2015 International Symposium on Nonlinear Theory and its Applications
NOLTA2015, Kowloon, Hong Kong, China, December 1-4, 2015

Analysis on the spread of disease on temporal networks

Koshi Abe†, Yutaka Shimada†, Kantaro Fujiwara† and Tohru Ikeguchi†

†Department of Management Science, Graduate School of Engineering, Tokyo University of Science
1-3 kagurazaka, Sinjuku-ku, Tokyo, Japan

Abstract—To avoid spreading an infectious disease, it is one of the important approaches to analyze mathematical models of the infectious disease. Many previous studies have already investigated the spread of infectious diseases on complex networks. However, these conventional studies mainly focus on static networks whose structures do not change with time, even though structures of complex networks change with time under realistic situations. In this paper, focusing on temporal networks whose structures change with time, we investigated the spread of infectious diseases on the real temporal networks observed from person-to-person interactions at a hospital and a high school. We also investigated what kinds of properties affect the spread of the infectious disease by using a simple mathematical model of the infectious disease. As a result, the degree correlation and the correlation of the number of contacts of connecting two vertices play a crucial role to the spread of the infectious disease on the temporal network.

1. Introduction

In the complex network science, various real phenomena are described as a network which consists of a set of vertices and a set of edges, for example, the Internet, prey-predator relations in an ecosystem [1], gene networks [1], economic systems, and face-to-face interactions between individuals [2]. Researches in the complex network science have clarified that several common features underly these real networks, such as small-world [3] and scale-free [4] properties.

Information diffusion and the spread of disease have also been discussed on the complex networks [5]. These previous researches mainly focus on static networks, namely the networks whose structures do not change with time. However, in the real networks, the structure of network changes with time [6].

In this paper, focusing on temporal changes in structures of networks, we investigated how infectious diseases spread over real temporal networks by using data obtained from face-to-face interactions between individuals recorded by the radio frequency identifier(RFID) in a hospital and a high school [7]. We employed a simple disease-transmission model and analyzed how the infectious disease spatiotemporally spreads on these real temporal networks. Then, we investigated what structural properties of the networks affect the spread of the infectious disease. As a result, the infectious disease spreads widely on network when the number of contacts is high and the degree correlation and the correlation of the number of contacts of connecting two vertices are negative.

2. Data

In this paper, we used two types of the data of the face-to-face contacts of individuals recorded by SocioPatterns [7–9]. The first one is observed from the face-to-face contacts between health care workers and patients at the hospital in Lyon, France. The second one is observed from those between high school students at the Lycée Thiers, Marseilles, France. The contacts between individuals were recorded every 20 seconds.

In the data of the hospital, the contact pattern was collected during five days from Monday, December 6, 2010 at 1:00 pm to Friday, December 10, 2010 at 2:00 pm. The subjects participating in the experiments were 27 nurses and nurses' aides, 11 medical doctors, 8 administrative staff members and 29 patients.

In the data of the high school, the contact pattern was collected during for four days (from Tuesday to Friday) in 2011 and seven days in 2012 (from Monday to Tuesday of the following week except for the Saturday and the Sunday). The number of subjects participating in the experiments in 2011 is 126, and that in 2012 is 180. The high school students are classified into three groups, PC, PC*, and PSI* in 2011, and into five groups, MP*1, MP*2, PC, PC* and PSI* in 2012. The groups MP*1 and MP*2 focus on the mathematics and the physics, PC and PC* focus on the physics and the chemistry, and PSI* focuses on engineering studies. Table 1 shows the number of students belonging to each class.

