Information Passing in Healthcare Social Networks

Gülay ÜNEL

Işık Üniversitesi, Enformasyon Teknolojileri Bölümü, İstanbul, Türkiye
gulay.unel@isikun.edu.tr
(Geliş/Received:27.06.2016; Kabul/Accepted:27.10.2016)
DOI: 10.17671/btd.25307

Abstract— In recent years, social networks have become one of the most influential developments in our lives. In this paper, we investigate trust relationships and information passing in online social networks that focus on health. Our results are based on a social network called doktorsitesi.com which is the one of the largest online social networks in Turkey about healthcare that is managed by professionals. We show that there is a connection among patients in terms of information passing in doktorlarsitesi.com and quantify this information passing. Our findings implicate that healthcare interactions are embedded in social networks, and the results on the existence of information passing on other types of e-commerce social networks also apply to healthcare social networks.

Keywords— information passing, healthcare, social networks

1. INTRODUCTION

The emergence of online social networks (OSNs) in the last years has led to a huge increase in the volume of information about individuals, their activities, connections amongst individuals or groups, and their opinions [1]. As a result, investors started to create online social networks for specialized purposes according to their needs. One of these specialized areas that use online social network structures is healthcare. Social networks hold considerable potential value for healthcare organizations because they can be used to reach stakeholders, aggregate information and leverage collaboration. According to a study 55% of surveyed Americans gets information about a therapy or condition online [2]. Hence, social networks have become an important topic of scientific research in social behavior on various domains including healthcare such as determining large outbreaks of infectious diseases, disability, accidents, and understanding how people react to crisis situations [3].

The fundamental process we focus on in this study is information passing which investigates the flow of social influence in commerce networks [4]. We analyze how information passing influences healthcare social networks in various cases such as: a patient asks a question to a doctor, the doctor answers the patient, then the patient sends a message to a friend, what is the likelihood that the friend will then ask a question to the same doctor?

We obtained a healthcare social network dataset from doktorlarsitesi.com to conduct our research which is one of the largest healthcare networks in Turkey. The network connects doctors and patients and provides a platform for searching for information, sharing information and messaging among its users.

In our study, we analyze activities of 25,512 unique patients from doktorsitesi.com. We quantify information passing in doktorlarsitesi.com using triadic closure processes. We show that there is a connection among patients in terms of information passing.

In addition to the existence of information passing we also investigated the relationship between the factors: communication strength and time difference, and information passing. Our results show that the stronger the communication between the two patients, the more likely that information passing will occur. In terms of the time difference it is expected that the larger the time difference between the interaction with the doctor and the message between patients, the lower the influence of the message on the interaction of the patient (who is on the receiver end of the message) with the doctor. Our results show that the probability of information passing success generally decreases with time.

Our findings implicate that healthcare interactions and transactions are embedded in social networks, and that patients' social connections affect the doctors they choose to interact with.

Understanding the use of social networks on purchasing behavior is a fundamental e-commerce research topic [5][6]. However, the research on the use of social networks on healthcare interactions is limited. Our work quantifies the different aspects of the relationship between social networks and healthcare connections.

Remainder of this paper is organized as follows. In Section 2, the related work from the literature for the technical presentation is presented. In Section 3, the case study on doktorlarsitesi.com is explained. In Section 4, the analysis and the results of the paper are explained, and finally the paper is concluded in Section 5.

2. RELATED WORK

The relationship between social networks and consumer behavior is a popular research topic in e-commerce, however the focus is mostly on product recommendation [7][8][9][11][12][13]. About this relationship, it is argued that economic transactions are embedded in social networks [10] and that the social graph is the most important feature in predicting consumer choice [4]. Our results are also parallel to this argument.

Information passing was first introduced for online shopping. A work on information passing and triadic closure [14][15][16][17] was on analyzing Taobao, a Chinese online shopping site that is one of the world's largest e-commerce networks [4]. The study, which was built upon works by [18] and [19], focused on the relationship between transactions and connections between the users. The presence of information passing is quantified by exploring triads and the directed closure process. The work also analyzed trust and as a result, price of trust [20][21][22][23] was defined as the extra amount a buyer is willing to pay for transaction with a highly rated seller. One of the results obtained was that higher rated sellers were able to sell their products with higher prices because sellers think that highly rated sellers may provide better services, such as replying to messages from customers in a timely fashion, or shipping products more frequently. In other words, buyers were willing to pay more to highly rated sellers to minimize transaction risk, thus sellers who maintained good reputations were financially rewarded.

