Inf Syst Front (2019) 21:229–240
DOI 10.1007/s10796-017-9740-8

Dynamic network analysis of online interactive platform

Mehmet N. Aydin[1] **· N. Ziya Perdahci**[2]

Published online: 15 February 2017

Abstract The widespread use of online interactive platforms including social networking applications, community support applications draw the attention of academics and businesses. The basic trust of this research is that the very nature of these platforms can be best described as a network of entangled interactions. We agree with scholars that these platforms and features necessitate the call for theory of network as a novel approach to better understand their underpinnings. We examine one of the leading online interactive health networks in Europe. We demonstrate that the interactive platform examined exhibits essential structural properties that characterize most real networks. In particular, we focus on the largest connected component, so-called a giant component (GC), to better understand network formation. Dynamic network analysis helps us to observe how the GC has evolved over time and to identify a particular pattern towards emerging a GC. We suggest that the network measures examined for the platform should be considered as novel and complementary metrics to those used in conventional web and social analytics. We argue that various stages of GC development can be a promising indicator of the strength and endurance of the interactions on the platform. Platform managers may take into account basic stages of the emergence of the GC to determine varying degrees of product attractiveness.

✉ Mehmet N. Aydin
mehmet.aydin@khas.edu.tr

1 Faculty of Engineering and Natural Sciences, Kadir Has University, Central Campus, Kadir Has Caddesi, Cibali, 34083 İstanbul, Turkey

2 Department of Informatics, Mimar Sinan Fine Arts University, Bomonti Campus, Silahşör Caddesi No:89, İstanbul 34380, Turkey

Keywords Online interactive platform · Theory of network · Social network analysis

1 Introduction

With the advance of ubiquities computing and promising information services, online human interactions have been a centre of attention for both research communities and business endeavours. To facilitate human interactions, online interactive platforms are promoted as, for instance, community support applications, social networking applications, micro blogs, mobile message exchanges. This research is concerned with practical and research challenges related to understanding the nature of interactive aspect of these platforms. Various features of platforms (e.g. information/advice seeking, message exchange, relationship formation) reflect this interactive aspect, but are subject to rules and other constraints in a given context (e.g., friendship, commerce, health domains).

Regarding practical challenges, often times managers face with uncertainties and complexities in managing these platforms since they are yet to be mature regarding sustainable business models, multifaceted services, a sheer number of involved interactions (Butler 2001; Jones et al. 2004; Kane et al. 2014). Conventional key performance indicators and web analytics measures are inadequate in truly reflecting the very nature of these platforms, which is a network of entangled interactions.

Recently, scholars in IS research have acknowledged the call for studying online interactive platforms from the perspective of theory of network (Johnson et al. 2014; Kane et al. 2014) or network science (Strogatz 2001; Barabási 2009; Wang et al. 2012). That is, the very notion of interactive behaviour on the platform requires further understanding of underlying network structure and dynamics. Simply, structure refers to a number of essential network characteristics (e.g. degree distribution, giant component) that characterize the network (e.g. random versus real networks) and dynamics is concerned with time evolution of the network with respect to structural characteristics. Dynamic network analysis is aimed to examine both structural and dynamic aspects of interactive platforms.

Of particular importance among structural characteristics is the notion of connected components, which indicates how strong and sustainable interactions are established on platforms. The largest of these components is central to this study as it plays a dominating role for both network development and platform management. This research is aimed to contribute to further understanding of establishment and sustainability of interactions on platforms. The motivating questions are as follows: What network measures are appropriate to examine an overall structure of the network underpinning a platform? How is the giant component established and evolved over time? Is there any peculiarity of giant development characterizing the network? Does the network reach a steady state? If so, what underpins this state? What research and practice implications can be discerned from this dynamic network analysis? In this regard, this research aims to contribute to both theory of networks and information systems research, but the key motivation is to advance an understanding and value of theory of network for Information System research. The practice related contribution to conduct dynamic network analysis is to associate the network findings with the business insights that may of concern to platform managers and decision makers or alike.

To proceed in answering these questions, we examine one of the leading online interactive health networks (www.doktorsitesi.com) in Europe. According to their website, Doktorsitesi has around two million registered members (presumably advice seeking patients) and 18,000 physicians. As shall be explained later on, the platform examined provides physicians and patients alike with special social networking features such as private messaging, tying, exchanging information via questions and answers, and other user-led content related services (twitter-like following, articles, videos). In this research we focus on the tie feature. The goal of the feature is to connect the health specialists who provide valuable health information to those who seek it through a private channel. This feature lets users to make request for a tie and/or to approve of a tie. In case a request is approved, it becomes a tie, which signifies a bilateral agreement with a reciprocal right to send an unlimited number of messages privately.

By conducting dynamic network analysis (DNA) we are able to surface intriguing interactions between physicians and patients alike on a macroscopic scale. To better examine tie establishment we make use of essential properties that characterize most real networks. We should note that the structures may reflect patterns of interactions that can be essential to explain how people learn, form opinions, and affect the others in real world social networks (Heidemann et al. 2012; Granovetter 1973).

2 Research background

Theory of network can be associated with various theoretical accounts including social network theory (Milgram 1967), network science (Barabási 2009), graph theory (Wasserman 1994). Theory of network helps in the examination of complex communication and interaction patterns in various contexts such as health information platforms (Li et al. 2008; van Mierlo 2014). As the academic endeavours contributing to a trans-discipline development of so-called network science, one can explore structural and dynamic (evolutionary) aspects of a network (that is, things and their relations) from appropriate perspectives including social, management and information systems. The impact of network theory on our understanding of real world complex systems is manifold. Thanks to graph theory, we can represent a complex network of interactions by a simple model. And it is thanks to the predictive power of network models we can study diverse phenomena and make predictions (Barabási 2009).

Several network models have been introduced to capture the evolution of real-world networks (e.g. economic networks, biological networks, social networks). The preferential attachment model of (Barabási and Albert 1999) is among the most studied of these models. The model requires two main constituents, growth and preferential attachment, to account for the emergence of heavy-tailed degree distributions in which the emergence of a giant component is automatic.

Although real-world complex networks look different microscopically, researchers discerned many commonalities in their characteristics on a macroscopic scale. Empirical results show that real-world networks are notable for several structural properties (Kossinets and Watts 2006). First, a high proportion of moderately tied nodes coexist with a small proportion of inordinately tied nodes, i.e. the connectivity between nodes follows highly skewed degree distributions compared to random networks (Simon 1955). Second, there exist short paths in these networks between most of the nodes, resulting in an overall low average path length compared to network size. This is widely recognised as the "small-world phenomenon" (Milgram 1967; Watts and Strogatz 1998). Third, real-world networks exhibit high clustering of connectivity compared to density of ties. Other commonly observed structural properties include one or few clusters of components.

