e-Health – For Continuity of Care
C. Lovis et al. (Eds.)

doi:10.3233/978-1-61499-432-9-443

Analysis of the Patients and Physicians Connection Network on an online Health Information Platform

Mehmet N. AYDIN[a,1] and N. Ziya PERDAHCI[b]
[a]*Faculty of Engineering and Natural Sciences, Kadir Has University, Istanbul*
[b]*Department of Informatics, Mimar Sinan Fine Arts University, Istanbul*

Abstract. Social network applications have gained popularity in the health domain as they bring health information seekers (patients and alike) and medication advice providers (physicians and other relevant actors) together. By employing a network science perspective, this research is aimed to understand an information network establishing connections among and between information seekers and providers. We found that such a connection network surfaces most of the essential characteristics of a typical complex network. Furthermore, a detailed structural analysis shows some intriguing relations and connection behaviours in the network. Implications of the findings are discussed from the perspectives of medical informatics and social network analysis.

Keywords. Complex networks, social network analysis, health information platforms, health informatics, patient experience, medication advice-seeking

Introduction

The effects of information and communications technologies (ICT) on health information and/or advice-seeking behaviors have been examined at such levels as organizational, group or individual [1]. At the individual level it is essential to understand such effects in terms of establishments of interactions within and between different health groups and actors such as physicians, nurses, patients or alike. Recent reports concerning the adoption of online health applications have shown people and organizations' significant interest in them [2]. 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-decision making [3]. Thus, one needs to find out if and how human interactions are established due to information and/or advice-seeking behaviors for health issues. Thanks to emerging online health social network platforms (e.g., HealthTap, WebMD, Doktorsitesi), which help in providing relevant data for the analysis of information and social networks.

[1] Corresponding Author.

This research is aimed to understand an extent to which characteristics are realized for one of the leading online health information websites (www.doktorsitesi.com) in Europe. According to their website, Doktorsitesi has around one million users (patients alike) and 15000 physicians. As shall be explained later on, the website examined provides physicians and patients alike with special features such as private messaging, connections, disclosed questions and answers, and other user-led content related services. In this research we focus on so-called the connection network. An established connection allows a user to send an unlimited number of messages to the user who accepts the connection request. By conducting social network analysis (SNA) we are able to surface intriguing interactions among and between medical practitioners and patients alike. The structures (patterns of interactions) underpinning such interactions turn out to be essential to explain how people learn, form opinions, and affect the others. This research focuses on the structural aspects of a network so that we can identify, if exists, salient features of a typical complex network such as small-world and giant component. As an alternative to conventional methods for analyzing networks in the health domain (that is, survey and observation for data collection and limited unit of analysis [4]), we make use of valuable data generated on a health information platform and employ relevant SNA [5] methods.

1. Methods

We obtained a dataset describing the connection requests of Doktorsitesi members. The set is composed of all records of established connections over the 3-month period from October to December 2012. For each connection, we have all information regarding the transaction that resulted in a connection, where the user identifiers of both parties and the timestamp of the record specify a transaction. Description of the network data and visual analysis of network diagrams are produced with Gephi [6], which is a visualization and exploration platform for SNA. It is open-source and free software.

We model the connection network as a graph of directed edges where patients (abbreviated to Ps) and physicians (abbreviated to Drs) are represented by nodes, and connection requests are represented by directed edges. Table 1 lists the basic statistics of the connection request network as well as the connections solely between Drs and Ps.

Two basic measures, node degrees and network density, provide useful, but limited, insights about the structure of the network. The former focuses on simple counts of in-, out-, maximum connections, and the latter captures how highly connected nodes are by calculating the percentage of all possible connections that are realized. To better examine an extent to which nodes are clustered (that is, clustering coefficient), one can measure number of triplets of nodes where three nodes are connected by two or three edges. Clusters of nodes can create a significant portion of the whole network (called a giant component) or small components [7]. 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 weakly connected.

Another important characteristic related to the global structure of a network is path length. Path length measures the distance between nodes in terms of number of connections in the network examined. Thus, it simply shows how far apart people (Ps and Drs) are. The small-world effect is present if the average path length between every node is around six [8]. The maximum geodesic distance, called diameter of a network, is the largest distance of all.