Al-Oufi et al. defined trust in online social network environment as follows: "Trust in a person is a commitment to an action based on a belief that the future actions of that person will lead to a good outcome" [24]. In addition, they adduced that trust in online social networks has three primary characteristics: transitivity, asymmetry, and personalization.

Additionally, in [25], the focus is moved toward time-aware trust prediction in evolving online trust networks. In this work the impact of considering the temporary progress of trust networks was examined explicitly in trust prediction tasks by using a supervised learning method.

Finally, [26] looks particularly at assigning trust in web-based social networks and investigates how trust in information can be mined and incorporated into applications.

We primarily focus on information passing in this study. However a trust model can also be built on top of this work based on information passing and interactions between the users of our domain

3. CASE STUDY: doktorlarsitesi.com

In this paper, we analyze a domain called doktorsitesi.com, which is the largest online social network about health in Turkey. The purpose of this Web site is to create and maintain interactions between patients and doctors. In this site, patients can freely ask questions to the doctors and get answers from them. They can also get in touch with the doctors. In addition to these, patients can check the articles or the videos of the doctors, share them online, and can send messages to each other.

Online social networks have two primary components, which are nodes that represent the users and edges that represent the relationships. In our case nodes represent the patients and the doctors, edges represent the relationships that we can divide to three groups each having its own characteristics: Patient to Patient, Doctor to Doctor, Patient to Doctor.

The dataset we used for our analysis includes 25,512 unique users who send question(s) to the doctor(s). The statistics of the dataset are given in Table 1.

Table 1. Dataset Statistics

Entity/Relationship	Cardinality
Patients	25,512
Messages	7997
Doctors	2494
Questions	60,690
Answers	46,222
Follow actions	3632
Thank actions	6287

In this work, our aim is to investigate the information passing and trust relationships among the nodes in our network. More formally, if patient P_1 asks question to doctor D_1 and then send message to another patient P_2, will patient P_2 then ask question to doctor D_1 as well?

In terms of data analysis, we primarily focus on the method used for TaoBao [4] network. Our aim is to investigate the information passing and trust relationships among the nodes in our network. In TaoBao network, the nodes can be buyers and sellers where buyers can share comments about the sellers, and send messages to other buyers. We used the method described in [4] to analyze information passing and trust among patients in our case.

Our main hypothesis is that the probability of a patient asking a question to a doctor has a relationship with information passing. In addition to the existence of information passing we also investigate the relationship between the factors: communication strength and time difference, and information passing. It can be hypothesized that the stronger the communication between the two patients, the more likely that information passing will occur. In terms of the time difference it is expected that the larger the time difference between the interaction with the doctor and the message from P_1 to P_2, the lower the influence of the message on the interaction of P_2 with the doctor.

We conduct our research on 4 steps. Each step is designed in terms of the different actions that can take place between the patients and the doctors to understand and investigate the information passing between the patients. We calculate the information passing success rate among patients for each step. Additionally, we calculate the number of exchanged messages among patients and the time difference between these messages in each step. The steps we defined are as follows:

- Step 1: Given that P_1 first asked a question to D_1 and received a reply from D_1, will P_2 ask a question to D_1?
- Step 2: Given that D_1 is followed by P_1, P_1 asked a question to D_1 and received a reply from D_1, will P_2 ask a question to D_1?
- Step 3: Given that P_1 asked a question to D_1, received a reply from D_1 and thanked to D_1, will P_2 ask a question to D_1?
- Step 4: Given that D_1 is followed by P_1, P_1 asked a question to D_1, received a reply from D_1, and thanked to D_1, will P_2 ask a question to D_1?

In all our steps, we focused on the patients who received answers from the doctors. There are 7.489 unanswered questions in our dataset, and if the doctor did not reply a patient's question, it means that the patient was not able to create a trust connection with another patient. Therefore, we skip the unanswered questions.