One of the most studied phenomenon in the theory of random graphs (Molloy and Reed 1998) and theory of networks (Kumar et al. 2010; Strogatz 2001) is emergence of a uniquely large network component. We refer to a connected component in a network as a cluster of nodes where there exist pathways between any two nodes in the cluster. For directed networks, if every node within a component is connected to every other node bi-directionally it is called a strongly connected component (SCC), otherwise it is called a weakly connected component (WCC) (Broder et al. 2000). Since our analysis excludes WCCs, all connected components are SCCs, and we will simply call them components.

Theoretically the sudden emergence of "giant" component is almost inevitable. Not surprisingly, much attention has been paid to the structure and dynamics of this component in empirical analysis of evolving systems ranging from co-authorship network of scientists to large online social networks (Kumar et al. 2010; Newman et al. 2002; Barabâsi et al. 2002; Leskovec et al. 2007; Kossinets and Watts 2006). Newman et al. (2002) emphasize the importance of giant component; it is that "group of connected vertices that fills a significant portion of the whole network and whose size scales up with the size of the whole network" (p.2568).

We have investigated emergence and evolution of the giant component (GC) of the tie network by recording its size and the rest of its relevant metrics as a function of time. We consider that a component is called GC if it satisfies two conditions: (i) the proportion of nodes it has to be above 50 % of the whole network, (ii) even the size of the second-largest component must never exceed a small amount per mil (that is, around half a per cent).

Besides GC we shall consider additional basic network measures. One of the basic measures is the average degree of nodes, which is simply arithmetic means of in-degree (i.e. making a tying wish), out-degree (i.e. approving of a tying wish), and maximum tying wishes.

Another important characteristic related to the global structure of a network is the average path length. Path length measures the distance between people in terms of number of connections in the network examined. The eccentricity of a node gives the shortest path i.e. the geodesic distance between any two nodes. The maximum eccentricity, called diameter of a network, is the largest distance of all, or the distance between people that are farthest from each other. The minimum eccentricity between people on the other hand is called the radius of the network. The small-world effect is present if the average of the geodesics between every node is around six (Kleinberg 2000).

3 Research context and method

3.1 Online interactive health platform

Interactive health platforms have gained popularity in health domain as they bring health information seekers (patients and alike) and medication advice providers (physicians and other relevant actors) together. Recent reports concerning the adoption of online health applications have shown people's and organizations' significant interest in them (Moss and Elias 2010). Of particular importance among these applications is health information and advice-seeking supporting applications (e.g. WebMD, Healthline) having a direct link to social network sites whereby information support is empowered by human relations or vice versa. This is not surprising especially for the health domain, since peers' opinions for medical practitioners and patients' experience for "like-minded others" are found to be valuable for health related decision-making (Bosslet et al. 2011). Thus, one needs to find out if and how human interactions are established on the platforms. Thanks to emerging online health interactive platforms (e.g. HealthTap, WebMD, Doktorsitesi), which help in providing relevant data for the analysis of information and social networks.

We should also note that on the business side, the decision makers (e.g. platform managers, owners, platform architects) may not be aware of the fact that they are "managing a complex system" of user interactions. Managers, not knowing they are dealing with a complex system, make heuristic decisions, may even terminate a service before it attains its maturity. Changing platform rules or user interfaces frequently can affect the user adoption and underlying network interactions (Fetterman 2009). All these pragmatic changes on platforms make it very difficult for researchers to obtain a data set that one can perform a sound empirical study. Nevertheless, while conducting this empirical study we model the underlying network carefully and aim to surface the global structure of the network.

We examine Doktorsitesi.com, which is one of the leading online interactive health networks (www.doktorsitesi.com) in Europe. Doktorsitesi.com allows its members to access the true identities of physicians but not patients or alike. During the examination of the platform physicians and patients are only granted access to patients' anonymous user profiles that consist of their age, gender, and education. The tying feature is a channel that helps patients connect with others, thereby allowing them to communicate anonymously through private messaging. The feature makes a member's ties list private.

We should note that the platform provides services in a country where digital reputation of physicians via rating or other means is against the law. Due to national regulations in the health sector, physicians are allowed to work full-time or part-time in different types of hospitals (state, university, private) or clinics.

3.2 Method

As an alternative to conventional methods for analysing online interactive platforms (that is, survey and observation for data collection (Chambers et al. 2012)), we make use of valuable data generated on a health information platform and conduct dynamic network analysis for our data set. We obtained a raw data set describing the tying feature of Doktorsitesi.com. The set is composed of activities of members who made use of the feature; either to make a request for a tie or to approve of a tie, collectively called tying wishes, over the 66-month period from January 2009 to July 2014.

For each of these members, we have a log of their tying wishes where user identifiers of both parties and the time stamp of a record constitute a transaction. We do not have any knowledge about the true identities of the members except that they are either medical practitioners or (presumably) patients. We believe that our data set is not influenced by subjective biases on the part of the platform members. That is, unlike traditional surveying which unavoidably introduces biased opinions (Marsden 1990), our approach can be considered as an effective method for probing the user engagement in adopting a service of an online platform.

Raw data was not readily available in the form of a temporal graph since a traditional relational database system is used to record tying transactions on the platform. Therefore, the first step of data preparation before starting dynamic network analysis was to transform records into a network data set. This is a very delicate and mostly iterative rather than a sequential process. After selecting the records relevant to the research focus, we imported the data and checked them for possible artefacts due to bugs in the software. We have detected several artefacts, self-ties, i.e. records of approved tying attempts of individuals to themselves, to name one, and cleaned them.

Description of the network data and visual analysis of network diagrams are produced with Gephi (Bastian et al. 2009), which is a visualization and exploration platform for network analysis. It is open-source and free software.

We model tying wishes and the resulting ties as a directed network. Our approach permits us to articulate both the members who initiate tying wishes and the resulting ties. In the present work, we focus on the tying interactions that all resulted in ties between members. Thus we have eliminated the tying wishes that have not resulted in a tie. Time stamps identify the creation of each tie, that is, when a tying approval occurs, giving the date and time of day. Since we consider the time stamp as a temporal parameter that gives us information on the time evolution of the tie network, we only take into account the date of each record, not its time. That is, we created a dynamic data set with the date of each record that helps us to explore the temporal evolution of the tie network. From this data set, we take several time snapshots, for which we visualize the network to determine if and when a giant component emerges and how it evolves thereon yearly. We also study the number of nodes, the number of edges, the number of components for the overall network, and two of its largest components for time snapshots monthly.