Table 1: The number of subjects belonging to each group in the high school.

	teacher	PC	PC*	PSI*	MP*1	MP*2	total
2011	8	31	45	42	0	0	126
2012	0	38	35	41	31	35	180

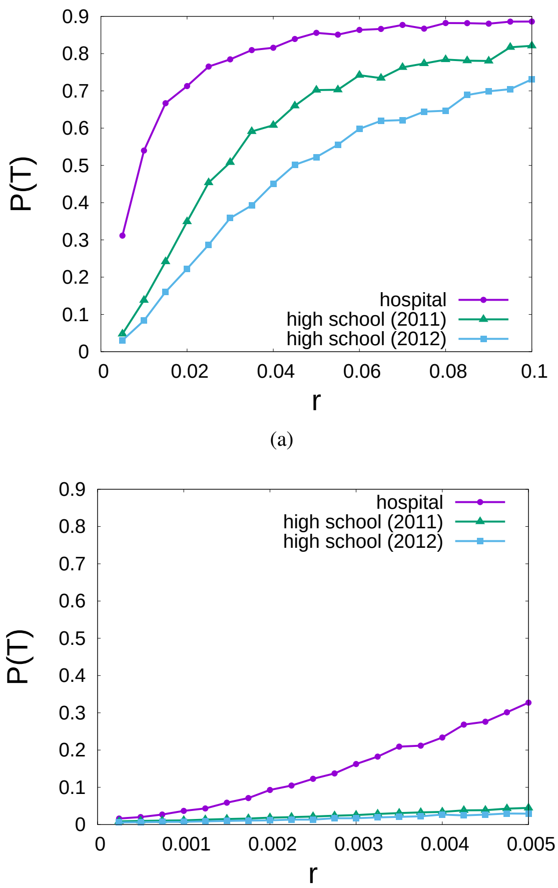
3. Methods

We analyzed how infectious diseases spatiotemporally spread on contact networks by using the real data and a simple mathematical model of the infectious disease. We constructed networks where vertices correspond to subjects and the contacts between the subjects are described by edges. The edge between the vertex i and the vertex j at time t is described by $l_{ij}(t) \in \{0, 1\}$, where if a contact exists, $l_{ij}(t) = 1$, otherwise $l_{ij}(t) = 0$. We also defined $S_i(t) \in \{0, 1\}$ as a state of the vertex i at time t. When the vertex i is infected with the infectious disease, $S_i(t) = 1$, otherwise $S_i(t) = 0$. Let r be the probability that a susceptible vertex changes to an infected one. The number of subjects is N. The number of infected persons at time t is $I(t)$ $(t = 0, 20, \cdots, T)$. The ratio of the number of infected persons at time t to the total number of subjects is $P(t) \equiv I(t)/N$. When the vertex i contacts with the vertex j under the situation that $S_i(t) = 1$ and $S_j(t) = 0$ at time t, the infectious disease is transmitted from the vertex i to the vertex j with the probability $rl_{ij}(t)$. When the vertex j is infected at t $(S_j(t) = 1)$, the state of the vertex j changes to the removed state after a fixed period τ. Removed vertices do not transmit the infectious disease to other vertices, and they are not infected again. We investigated how the infectious disease spreads on the contact networks by $P(t)$.

4. Results

4.1. The relation between the infection rate r and the normalized number of infected persons $P(T)$

We first conducted experiments in the condition where $\tau = 72$ $[h]$. We calculated the final ratio of the number of infected persons to the total number of subjects $P(T)$. Figure 1 shows how $P(T)$ changes when the infection rate r changes. From Fig. 1(a), $P(T)$ in the case of the hospital is higher than that of the high school in all values of the infectious rate r. In addition, the number of infected persons of $P(T)$ in the case of the hospital increase more quickly than that of the high school in $0 \le r \le 0.03$. We next investigated how $P(T)$ increases when the values of r are small $(0.00025 \le r \le 0.005)$. From Fig. 1(b), we found that $P(T)$ of the hospital increases when the infectious rate r increases. However, $P(T)$ of the high school does not increase even though the infectious rate r increases. One of possible causes of the different tendency of the spread of the infectious diseases between the hospital and the high school is the difference between the number of contacts in the hospital and that in the high school. When the number of contacts is large, the suspectible vertices are likely to be infected from other infected vertces, because the chances that the vertex contacts with the infected vertices increase. Indeed, the number of contacts in the hospital is 32,424 during the five days. On the other hand, the number of contacts in the high school in 2011 is 28,561 during the four days and that in 2012 is 45,047 during the seven days. Then, the average number of contacts is 86.46 in the hospital network, 56.67 in the high school network in 2011, and 35.75 in the high school network in 2012.



