liljeros F, Makse HA (2009) Scaling laws of human interaction activity. Proc Natl Acad Sci U S A 106:12640
Saramäki J, Kivelä M, Onnela J-P, Kaski K, Kertész J (2007) Generalizations of the clustering coefficient to weighted complex networks. Phys Rev E 75:027105
Sarkar P, Moore A (2005) Dynamic social network analysis using latent space models, Special edn. Link Mining, In SIGKDD Explorations
Scellato S, Noulas A, Lambiotte R, Mascolo C (2011) Socio-spatial properties of online location-based social networks. In: ICWSM'11, Fifth international AAAI conference on weblogs and social media.
Scott J (2000) Social network analysis: a handbook. Sage Publications, London
Snijders, Tom AB, The Statistical Evaluation of Social Network Dynamics. pp. 361–395 in Sociological Methodology (2001), edited by M.E. Sobel and M.P. Becker. Boston and London: Basil Blackwell.
Solé RV, Pastor-Satorras R, Smith E, Kepler TB (2002) A model of large-scale proteome evolution. Adv Complex Syst 5:43
Sporns O (2003) Network analysis, complexity, and brain function. Complexity 8:56
Strogatz SH (2001) Exploring complex networks. Nature 410:268–276
Szabó G, Barabási A-L (2006) Network effects in service usage.arXiv:physics/0611177.
Tang J, Musolesi M, Mascolo C, Latora V (2010) Characterising temporal distance and reachability in mobile and online social networks. ACM SIGCOMM Comput Commun Rev 40:1
Tang J, Musolesi M, Mascolo C, Latora V (2009) Temporal distance metrics for social network analysis. In: Proceedings of the 2nd ACM SIGCOMM workshop on online social networks (WOSN'09), Barcelona, Spain.
Tang W, Zhuang H, Tang J (2011) Learning to infer social ties in large networks. In: Proc. ECML PKDD'11. Proceedings of the 2011 european conference on machine learning and knowledge discovery in databases, vol Part III. pp 381–397.
Tantipathananand C, Berger-Wolf T, Kempe D (2007) A framework for community identification in dynamic social networks. In: Proc. KDD '07 Proceedings of the 13th ACM SIGKDD international conference on knowledge discovery and data mining, pp 717–726.
Toivonen R, Onnela J-P, Saramäki J, Hyvönen J, Kaski K (2006) A model for social networks. Physica A 371:851

Toivonen R, Kumpula J, Saramäki J, Onnela J-P, Kertész J, Kaski K (2007) The role of edge weights in social networks: modelling structure and dynamics. In: Proceedings of SPIE: noise and stochastics in complex systems and finance, Proc. of SPIE 6601, 660110.
Ugander J, Karrer B, Backstrom L, Marlow C (2011) The anatomy of the facebook social graph. preprint:arXiv:1111.4503v1.
Valverde S, Solé RV (2005) Network motifs in computational graphs: a case study in software architecture. Phys Rev E 72:026107
Vázquez A, Flammini A, Maritan A, Vespignani A (2003) Modeling of protein interaction networks. ComPlexUs 1:38
Vázquez A, Dobrin R, Sergi D, Eckmann J-P, Oltvai Z-N, Barabási A-L (2004) The topological relationship between the large-scale attributes and local interaction patterns of complex networks. Proc Natl Acad Sci U S A 101:17940–17945
Vázquez A, Oliveira J, Dezsö Z, Goh K-I, Kondor I, Barabási A-L (2006) Modeling bursts and heavy tails in human dynamics. Phys Rev E 73:036127
Vázquez A, Rácz B, Lukács A, Barabási A-L (2007) Impact of non-Poissonian activity patterns on spreading processes. Phys Rev Lett 98:158702
Volkovich Y, Scellato S, Laniado D, Mascolo C, Kaltenbrunner A (2012) The length of bridge ties: structural and geographic properties of online social interactions. Proceedings of the sixth international AAAI conference on weblogs and social media, In
Waleaevens J, Demoor T, Maertens T, Bruneel H (2012) Stochastic queuing-theory approach to human dynamics. Phys Rev E 85:021139
Wasserman S, Faust K (1994) Social networks analysis. Cambridge University Press, Cambridge
Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. Nature 393:440–442
Watts DJ (1999) Small worlds: the dynamics of networks between order and randomness. Princeton University Press, Princeton, NJ
Watts DJ (2004) The "new" science of networks. Annu Rev Sociol 30:243–270
Wellman B, Wortley S (1990) Different strokes from different folks: community ties and social support. Am J Sociol 96:558–588
Wellman B, Salaff J, Dimitrova L, Garton L, Gulia M, Haythornthwaite C (1996) Computer networks as social networks: collaborative work, telework, and virtual community. Annu Rev Sociol 22:213
Wellman B (2001) Computer networks as social networks. Science 293:2031
Wellman B (2007) The network is personal: introduction to a special issue of social networks. Soc Netw 29(3):349–356
Wellman B, Haythornthwaite C (2003) The internet in everyday life. Blackwell, Oxford
West D (1995) Introduction to graph theory. Prentice-Hall, Englewood Cliffs, NJ
Wilkinson D, Huberman B (2004) A method for finding communities of related genes. Proc Natl Acad Sci U S A 101:5241–5248
Wu T, Zhoud C, Xiaob J, Kurthsa J, Schellnhubera H (2010) Evidence for a bimodal distribution in human communication. Proc Natl Acad Sci U S A 107:18803
Wuchty S (2009) What is a social tie? Proc Natl Acad Sci U S A 106:15099–15100
Wuchty S, Uzzi B (2011) Human communication dynamics in digital footsteps: a study of the agreement between self-reported ties and email networks. PLoS ONE 6(11):e26972
Yook S, Jeong J, Barabási A-L, Tu Y (2001) Weighted evolving networks. Phys Rev Lett 86:5835
Zhao Q, Oliver N (2010) Communication motifs: a novel approach to characterize mobile communications. In: NetMob2010.

http://www.springer.com/978-3-319-00109-8

Temporal Patterns of Communication in Social Networks
Miritello, G.
2013, XIV, 153 p., Hardcover
ISBN: 978-3-319-00109-8