A graph of directed edges models the tie network (see Fig. 1). On the graph (as shall be seen in Figs. 2, 3 and 5), black nodes represent patients (abbreviated to P) and grey nodes represent physicians (abbreviated to Dr). Black directed edges of the tie network represent the tying wishes. Grey directed edges, on the other hand, represent the tying approvals. Our colouring convention is just for the sake of obtaining decent network visualization (X. Chen and Yang 2010). However, from the network science point of view such sophistication of network representation would mean the possibility of measuring an endless number of statistics.

Since in our data set every node is connected to every other bi-directionally all of the social network analysis metrics are calculated by making the network undirected where each pair of directed edges of tying wishes corresponds to a single undirected tie edge.

4 Findings

4.1 Basic network measures

The tie network focuses on those tying requests that are reciprocated. We present basic characteristics of the overall tie network as of June 2014 in Table 1. To better visualize components we provide network graphs by ForceAtlas 2 layout algorithm (see Figs. 2, 3, and 4). This algorithm provides a visual representation that brings out the giant component, placing it at center.

The tie network of Fig. 2 has $N = 11{,}559$ platform members as nodes and $L = 13{,}750$ ties (half of the directed edges) between them. In consequence the average degree of the tie network is 2.4, indicating that a common member interacts with two to three other members. It is thanks to highly skewed degree distribution with a small number of high-degree nodes, the majority of members have engaged in only one tie. To be more specific, 72% members have engaged in only one tie, 12% of them in two ties, and 2.8% of the tie network have more than ten ties. In this network of tying interactions, 71% of all nodes and 86% of all tie edges belong to the largest component, which will turn out to be a giant component (GC) (the central part of Fig. 2).

The average path length between all pair of nodes within the tie network is 5.72 and maximal distances vary between graph diameter of 17 and graph radius of 1. The average path length of the network is 5.72. This means that members are "less than six people apart" should they wish to communicate with each other.

Regarding GC, the increase in the average degree from 2.4 to 2.9 (see Table 1) is expected because interactions are dense compared to the overall network. It is no coincidence that average path length and diameter of the overall network and GC are almost the same due to the fact the latter GC dominates the tying interactions. The minimum eccentricity (that is, radius) of 1 of the overall network stems from large numbers of small-sized components. In contrast to these small-sized components, within the GC of 8239 nodes, maximal geodesic distances between nodes vary between radius of 9 and diameter of 17 with a mean distance of 5.73 nodes. The GC is a small-world network where nodes are on the average about six handshakes away from each other.

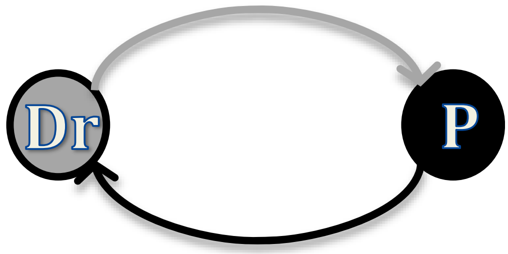

Fig. 1 Network model description of tying interactions: Grey nodes represent physicians, black nodes represent patients or alike. Directed edges between nodes represent tying wishes (black) and tying approvals (grey). A pair of directed edges between members indicates that they are tied. The figure illustrates one tying example.

For a more detailed network analysis of the online service of tying feature regarding the interaction patterns among physicians and patients, the interested reader may refer to (Aydin and Perdahci 2016).

4.2 Emergence of giant component and its evolution

Network visualization helps us to observe how the GC is established, grows and evolves through the timeline. We can point out few significant observable facts in the visualizations provided. Figure 3 includes six graphs, illustrating how the tie network has shaped since its inspection near the beginning of the year 2009 while Table 2 lists key network measures for each snapshot.

At the beginning (Fig. 3a), the network is quite fragmented; with a total of 105 components (Table 2). There appear three large components at the centre, none of which can be classified as a GC. These components comprise (to the nearest integer value) 25%, 7%, and 4% of the network respectively. Figure 3b is the snapshot taken two months later where we now recognize only one large component at the centre. This component is comprised of 53% of the network; it is already bigger than the totality of the other 166 components of the time. We recognize this as a yet-to-be born GC, because most of the little fragments account for some fraction of a per cent of the network (see Table 2).

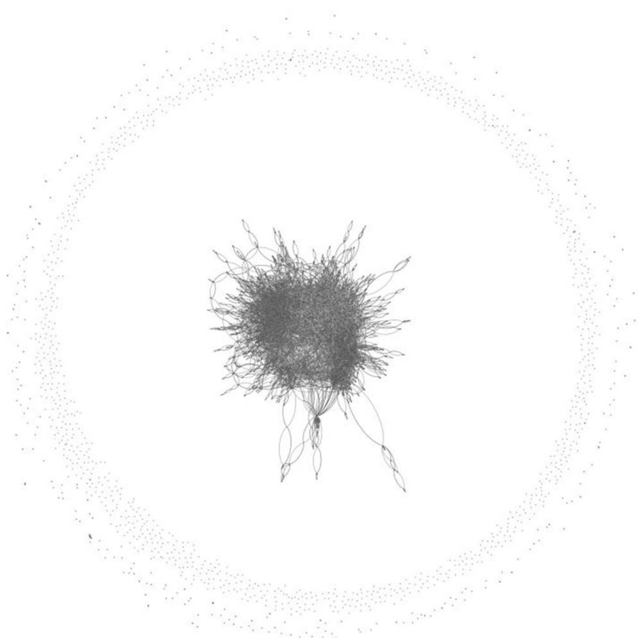

Fig. 2 Graph of the overall tie network as of June 2014

Nevertheless, the progression of the network apparently follows a particular pattern towards making a GC. With hindsight, we continue exploring the evolution of the network by taking more snapshots and recognize the emergence of the sought-after GC nearly 17 months after the inspection of the feature (see Fig. 3c). At this point, the GC comprises 60% of nodes and 75% of ties, and even the largest fragment of the rest of the 821 components accounts for just half a per cent of the whole network. Subsequent snapshots taken a year apart reveal that the GC enters what can be called as the "giant regime" state, where it spends four more years till the summer of 2014 as a stable GC. Of course, this state of affairs is a sign of stability, as the network is like an open system that grows out of both interactions between existing and newly added nodes.