(a)

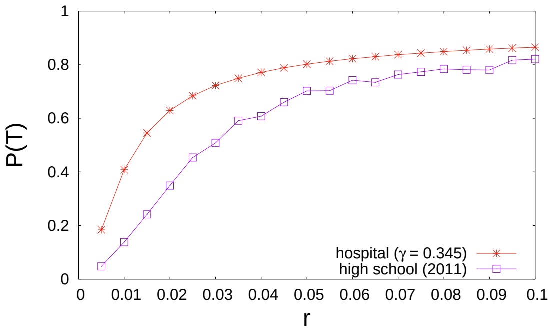
(b)

Figure 1: The relation between infection rate r and the ratio of the number of infected persons at time T, $P(T)$. The range of r is (a) $0.005 \le r \le 0.1$, (b) $0.00025 \le r \le 0.005$. We simulated 1,000 trials for each condition and calculated those of average.

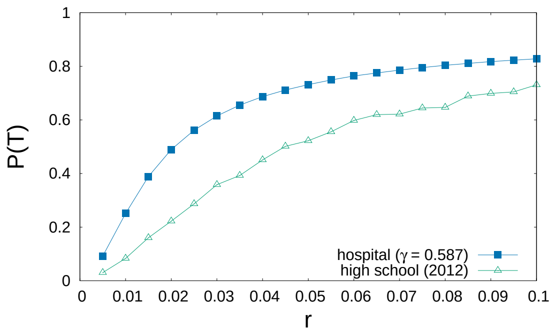
4.2. The relation between $P(T)$ and the average number of contacts

We investigated how the average number of contacts affects the ratio of the number of infected persons at time T, $P(T)$. To eliminate the differences in the number of contacts between the hospital and the high school, we randomly removed the edges in the hospital network to make the average number of contacts per vertex in the hospital network the same as that in the high school in 2011 (and also in 2012) as much as possible. The ratio of the number of removed edges to the total number of edges is γ. We defined a reduced network as a contact data randomly removed edges from the original data. We first compared the hospital and the high school in 2011. When $\gamma = 0.345$, the number of edges of the reduced hospital network is 21,225. Then, the average number of contacts in a day of the reduced the hospital network $(\gamma = 0.345)$ is equal to the average number of contacts in a day of the high school in 2011. Next, we also compared the average number of contacts in a day of hospital and the average number of contacts in a day of high school in 2012. When $\gamma = 0.587$, the number of edges of the reduced hospital network is 13,407. Then, the average number of the contacts in a day of the reduced hospital network ($\gamma = 0.587$) is equal to that of the high school in 2012. We conducted the same experiments as those in Section **4.1**, however we used the reduced hospital network instead of the original data collected in the hospital.

From Fig. 2, the value of the ratio of the number of infected persons at time T, $P(T)$, of the reduced hospital network are still higher than those of the high school in all values of the infectious rate r. In addition, $P(T)$ of the reduced hospital network increases more quickly than the high school network in $0 \leq r \leq 0.03$. Figure 3 shows that $P(T)$ of the reduced hospital network increases when the value of r increases in $0.00025 \leq r \leq 0.005$. However, $P(T)$ of the high school in both 2011 and 2012 increase just a little where $0.00025 \leq r \leq 0.005$. From these results, the infectious disease likely spread on the reduced hospital networks compared to the high school networks. We also investigate that how the spread of the infectious disease is suppressed by the number of contacts decreases. Figure 4 shows that $P(T)$ of the hospital decreases when the number of contacts decreases. These results show that the number of contacts relates to the spread of the infectious disease.