For the questions we analyzed, we also used timestamps for the events except follow and thank events. For instance, for the question: "Given that P_1 first asked a question to D_1 and received a reply from D_1, will P_2 ask question to D_1?" the timestamps we used are:

- T_0: P_1 asked question to D_1
- T_1: D_1 answered the question
- T_2: P_1 sent message to P_2
- T_3: P_2 asked question to D_1

where $T_0<T_1<T_2<T_3$.

4. RESULTS

In this section, we provide our results based on TaoBao [4] trust model. We provide the information passing success rate that is also called the triangle probability and details about the influences on information passing for our steps mentioned in the previous section. In terms of the variables that influence information passing we investigate communication strength and time difference.

Step 1: The information passing success rate of our network for this step can be defined as $Prob(E_1|E_2)$ where E_1 is the number of patients P_2 who ask a question to a doctor D_1 at T_3 , E_2 is the number of patients P_1 who ask a question to D_1 at T_0 and D_1 answers P_1 at T_1 and P_1 messages to P_2 at T_2. $Prob(E_1|E_2)$ for this step is 0.098.

For this step, we also computed a base probability value for comparison. We used the probability of a patient asking a question to a doctor on the days that the patients in P_2 role asked question as our base of comparison which turned as 0.0037.

For our steps Figure 1, Figure3, Figure 5, and Figure 7 show the relationship between the number of messages exchanged between the patients P_1 and P_2 and the triangle probability. Figure 2, Figure 4, Figure 6, and Figure 8 give the relationship between the time difference between reply from the doctor and the initial message sent to P_2 and the triangle probability.

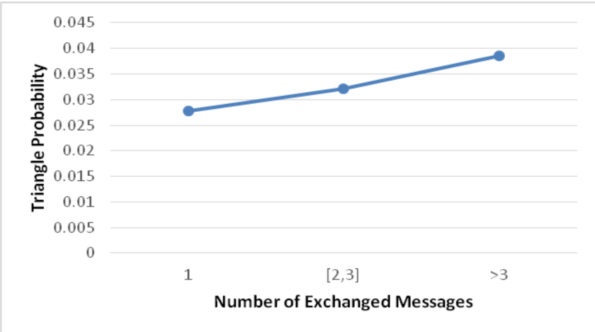

Figure 1. Number of Messages vs. Triangle Probability for Step 1

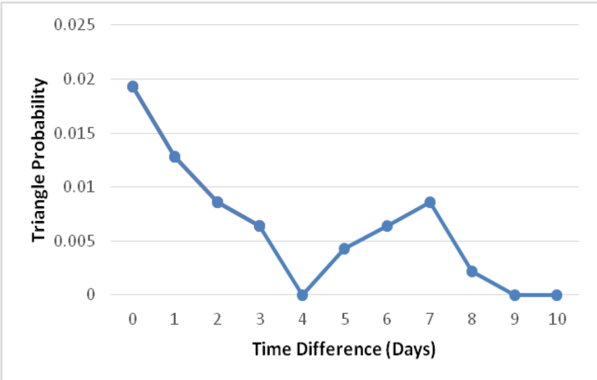

Figure 2. Time Difference vs. Triangle Probability for Step 1

Step 2: The information passing success rate of our network for this step can be defined as $Prob(E_1|E_2)$ where E_1 is the number of patients P_2 who ask a question to a doctor D_1 at T_3, E_2 is the number of patients P_1 who follow D_1, P_1 asks question to D_1 at T_0 and D_1 answers P_1 at T_1 and P_1 messages P_2 at T_2. $Prob(E_1|E_2)$ is 0.06.