With the idea that this phase might indicate a second wave of GC formation we investigated the year 2012 as a separate case (see Fig. 4). The network metrics (see Table 1) suggest that the interactions get stronger and people get more closely tied to each other with a higher average degree per node, smaller average path length, diameter, and radius. The largest component of the period is eligible to be called a GC as it accounts for 77% of the nodes and 90% of the edges (see the centre of Fig. 4) while the largest of the rest of 212 components is only 8 per mil (0.8%).

To better understand the transition from small-sized network components to GC, we have taken snapshots of the largest component for about a week apart and recorded the proportion of nodes and edges it absorbs for the first three months. Thus, Fig. 5a shows proportion of nodes and edges accumulated by the largest component across time. As Fig. 5a illustrates, the largest component exhibits a rapid growth from 10% to 50% per cent of nodes and edges within the first three months.

One can easily notice that the yet-to-be born giant signifies the beginning of the second stage of the evolution, which involves a spike followed by one year of fluctuations before entering to giant regime at around middle of 2010. That means, it has taken about one-and-half year for the giant to reach its mature state.

Figure 5b shows the total number of new nodes added to the tie network, which is of particular interest to examine the network evolution further. We have recorded both the total number of new nodes that enter into the network and the number of new nodes that enter directly into the GC every month as shown in Fig. 5b. Notice that before the inception of the GC at around the middle of the year 2010, there is a

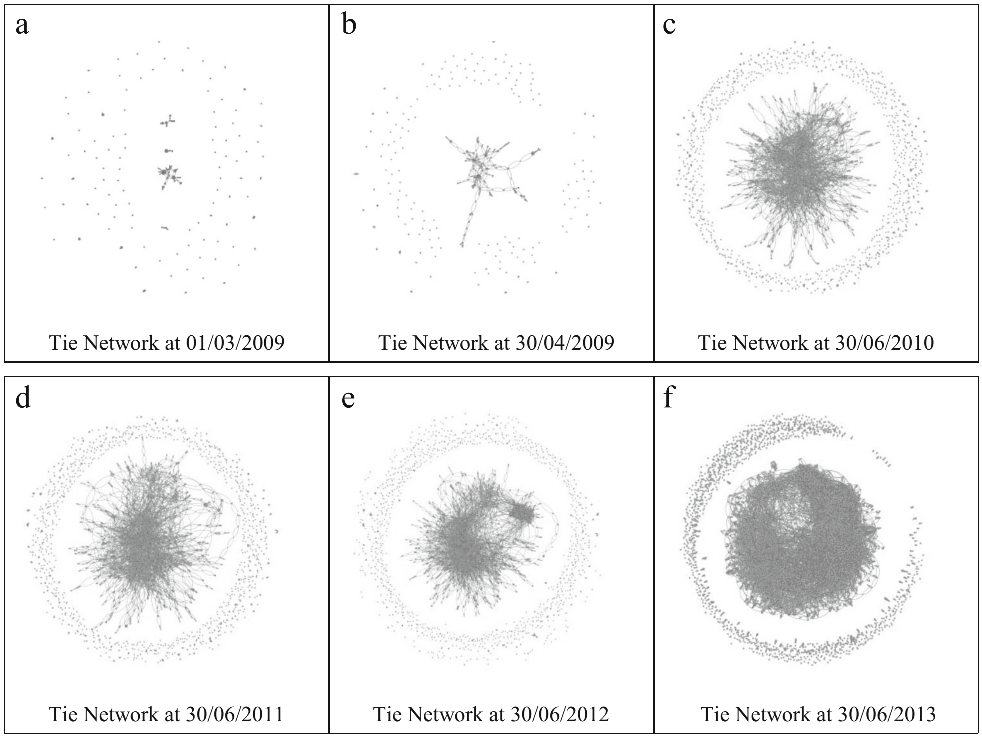

Fig. 3 Evolution of the Tie Network and GC across timeline

significant difference in the number of nodes that enter into the network and the GC. However, GC has absorbed a major proportion of nodes since its inspection. As one can see in Fig. 5b, occasionally the GC absorbs more than the total number of newly added nodes. This happens because both existing outsider nodes (that is, nodes lying in the outer rim) and newly added nodes can contribute to the growth of GC when some newly entered nodes establish ties that cause small-sized components to join the GC.

Several observations on Fig. 5b compelled us to make further investigations. First, there seems to be no tying activity during a large portion of the summer of 2012. This was because of the unavailability of the tie feature during information technology migration. Second, in the last quarter of 2013 there seems to be noticeable deviations from the normal state of affairs that prevail since the inception of the GC. As it happens, the platform owners changed the user interface during the last quarter of 2013, and provided the platform members with a new feature that lets them to view who ties to whom so long as they are tied since the second quarter of 2014.

As one can see that Fig. 5b signifies an erratic pattern. This raises the issue of what underpins this pattern. Given that the network is constituted by interactions between two types of nodes (physicians and patients), we show the ratio of the number of patients to the number of physicians joining GC in Fig. 5c. Physicians and patients who start using the tie feature do not seem to follow a particular pattern.

Yet another macroscopic parameter of interest by which we can describe the network evolution is the ratio of the number of patients to the number of physicians. We study the patients-to-physicians ratio by taking time snapshots monthly (see Fig. 5c). Unlike Fig. 5b that ratio follows a regular pattern. After an initial regime of about one and a half years, the ratio stabilizes in an almost steady state with an approximately constant value of four patients to physicians. Remarkably, the ratio stabilizes at about the same time as we consider being GC's inception in the middle of 2010.

These results are encouraging enough to merit further discussion.