2. Results

We present basic characteristics of the connection network for the overall network and two specific connections types (Table 1). To better visualize edges between nodes we provide network models by alternative layouts. Fig 1 utilizes Force Atlas layout algorithm, which produces readable spatial arrangements of nodes.

Table 1. Summary of SNA measures for the connection network

Network	Nodes	Edges	Av Degree	Density	Av Path Length	Av Clustering Coeff
Overall	818	1768	2.161	0.003	3.629	0.022
Drs to Drs	42	52	1.238	0.03	1.587	0.000
Ps to Ps	314	568	1.809	0.006	3.479	0.004

In fig 2 the graph is laid out using a customized layout algorithm (alphabetical ordering of the nodes of the same type on two vertical lines) where the Dr nodes are on the right and Ps are on the left. In this visualization, interactions between the same node types are clearly visible.

The directed network of connections has N = 818 nodes, 572 (70%) of which being Ps and 246 (30%) being Drs, hence the average degree is 2.161, suggesting that a typical member accepts two connection requests and makes two connections requests. Yet, with the help of the degree distribution, we see that this number is misleading, because the majority of actors have less than two requests. To be more specific, 70% of them made fever than two connection requests, and 74% of them accepted fever than two connection requests. These nodes coexist with highly connected nodes, or hubs. The largest two of these hubs have 318 (in 171, out 147) and 235 (in 115, out 120) connections. This is a characteristic seen in many real networks.

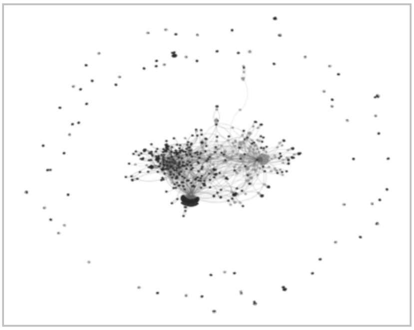

Figure 1. Network by force atlas layout

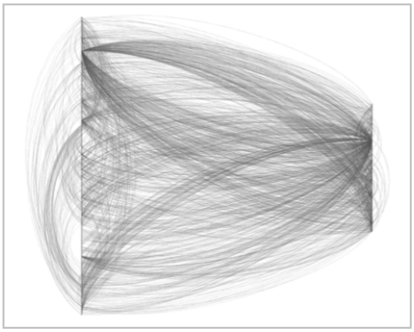

Figure 2. Network by customized layout

The connection network is, to a large extend, weakly connected (fig 1). Overall, there are 73 weakly connected components and 268 SCCs. One of the characteristics of real networks is the emergence of a giant component. In the present work, we observe that the giant connected component comprises 79% of the network. More importantly, 75% of these nodes belong to giant SSC (middle of fig 1). 77% of Ps 84% of Drs are in the giant connected component whereas 53% of Ps and 74% of Dr contribute to the giant SCC. The average path length between all pair of nodes within the SCC is 3.592 and maximal distances vary between graph diameter of 8 and graph radius of 5. A little bit more exploration reveals that around 55% of the people can reach each other in no more than 8 hops, 40% in no more than 6 hops. It is a *small world* [8].

The density of the whole network is 0.003. It is a sparse network like most of the real networks. The average clustering coefficient is 0.022. Note that clustering is measured by making the network undirected, which is not uncommon in network science [5]. This clustering value is nearly an order of magnitude higher than the clustering of a random graph of the same number of nodes and edges. Connections between Ps and Dr alone are even sparser (fig 2). There is less clustering between Ps and none among Drs (Table 1). On the whole, 66% of the connection requests are reciprocated. 81% of the connection requests between patients and physicians are reciprocated; whereas only 41% of them are reciprocated among patients, and 73% are reciprocated among few physicians.

3. Discussion

The connection network examined exhibits essential properties that characterize most complex networks (e.g. WWW, Power Grid, Movie actors etc.) [9], namely:

i. *Smallworlds*: Despite the large size of the network there exist short paths between any of two platform users (Ps or Dr).
ii. *Clustering:* Small groups form within a network due to various reasons. Within the health context local clusters may be due to, among other things, "like-minded others" sharing common health issues.
iii. *Degree distribution:* The degree distribution of the network examined exhibits a *hub-and-spoke* connection pattern; there exists hubs (nodes with large number of connections) with surrounding spokes. The largest hub is a physician while the next is a patient. One would argue that the reason behind this pattern should be *preferential attachment* [9], which is not atypical for complex networks. However, the key question still remains: what are the underlying reasons for this attachment? One can suspect such other reasons as popularity of the physician for some reasons, misuse of the platform by a spammer, marketing and advertisement.