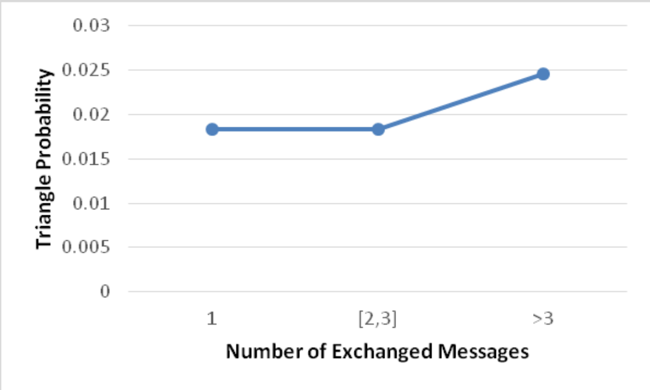

Figure 3. Number of Messages vs. Triangle Probability for Step 2

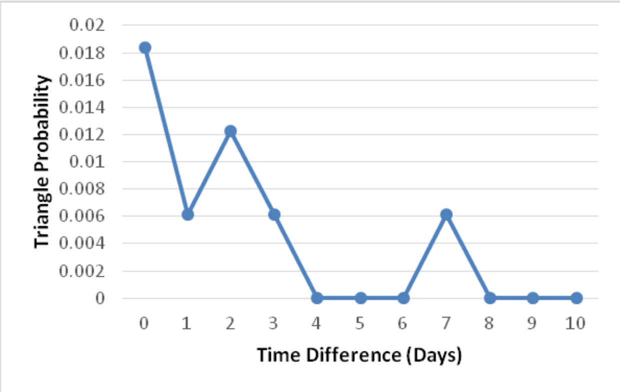

Figure 4. Time Difference vs. Triangle Probability for Step 2

Step 3: The information passing success rate of our network for this step can be defined as Prob($E_1|E_2$) where E_1 is the number of patients P_2 who ask a question to a doctor D_1 at T_3, E_2 is the number of patients P_1 who ask a question to D_1 at T_0 and D_1 answers P_1 at T_1, P_1 thanks to D_1, and P_1 messages P_2 at T_2. Prob($E_1|E_2$) is 0.58.

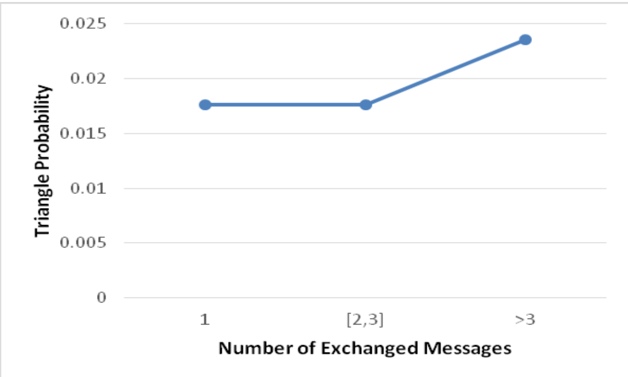

Figure 5. Number of Messages vs. Triangle Probability for Step 3

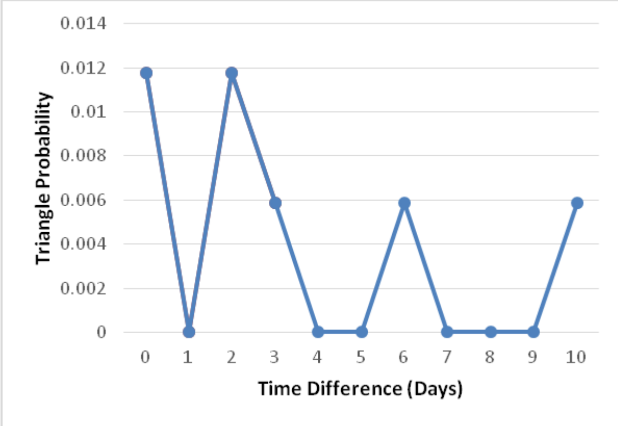

Figure 6. Time Difference vs. Triangle Probability for Step 3

Step 4: The information passing success rate of our network for this step can be defined as Prob($E_1|E_2$) where E_1 is the number of patients P_2 who ask a question to a doctor D_1 at T_3, E_2 is the number of patients P_1 who follow D_1, P_1 asks question to D_1 at T_0 and D_1 answers P_1 at T_1, P_1 thanks to D_1, and P_1 messages P_2 at T_2. Prob($E_1|E_2$) calculated for this step is 0.02.