5 Discussion

Recently, as a new type of information systems, we have witnessed a sheer number of online interactive platforms that provide their users with special networking and interpersonal communication features. The widespread use and economic impact of these platforms signifies them as a research and industry phenomenon. We agree with scholars that these platforms and features necessitate the call for theory of network as a novel approach to better understand their underpinnings (Johnson et al. 2014; Kane et al. 2014). This paper contributes to further understanding of these platforms from the perspective of theory of network and demonstrates the value and appropriateness of a number of network measures examined. As such, the present study outlines a number of essential network measures and

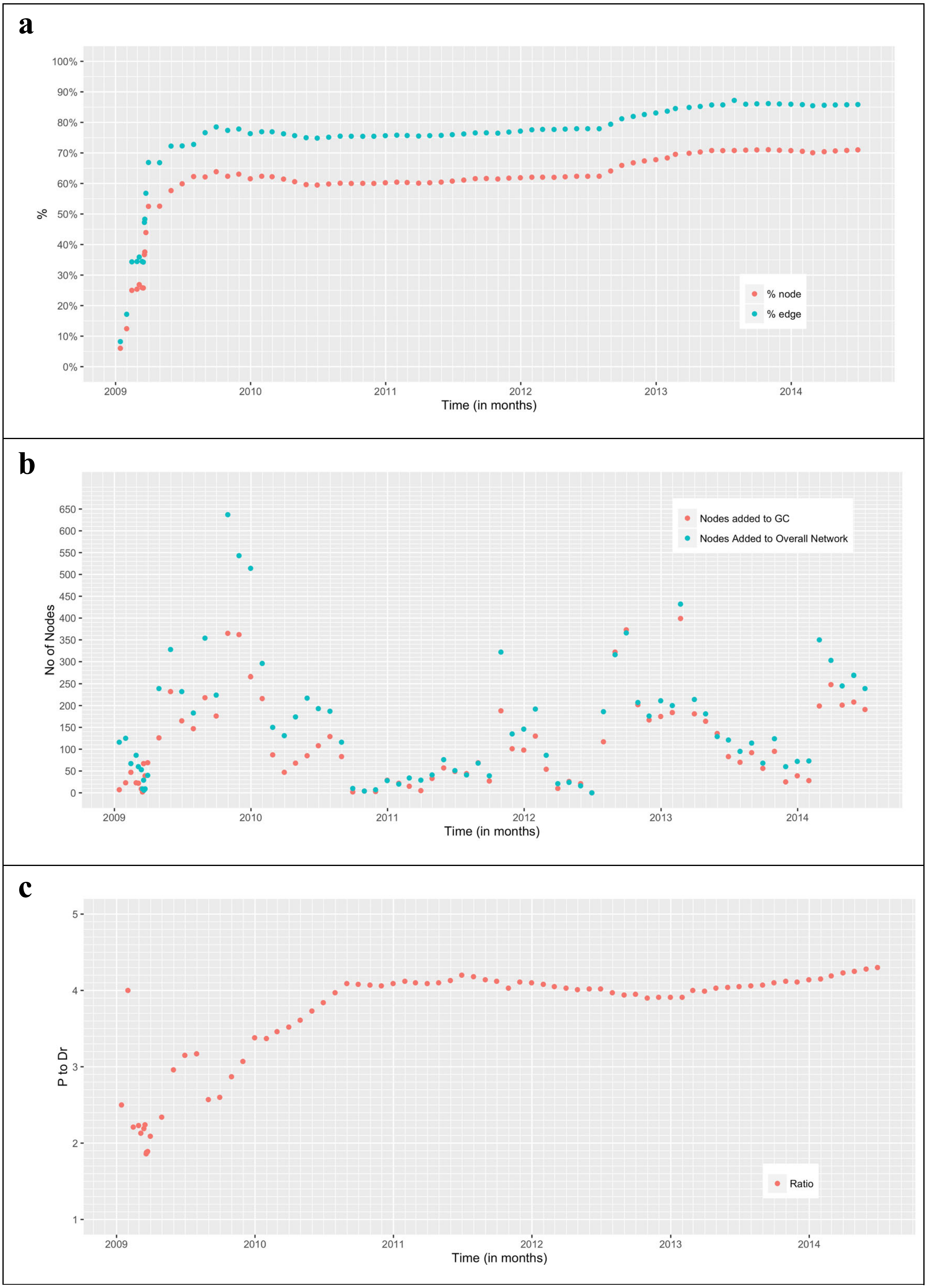

Fig. 5 **a** Monthly percentages of nodes and edges in the largest component in the period January 2009 to June 2014. Fig. 5**b** Monthly number of new nodes added to the tie network and to the largest component in the period January 2009 to June 2014. Fig. 5**c** Monthly ratio of the number of patients to the number of physicians joining GC in the period January 2009 to June 2014

Table 1 Summary of basic measures for the overall tie network and Giant Component (GC)

	Nodes	Directed Edges	Average Degree	Average Path Length	Diameter	Radius
Overall	11,559	27,500	2.4	5.72	17	1
GC	8209	23,606	2.9	5.73	17	9
GC of 2012	1652	5146	3.1	4.30	13	7

emphasizes importance of a giant component for these platforms. That is, an examination of giant component establishment deserves a special attention from theory and practice point of view and thus can be of a great interest to academics and practitioners.

5.1 Implications for theory of networks

The tie feature of Doktorsitesi.com is just one example of ever increasing online services that give rise to an obscure complex structure. We believe that a directed graph is the right representation (see Fig. 1) one could employ to capture the key properties of these services. Choosing the right representation is not a trivial task; ambiguous network models such as the one proposed (Chau and Xu 2012) should be carefully articulated and contextualized within the frame of theory of network (Strogatz 2001; Kim and Wilhelm 2008).

Giant components are not exclusive to real-world complex networks. Even the simplest random network model predicts their existence under certain conditions (Newman et al. 2002). Yet, giant components of random networks are fundamentally different from their real-world counterparts, in that as random networks that have comparable average degrees to real-world networks evolve, the giant component consumes all the other components until there remains only one component in the network (Barrat et al. 2008). This is in stark contrast to our empirical case; the proportion of new nodes entering into the GC follows a stationary trend in a narrow band between sixty and 70 % (see Fig. 4a), as people do not create ties at random but consciously. We believe that this narrow band of fluctuations reflects a fundamental phenomenon of the underlying tying interactions.

Both empirical and theoretical studies of the evolution of such complex systems are still few. One pioneering work studies the dynamics of component formation for two online social networks (Kumar et al. 2010). Similar to our findings, they state that the fraction of nodes in the giant component remains almost constant once a steady state is reached (Kumar, pp.614), although this observation is not so obvious on the figures presented. One other work (Leskovec et al. 2007) introduces two empirical observations: (i) growing networks become denser over time, (ii) the network diameter decreases as a network grows. These empirical observations are also in agreement with our observations as seen in Table 1, the tie network of 2012, and in Fig. 4a the densification of edges of the GC in later years.

The ratio of patients to physicians remains almost constant in a narrow band around four once the tie network enters into the giant regime. It is a remarkable phenomenon that the tying interactions of patients and physicians coalesce into a giant component with an accompanying approximately constant ratio. To the best of our knowledge, no research group has ever had the opportunity to make an empirical study of online interactive platforms of similar nature. On the one hand, we contend that the existence of this ratio is one of the characteristics of the tie network. On the other hand, it would not be unexpected to see the presence of almost constant quantities in real-world evolving networks. Thus, the relevant research question in search of these could be: Which network measure or a combination of network measures remain approximately constant as a network evolves?