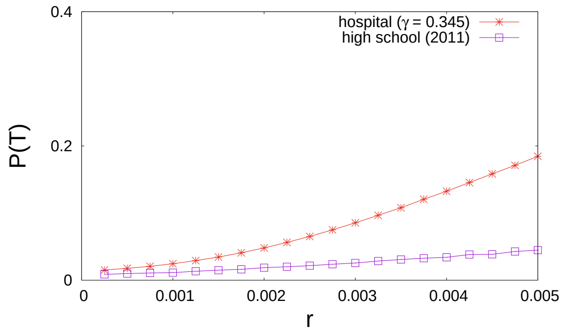


(a)

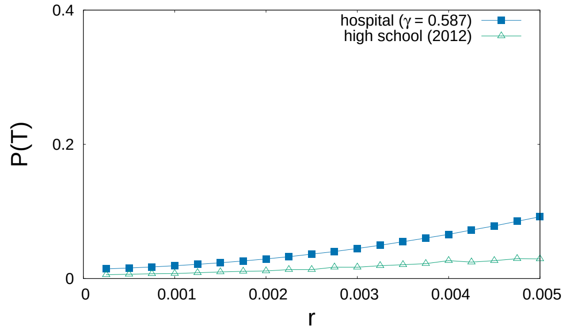


(b)

Figure 2: The results of the ratio of the number of infected persons at time T, $P(T)$ as a function of infection rate r for (a) the hospital network ($\gamma = 0.345$) and the high school network (2011), (b) the reduced hospital network ($\gamma = 0.587$) and the high school network (2012).



(a)



(b)

Figure 3: The results of the ratio of the number of infected persons at time T, $P(T)$ as a function of infection rate r for (a) the hospital network ($\gamma = 0.345$) and the high school network (2011), (b) the reduced hospital network ($\gamma = 0.587$) and the high school network (2012).

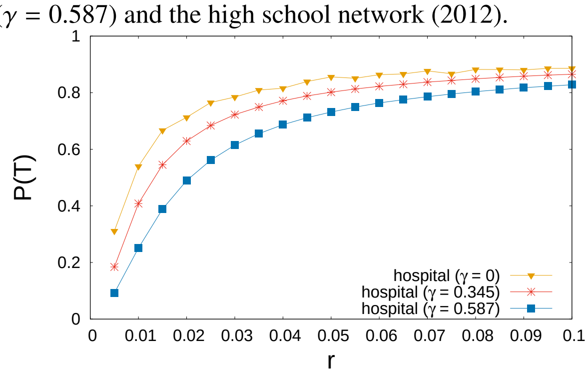


Figure 4: The results of the ratio of the number of infected persons at time T, $P(T)$ as a function of infection rate r for the hospital network ($\gamma = 0$), the reduced hospital networks ($\gamma = 0.355$ and $\gamma = 0.587$).

4.3. Clustering coefficient, degree correlation and contact correlation

We next investigated what structural properties of temporal networks affect the spread of the infectious disease. Table 2 shows the results of clustering coefficients [3]. The clustering coefficient is normalized by the clustering coeeficient from the original network devided the clustering coefficient from the randomized network. When the clustering coefficient is high, it is natural to expect that the infectious disease spreads widely, because the infected ver-

Table 2: The clustering coefficient, the degree correlation and the contact correlation of static contact networks whose edges are summarized temporally.

	hospital	hospital (γ = 0.355)	hospital (γ = 0.587)	highschool2011	highschool2012
clustering coefficient	1.84	1.84	1.90	2.71	3.35
degree correlation	-0.0540	-0.0500	-0.0447	0.1527	0.0495
contact correlation	-0.0637	-0.0608	-0.0586	0.0623	0.1380

tices belonging to a cluster, in which vertices mutually connects to each other, can easily transmit the infectious disease to other vertices belonging to the same cluster. However, against our expectation, the value of the clustering coefficient of the high school is higher than that of the hospital from Table 2. Then, we focused on the degree correlation [10] and the contact correlation. The degree correlation measures the correlation between degrees of two connected vertices. The degree correlation R is described as follows