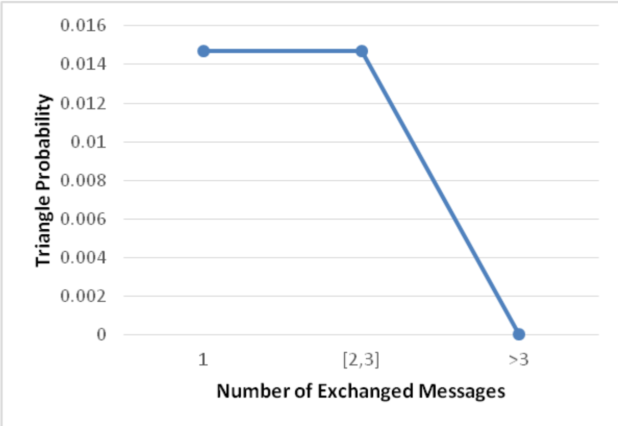

Figure 7. Number of Messages vs. Triangle Probability for Step 4

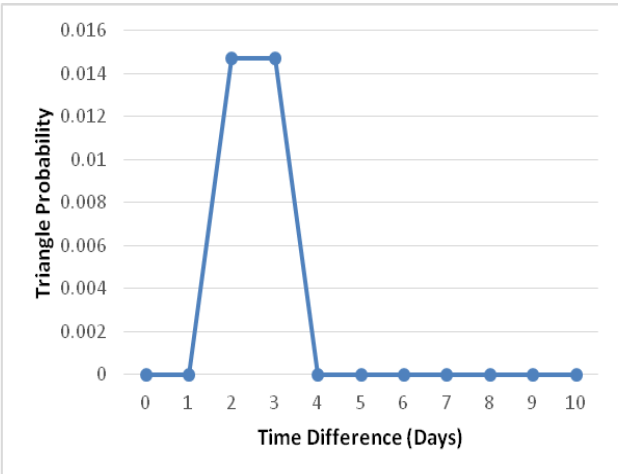

Figure 8. Time Difference vs. Triangle Probability for Step 4

Following our examination of the results we can argue that our dataset has parallel properties to the TaoBao dataset in terms of information passing. The information passing success rate we computed for our dataset in Step 1 is 0.098 whereas it is 0.00203 for the TaoBao network. One reason for the difference in the rate can be the size difference between the two datasets where ours is much smaller than the TaoBao dataset which is also the most significant limitation of our work. Another reason can be related to the internal properties of the two networks such as their use. Since our network is used primarily for getting advice from professionals and the type of all the relationships are based on messages rather than trades and messages in the TaoBao case, this can also be a reason for the higher rate of information passing in our case.

In terms of the relationship between the factors: communication strength and time difference, and information passing; it was hypothesized that the stronger the communication between the two patients, the more likely that information passing will occur. Our results are parallel to this expectation as it was also the case in [4]. In terms of the time difference it was expected that the larger the time difference between the interaction with the doctor and the message from P_1 to P_2, the lower the influence of the message on the interaction of P_2 with the doctor. Our results show that the probability of information passing success generally decreases with time as it was also the case in [4].

The probabilities we calculated provide us a better understanding of the influence of information passing in healthcare social networks. They also imply that healthcare interactions and transactions are embedded in dynamic social networks. Hence, our results can also be used to build more intelligent healthcare systems based on social networks.

5. CONCLUSIONS AND FUTURE WORK

Our work analyses the activities of 25,512 unique patients from doktorsitesi.com. Through the study of Taobao case and using the trust methodology of Taobao, we verified that there is a connection among patients in terms of information passing and trust. Our results show the existence of information passing in our network. In terms of the relationship between information passing and related parameters, namely communication strength and time difference we also show that our results conform to the expectations as the Taobao network.

We hope our study will motivate the future research into online social networks on health. Future areas of related study include: analyzing the trust relationships in doktorsitesi.com among doctors and the success of information passing from the doctors' perspective, analyzing the relationship between the information passing success rate and strength of the relationship among doctors, and investigating the trust relationship in the online social networks supported by professionals and non-professionals.

REFERENCES

[1] S. Bhagat, G. Cormode, S. Muthukrishnan, "Node Classification in Social Networks", *Social Network Data Analytics*, IBM T. J. Watson Research Center, NY 10598, 115-148, 2011.