Kumar et al. resort to a biased preferential attachment to explain the network evolution, while Leskovec et al. (2007) propose the Community Guided Attachment and the Forest Fire models to describe the empirical findings of evolving networks. (Ma et al. 2012) propose yet another network growth model specifically for software engineering domain. We contend that the descriptions that these mechanistic models provide would be too restrictive to simulate the interactions of the tie network, because various complications add to complexity (Strogatz 2001). One such complication is the node diversity: basically there are two kinds of nodes, anonymous registered members and physicians, who discover the site via different channels. It is found that most of the members reach online interactive platforms through organic (non-sponsored) Web search results (Ghose and Yang 2009). We suspect the same is applied to our case. On the other hand, physicians may become aware of the platform easily through professional channels. According to a recent study (Johnson et al. 2014) a multiple mechanism model ranging from theory of homophily to preferential attachment should have a better chance of simulating the role of social mechanisms in tie establishments on online communities. However, even a multiple mechanism model cannot fully capture the tying mechanisms as the model ignores the node diversity. As the number of empirical studies of evolving real-world complex networks within the context of management information systems increases, theory of network will have the opportunity to produce models that could describe them.

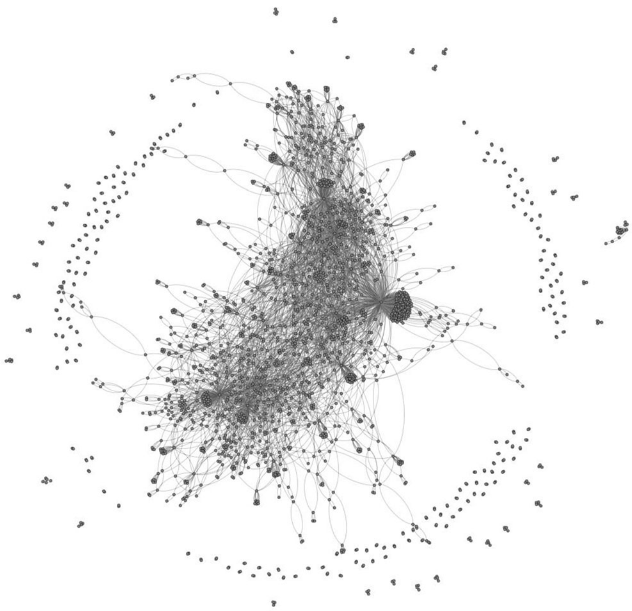

Fig. 4 The tie network for the period 01/01/2012–31/12/2012

5.2 Implications for business insights

The value of network measures for platforms would be of interest to actors including users, managers, designers, and sponsors. One promising venue for utilizing these measures is platform related business analytics aiming to turn data into insight and insight into decision (Sharma et al. 2014). As these platforms operate in an ever-changing market, platform owners and managers may use network measures as complementary analytics for decision-making (Table 3).

We should note that the network measures examined for the platform are complementary to the metrics used in conventional web and social analytics (e.g. engagement, bounce rate, goal conversion, key influencer (Zeng et al. 2010; Sen et al. 2006). Giant component development can be a promising indicator of how well the platform is received by target users. Basic stages of its development may give indication of the product attractiveness with varying degrees. Starting with its inception stage the platform manager can monitor its development as a typical product life cycle (PLC), for which one can find its well-known models in management and marketing (Anderson and Zeithaml 1984; Orbach and Fruchter 2014). It should be noted that one can have reservation for studying giant component development vis-à-vis PLC, but the focus of the findings is here to point out the very notion of giant component and its value to practice in a business intelligence context (H. Chen et al. 2012). For instance, with the help of the examined measures managers can be aware of product maturity status of the platform. In this regard, managers may be interested in knowing how long it will take to reach a giant regime, which can be considered as an indication of product maturity. Similarly, as observed in the platform case, significant changes in giant component pattern such as spikes can be early warnings of certain behaviours, which require a close examination of what causes these changes.

Another implication of the use of giant component is valuation of these platforms (Enders et al. 2008). In addition to conventional web analytics metrics such as unique visits, visitors, and traffic sources, platform owners and other stakeholders can use the giant component as a novel metric to monitor progression of the network, so the platform. Gneiser et al. (2012) state importance of users' connections to evaluate their contributions to the network, so to the platform value (Gneiser et al. 2012). It is worth noticing that the way we employ giant component for platform valuation is different than those approaches that are often based on node-specific contribution using the PageRank algorithm (Kimura et al. 2009). Nevertheless, there is a room for identifying and operationalizing various network specific measures referring to overall interconnectedness of users.

The empirical finding that the emergence of a giant component is a repeating process (see Fig. 4) may be related to the self-sustainability of a service. The interactions of early adopters coalesce into a giant component in about sixteen months. Later on, tying interactions of new individuals plus those of the experienced users result in the emergence of a "sub-giant" component in about a year. The GC may be seen as the culmination of these "sub-giants" over many years. Tying interactions sustain the presence of just a single large component. We never observe, for example, the emergence of

Table 2 The size and component measures for the overall tie network and GC across timeline

Time	Overall Network		Largest Component		2nd Largest Component		#Components
	#Nodes	#Edges	%Nodes	%Edges	%Nodes	%Edges	
1/3/2009	408	622	25.00	34.08	7.11	9.65	105
30/04/09	839	1398	52.56	66.81	1.07	1.14	167
30/06/10	5015	9654	59.48	74.85	0.26	0.25	822
30/06/11	5618	10,900	60.77	75.98	0.25	0.24	898
30/06/12	6708	13,516	62.34	77.94	0.30	0.28	1039
30/06/13	9447	22,508	70.76	85.71	0.21	0.17	1160

Table 3 Summary of Implications for Practice

Key Findings	Implications for Practice
Stages of giant component development and significant changes as noteworthy patterns	Product attractiveness with respect to Product Life Cycle marketing (Anderson and Zeithaml 1984; Orbach and Fruchter 2014)
Giant component related network measures	Valuation of the platforms (Gneiser et al. 2012)
Emergence of giant component and stage changes during network evolution	The effects of platform feature and design decisions (Wijnhoven and Kraaijenbrink 2008)
Approximately constant ratio	Platform Service Attractiveness (Steinfield et al. 2008) and predictability of certain events in business

another giant caused by the interactions of the "rest of the individuals" (see the outer rim of Fig. 2).