One of the key implications of the realization of these properties is that the connections on the platform are not by chance but consequences of conscious choices. The results concerning reciprocated connections help us interrogate online behavior of the users examined. Physicians and patients tend to have reciprocated connections with each other. The fact that there are few reciprocated requests among Drs may suggest that they are reluctant to establish connections with their peers. The same holds true for patients as well. We argue that one of the reasons behind few Dr-to-Dr reciprocated requests can be explained by physicians' perception on health information platforms, which is already acknowledged [3] that the majority of the physicians surveyed (68%) in the US concerning attitudes toward online social networks did not agree that it was ethically acceptable to interact with patients on OSN, either for social or patient-care reasons. This suggests that for further research one should take into account physicians' perceptions on OSN with respect to ethics and/or professional use while examining reciprocations among physicians. One can further argue that platform may serve physicians an effective channel to attract patients.

Regarding few patients' reciprocated-connections, one may argue that patients having similar health related issues are not inclined to exchange messages indefinitely. Although Drs and Ps are motivated for different reasons they tend to make reciprocated connections. On the Drs' site the motivation is the opportunity to introduce themselves to as many patients as they could, on the Ps' site the opportunity to send as many messages as they like to get valuable information. It is also found [3] that the majority of both online and offline health information seekers report reliance upon healthcare professionals as a source of health information.

4. Conclusion

In this research we have shown that network science can contribute considerably to our understanding of an online health information exchange platform. Our attempt to unravel the glue that holds platform users together led us to the conclusion that patients seek information, physicians make use of connections to make patients aware of their expertise.

Further work with a larger data set is needed to deepen our understanding of evolutionary nature of the interactions. Incorporation of time stamps is needed to explore dynamics of the network so as to identify critical changes in the network at a local and global scale. That is, it would be essential to identify key influencers, their roles and information passing behaviors by employing centrality measures such as betweenness, closeness, and brokerage roles.

For comparison purposes and other aspects of the health information platform, one can examine other types of the interactions (questions and answers network, messages network) that may reveal [2]. Another potential research area is the analysis of modularity or formation of communities and cliques to articulate social identity development in online health platforms. Thanks to advances in SNA, the endeavor for better understanding of the structure and dynamics of health actors' interactions will illuminate the future state of health social networks.

Acknowledgements

The authors are grateful to M.D. Erden Asena for his valuable comments.

References

[1] Dunn AG, Westbrook JI. Interpreting social network metrics in healthcare organisations: A review and guide to validating small networks. Soc Sci Med. 2011 Apr;72(7):1064-8.
[2] Moss J, Elias B. Information networks in intensive care: A network analysis of information exchange patterns, AMIA Annu Symp Proc. 2010 Nov 13;2010:522–6.
[3] Bosslet GT, Torke AM, Hickman SE, Terry CL, Helft PR. The patient–doctor relationship and online social networks: Results of a national survey. J Gen Intern Med. 2011 Oct;26(10):1168-74.
[4] Chambers D, Wilson P, Thompson C, Harden M. Social network analysis in healthcare settings: A systematic scoping review. PLoS One, 2012 Aug 3; 7(8):e41911.
[5] Leskovec J, Huttenlocher D, Kleinberg J. Signed networks in social media. Proc of the SIGCHI Conf on Hum Factors in Comput Syst, ACM, 2010 Apr; 1361-70.
[6] Bastian M, Heymann S, Jacomy M. Gephi: An open source software for exploring and manipulating networks. Proc of the 3rd Intl AAAI Conf on Weblogs and Social Media, 2009 May, 361-2.
[7] Newman MEJ, D. Watts DJ, Strogatz SH. Random graph models of social networks. Proc of the National Academy of Sciences of the USA, 2002;99(Suppl 1):2566-72.
[8] Milgram S. The small world problem. Psychol Tod. 1967 May;1(1), 61-7.
[9] Albert R, Barabasi A-L. Statistical mechanics of complex networks. Rev Mod Phys. 2002 Jan 30;74(1):47-97.