$$R = \frac{4M \sum_{v,v' \in E} k_v k_{v'} - \left[\sum_{v,v' \in E} (k_v + k_{v'}) \right]^2}{2M \sum_{v,v' \in E} (k_v^2 + k_{v'}^2) - \left[\sum_{v,v' \in E} (k_v + k_{v}^{'}) \right]^2}, \quad (1)$$

where E is a set of edges, v and $v^{'}$ are vertices, M is the total number of edges and k_v is the degree of v. When the degree correlation is negative, the vertex with high (low) degree is likely to connect to the vertex with low (high) degree. When the degree correlation is positive, the vertex with high (low) degree is likely to connect to the vertex with high (low) degree. From Table 2, the degree correlation of the hospital network is negative, but that of the high school is positive. In case that the degree correlation is positive, if the degree of the first infected person is low, the probability that the infectious disease is transmitted to the vertices with high degrees is low. On the other hand, in case that the degree correlation is negative, if the degree of the first infected person is low, the probability that the infectious disease is transmitted to the vertices with high degree is high. From these results, the infectious disease might spread widely in the hospital network.

We also calculated the contact correlation defined by

$$R_c = \frac{4M \sum_{v,v' \in E} c_v c_{v'} - \left[\sum_{v,v' \in E} (c_v + c_{v'}) \right]^2}{2M \sum_{v,v' \in E} (c_v^2 + c_{v'}^2) - \left[\sum_{v,v' \in E} (c_v + c_{v}^{'}) \right]^2}, \quad (2)$$

where c_v is the number of contacts of the vertex v. From Table 2, R and R_c of the network of the hospital are negative, however that of the high school are positive. From these results we expect that the degree correlation and the contact correlation play a crucial role to the spread of the infectious disease on the temporal network.

5. Conclusion

We constructed temporal networks by using the data of person-to-person interactions recorded by RFID [2]. We investigated that the spread of infectious disease on the temporal networks by using a simple mathematical model of the infectious disease on the temporal networks. We further investigated what properties of temporal networks affect the spread of the infectious disease by comparing the structural measures of the hospital and the high school networks. As the result, the infectious disease likely spread easily on the networks when the contact number is high, degree correlation is negative and contact correlation is negative.

Acknowledgment

The authors would like to thank Mr. S. Ogawa and AGS Corp. for their kind encouragements on this research. The research of Y.S. is supported by Grant-in-Aid for Research Activity Start-up (No. 26880020) from JSPS. The research of K.F. was partially supported by Grant-in-Aid for Challenging Exploratory Research (No 15K12137) from MEXT of Japan. The research of T.I. was partially supported by Grant-in-Aid for Exploratory Research (No. 24650116) from JSPS.

References

[1] U. Alon, "An introduction to systems biology - design principles of biological circuits," CRC Press, 2006.

[2] A. Barrat, C. Cattuto, V. Colizza, J-S. Pinton, W. V. Broeck, A. Vespignani, *PLOS ONE*, **5**, e11596, 2010.

[3] D. J. Watts and S. H. Strogatz, *Nature*, **393**, 440–442, 1998.

[4] A-L. Barabási, and R. Albert, *Science*, **286**, 509–512, 1999.

[5] D. Centola, *Science*, **329**, 1194–1197, 2010.

[6] P. Holme and J. Saramäki, *Physics Reports*, **519**, 97–125, 2012.

[7] http://www.sociopatterns.org/

[8] P. Vanhems, A. Barrat, C. Cattuto, J-S. Pinton, N. Khanafer, C. Régis, B. Kim, B. Comte, N. Voirin, *PLOS ONE*, **8**, e73970, 2013.

[9] J. Fournet and A. Barrat, *PLOS ONE*, **9**, e107878, 2014.

[10] M. E. J. Newman, *Physical review letters*, **89**, 208701, 2002.