Another practical outcome of monitoring giant component development is to see the effects of platform feature and design decisions and changes on platform performance and value (Wijnhoven and Kraaijenbrink 2008). Noticeably, conventional design principles need to be revisited with regard their importance to the progress of giant component.

One can also relate theory of network to the ability and success of a service to attract customers. An online service is said to be attractive to customers if it has achieved to be the centre of attention of a major proportion of people that are all connected to each other (that is, GC). This just means that the company (deliberately/accidently) offers a service that works for the best interest of the majority of the members, that their interactions through the platform are intense enough to get past a certain threshold beyond which a networking effect is realized that connects a majority of people through (rather) short pathways. Whatever the underlying reason (their common best interests) for selecting to use the product is, the company can be considered to achieve its goal that they have created (virtual) customers loyal to their product/service, who spent a considerable amount of time interacting with it. Of course, the real customer is not people who have been attracted to the service but the people who would pay for the amount of time members spend interacting with the product (Steinfield et al. 2008).

Perhaps knowing the differences in microscopic and macroscopic state of affairs underlying the network dynamics can help managers to make predictions about investment in target audience with respect to an optimum ratio of users interacting on the platform. The platform examined shows that at the microscopic level the number of users joining the largest component is not predictable but the ratio of the number of patients to the number of physicians is surprisingly steady and may continue this way in the future. If that is not the case, this can be considered as early warnings for managers to explore the network dynamics further. Prediction and predictability of certain events in other networks have been already examined (Colizza et al. 2006) and in this regard one can further explore the importance of the metrics and ratios for platform related predictions.

6 Conclusion

The widespread use of online interactive platforms including social networking applications, community support applications draw the attention of academics and businesses. The basic trust of this research is that the very nature of these platforms can be best described as a network of entangled interactions. We contend that theory of network can contribute considerably to our understanding of online interactive platforms. In this regard, this research is aimed to show an extent to which a dynamic network analysis can advance an understanding and value of theory of network in the IS research context.

In this research we demonstrate that the interactive platform examined exhibits essential structural properties that characterize most real networks. In particular, we focus on the largest connected component, so-called a giant component, to better understand network formation. Dynamic network analysis helps us to observe the inception and evolution of the GC over time and to identify a particular pattern (including inception, growth, maturity) towards making a GC. Although tying interactions of two fundamentally different kinds of platform members seem to have resulted in an erratic network growth pattern, we have discovered that there is an order to these interactions: however different interactions are at the microscopic level, at the macroscopic level members engage in tying relations that preserve a four patients-to-one physician ratio.

We suggest that the network measures examined for the platform should be considered as novel and complementary metrics to those ones used in conventional web and social analytics. We argue that various stages of GC development can be a promising indicator of the strength and endurance of the interactions on the platform. Platform managers may take into account basic stages of giant development to determine varying degrees of product attractiveness.

We assert that the ties on the interactive health network examined are not randomly established, but consequences of consciously made tie choices of physicians and patients. Tie establishment in the platform has to do with a multiple mechanism including homophily (Johnson et al. 2014) and is subject to further investigation to articulate possibly intriguing interactions between physicians and patients.

For an Internet service to become successful the prerequisite is to combine the right content with right technologies. The question that still remains is how to measure the success and continuity of the business. Researchers in the information systems research domain should bridge the gap between theory of network and issues of managing online interactive platforms. Appropriate descriptions for each network phenomenon observed must be produced in an effort to make them part of business intelligence.

Acknowledgements The authors were supported by Kadir Has Scientific Research Project Grant 2014-BAP-05.

The authors are grateful to the CEO of Doktorsitesi.com for his valuable comments and providing data. Complete mathematical specifications and data are available from the corresponding author on request.

References

Anderson, C. R., & Zeithaml, C. P. (1984). Stage of the product life cycle, business strategy, and business performance. *Academy of Management Journal, 27*(1), 5–24.

Aydin, M. N., & Perdahci, N. Z. (2016). Network analysis of an interactive health network. *Journal of Internet Social Networking & Virtual Communities, 2016*(2016), 1.

Barabási, A.-L. (2009). Scale-free networks: a decade and beyond. *Science, 325*(5939), 412.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286*(5439), 509–512.

Barabâsi, A.-L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., & Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications, 311*(3), 590–614.

Barrat, A., Barthelemy, M., & Vespignani, A. (2008). *Dynamical processes on complex networks (Vol. 1)*. Cambridge: Cambridge University Press.

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. *ICWSM, 8*, 361–362.

Bosslet, G. T., Torke, A. M., Hickman, S. E., Terry, C. L., & Helft, P. R. (2011). The patient–doctor relationship and online social networks: results of a national survey. *Journal of General Internal Medicine, 26*(10), 1168–1174.

Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., et al. (2000). Graph structure in the web. *Computer Networks, 33*(1), 309–320.

Butler, B. S. (2001). Membership size, communication activity, and sustainability: a resource-based model of online social structures. *Information Systems Research, 12*(4), 346–362.

Chambers, D., Wilson, P., Thompson, C., & Harden, M. (2012). Social network analysis in healthcare settings: a systematic scoping review. *PloS One, 7*(8), e41911.

Chau, M., & Xu, J. (2012). Business intelligence in blogs: understanding consumer interactions and communities. *MIS Quarterly: Management Information Systems, 36*(4), 1189–1216.

Chen, X., & Yang, C.-Z. (2010). Visualization of social networks. In B. Furht (Ed.), *Handbook of social network technologies and applications* (pp. 585–610): Springer, Boston.

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: from big data to big impact. *MIS Quarterly, 36*(4), 1165–1188.

Colizza, V., Barrat, A., Barthélemy, M., & Vespignani, A. (2006). The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences of the United States of America, 103*(7), 2015–2020.

Enders, A., Hungenberg, H., Denker, H.-P., & Mauch, S. (2008). The long tail of social networking.: revenue models of social networking sites. *European Management Journal, 26*(3), 199–211.

Fetterman, D. (2009). *Data grows up: The architecture of the Facebook platform* (pp. 89–109). Inc: O'Reilly Media.

Ghose, A., & Yang, S. (2009). An empirical analysis of search engine advertising: sponsored search in electronic markets. *Management Science, 55*(10), 1605–1622.

Gneiser, M., Heidemann, J., Klier, M., Landherr, A., & Probst, F. (2012). Valuation of online social networks taking into account users' interconnectedness. *Information Systems and e-Business Management, 10*(1), 61–84.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 1360–1380.

Heidemann, J., Klier, M., & Probst, F. (2012). Online social networks: a survey of a global phenomenon. *Computer Networks, 56*(18), 3866–3878.