[2] P. H. Keckley, M. Hoffmann, "Social Networks in Health Care: Communication, collaboration and insights", Deloitte Center for Health Solutions, 2010.

[3] I. Masic, S. Sivic, S. Toromanovic, T. Borojevic, H. Pandza, "Social Networks in Improvement of Health Care", Avicena, 2012.

[4] S. Guo, M. Wang, J. Leskovec, "The Role of Social Networks in Online Shopping: Information Passing, Price of Trust, and Consumer Choice", Stanford University, Stanford, CA, USA, 2011.

[5] P. DiMaggio, H. Louch. "Socially embedded consumer transactions: for what kinds of purchases do people most often use networks?", *Amer. Soc. Rev.*, 63(5), 619–637, 1998.

[6] T. Hennig-Thurau, K. Gwinner, G. Walsh, D. Gremler. "Electronic word-of-mouth via consumer-opinion platforms:'What motivates consumers to articulate themselves on the Internet?", *J. of Interactive Marketing*, 18(1), 38–52, 2004.

[7] B. Sarwar, G. Karypis, J. Konstan, J. Riedl. "Analysis of recommendation algorithms for e-commerce". EC, 2000.

[8] J. Schafer, J. Konstan, J. Riedi. "Recommender systems in e-commerce". EC, 1999.

[9] J. Zheng, X. Wu, J. Niu, A. Bolivar. "Substitutes or complements: another step forward in recommendations". EC, 2009.

[10] M. Granovetter. "Economic Action and Social Structure: The Problem of Embeddedness". *Amer. J. of Sociology*, 1985.

[11] S. Hill, F. Provost, C. Volinsky. "Network-based marketing: Identifying likely adopters via consumer networks". *Statistical Science*, 21(2), 256–276, 2006.

[12] J. Leskovec, L. Adamic, B. Huberman. "The dynamics of viral marketing". *ACM TWEB*, 1(1), 5, 2007.

[13] R. Bhatt, V. Chaoji, R. Parekh. "Predicting product adoption in large-scale social networks". CIKM, 2010.

[14] D. Liben-Nowell, J. Kleinberg. "The link-prediction problem for social networks". CIKM, 2003.

[15] J. Leskovec, L. Backstrom, R. Kumar, A. Tomkins. "Microscopic evolution of social networks". KDD, 2008.

[16] J. Leskovec, D. Huttenlocher, J. Kleinberg. "Predicting positive and negative links in online social networks". WWW, 2010.

[17] D. Romero, J. Kleinberg. "The directed closure process in information networks with an analysis of link formation on twitter". ICWSM, 2010.

[18] A. Rapoport. "Spread of information through a population with socio-structural bias: I. Assumption of transitivity". *Bulletin of Mathematical Biology*, 15(4), 523–533, 1953.

[19] M. Granovetter. "The strength of weak ties". *American Journal of Sociology*, 78(6), 1360–1380, 1973.

[20] P. Resnick, R. Zeckhauser. "Trust among strangers in

Internet transactions: Empirical analysis of eBay's reputation system". *Adv. in Appl. Microecon.: A Research Ann.*, 2002.
[21] P. Resnick, R. Zeckhauser, J. Swanson, K. Lockwood. "The value of reputation on eBay: A controlled experiment". *Experimental Economics*, 9(2), 79–101, 2006.
[22] D. Houser, J. Wooders. "Reputation in auctions: Theory, and evidence from eBay". *J. of Econ. & Man. Strat.*, 2006.
[23] D. Lucking-Reiley, D. Bryan, N. Prasad, and D. Reeves. "Pennies from ebay: the determinants of price in online auctions". *The J. of Industrial Economics*, 2007.
[24] S. Al-Oufi, H. Kim, A. El Saddik, "A Group Trust Metric For Identifying People of Trust in Online Social Networks", *Expert Systems with Applications*, 39(18), 13173–13181, 2012.
[25] K. Zolfaghar, A. Aghaie, "Evolution of Trust Networks in Social Web Applications Using Supervised Learning", *Procedia Computer Science*, 3, 833–839, 2011.
[26] J. Golbeck, J. Hendler, "Inferring Trust Relationships in Web-based Social Networks", *ACM Trans. Internet Technol.*, 6(4), 497–529, 2006.