Johnson, S. L., Faraj, S., & Kudaravalli, S. (2014). Emergence of Power Laws in Online Communities: The Role of Social Mechanisms and Preferential Attachment. *MIS Quarterly, 38*(3), 795–808.

Jones, Q., Ravid, G., & Rafaeli, S. (2004). Information overload and the message dynamics of online interaction spaces: a theoretical model and empirical exploration. *Information Systems Research, 15*(2), 194–210.

Kane, G. C., Alavi, M., Labianca, G., & Borgatti, S. P. (2014). What's different about social media networks? A framework and research agenda. *MIS Quarterly, 38*(1), 275–304.

Kim, J., & Wilhelm, T. (2008). What is a complex graph? *Physica A: Statistical Mechanics and its Applications, 387*(11), 2637–2652.

Kimura, M., Saito, K., Nakano, R., & Motoda, H. (2009). Finding influential nodes in a social network from information diffusion data. In M. J. Young, J. Salerno, H. Liu (Eds.), *Social Computing and Behavioral Modeling* (pp. 1–8): Springer, Boston.

Kleinberg, J. (2000). The small-world phenomenon: An algorithmic perspective. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing,* (pp 163–170): ACM, New York.

Kossinets, G., & Watts, D. J. (2006). Empirical analysis of an evolving social network. *Science, 311*(5757), 88–90.

Kumar, R., Novak, J., & Tomkins, A. (2010). P. S. Yu, J. Han, C. Faloutsos (Eds.), Structure and evolution of online social networks. In *Link mining: models, algorithms, and applications* (pp 337–357): Springer, New York.

Leskovec, J., Kleinberg, J., & Faloutsos, C. (2007). Graph evolution: densification and shrinking diameters. *ACM Transactions on Knowledge Discovery from Data (TKDD), 1*(1), 2.

Li, L., Xu, L., Jeng, H. A., Naik, D., Allen, T., & Frontini, M. (2008). Creation of environmental health information system for public health service: a pilot study. *Information Systems Frontiers, 10*(5), 531–542.

Ma, J., Zeng, D., & Zhao, H. (2012). Modeling the growth of complex software function dependency networks. *Information Systems Frontiers, 14*(2), 301–315.

Marsden, P. V. (1990). Network data and measurement. *Annual Review of Sociology*. doi:10.1146/annurev.so.16.080190.002251.

Milgram, S. (1967). The small world problem. *Psychology today, 2*(1), 60–67.

Molloy, M., & Reed, B. (1998). The size of the giant component of a random graph with a given degree sequence. *Combinatorics, Probability and Computing, 7*(03), 295–305.

Moss, J., & Elias, B. (2010). Information networks in intensive care: a network analysis of information exchange patterns. In *AMIA annual symposium proceedings,* (Vol. 2010, pp. 522–526): American medical informatics association.

Newman, M. E., Watts, D. J., & Strogatz, S. H. (2002). Random graph models of social networks. *Proceedings of the National Academy of Sciences, 99*(suppl 1), 2566–2572.

Orbach, Y., & Fruchter, G. E. (2014). Predicting product life cycle patterns. *Marketing Letters, 25*(1), 37–52.

Sen, A., Dacin, P. A., & Pattichis, C. (2006). Current trends in web data analysis. *Communications of the ACM, 49*(11), 85–91.

Sharma, R., Mithas, S., & Kankanhalli, A. (2014). Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations. *European Journal of Information Systems, 23*(4), 433–441.

Simon, H. A. (1955). On a class of skew distribution functions. Oxford University Press on behalf of *Biometrika Trust*. doi:10.2307/2333389.

Steinfield, C., Ellison, N. B., & Lampe, C. (2008). Social capital, self-esteem, and use of online social network sites: a longitudinal analysis. *Journal of Applied Developmental Psychology, 29*(6), 434–445.

Strogatz, S. H. (2001). Exploring complex networks. *Nature, 410*(6825), 268–276.

van Mierlo, T. (2014). The 1% rule in four digital health social networks: an observational study. *Journal of Medical Internet Research, 16*(2).

Wang, Y., Zeng, D., Zhu, B., Zheng, X., & Wang, F. (2012). Patterns of news dissemination through online news media: a case study in China. *Information Systems Frontiers*, 1–14.

Wasserman, S. & Faust K. (1994). *Social network analysis: Methods and applications* (Vol. 8): Cambridge: Cambridge University Press.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’networks. *Nature, 393*(6684), 440–442.

Wijnhoven, F., & Kraaijenbrink, J. (2008). Product-oriented design theory for digital information services: a literature review. *Internet research, 18*(1), 93–120.

Zeng, D., Chen, H., Lusch, R., & Li, S.-H. (2010). Social media analytics and intelligence. *Intelligent Systems, IEEE, 25*(6), 13–16.

Mehmet N. Aydin is Associate Professor of MIS at faculty of Engineering and Natural Sciences and the Director of iNEO Innovation and Entrepreneurship Technology Transfer Office at Kadir Has University and has provided organizations with proven expertise in the emerging field of cloud computing, Corporate Innovations Systems and Technology Management, Network Science. He holds PhD in Management Information Systems from University of Twente, The Netherlands. Dr. Aydin worked at the Institute for Innovation and Technology Management, Ryerson University, Toronto, Canada and at the Department of Information Systems and Change Management at University of Twente, the Netherlands. He has published over 50 articles as journal papers (e.g., Journal of Database Management, Information Frontiers, Journal of Enterprise Management), book chapters (e.g., Springer series on Lecture Notes in Computer Science) conference proceedings (IFIP 8.1, CAISE). His contribution to the research field of method engineering is laudable and cited in various studies. He can be reached at mehmet.aydin@khas.edu.tr.

Nazim Ziya Perdahci is Assistant Professor of IT at the Department of Informatics at Mimar Sinan Fine Arts University and holds a PhD degree (2002) in Physics from the Bogazici University, Istanbul, Turkey. As part of his PhD work, he worked at the University of California, Irvine, USA on fractal analysis of fMRI as complex systems. Prior to joining the Mimar Sinan Fine Arts University, Dr. Perdahci worked as an Assistant Professor of IT at the Department of Information Technologies of the Işık University, Istanbul, Turkey. He has published ten articles as journal and conference papers. His current research focus is on network science in relation to management information systems and he also has two patents. He has opened an elective course about complex networks and data science recently at the Kadir Has University as part time faculty. He can be reached at nz.perdahci@msgsu.edu.tr