MobiCom 2010 Poster: A Social Community Based Approach for Reducing the Propagation of Infectious Diseases in Healthcare

Jie Yang[a]   **Yanzhi Ren**[a]   **Yingying Chen**[a]   **Mooi Choo Chuah**[b]

jyang@stevens.edu   yren2@stevens.edu   yingying.chen@stevens.edu   chuah@cse.lehigh.edu

[a]Department of Electrical and Computer Engineering, Stevens Institute of Technology
[b]Department of Computer Science and Engineering, Lehigh University

In this work, we investigate the feasibility and effectiveness of using human contact traces collected from mobile phones to derive social community information to reduce the disease propagation rate in the healthcare domain. Our trace-driven simulation results showed that our scheme is highly effective, and thus opens more opportunities for utilizing social relationships information to support healthcare related applications.

I. Introduction

The recent years have witnessed an explosion of the usage of mobile sensors in people's daily life. With the rapid development of mobile sensing technology, the collected sensing data can be comprehensive enough for mining of human behaviors. For instance, it is convenient to exploit the mobile devices equipped with Bluetooth technology to discover the encounter events between people such that their social relationships can be derived and analyzed. The discovered social relationships can be used to extract social communities to assist in the development of a variety of applications, especially in healthcare.

In the healthcare domain, the transmission of an infectious disease in the public places is a severe problem related to life or death and can cause panic in the whole society if not handled properly. Due to the typical characteristic of a slow start and then exponential propagation of the disease [2, 5], mitigating the infectious disease at its early stage is critical. Since many biological diseases propagate through human interaction, the social communities derived from people's daily life may be utilized to choose the set of people that need to be vaccinated or alerted such that we can reduce the rate of a disease propagation effectively and economically.

Existing work addresses the problem of infectious disease propagation by modeling the disease spread over populations and the control measures [2]. We are not aware of any prior work that exploits the social relationships to reduce the biological disease propagation rate. In our work, we design a social community-based method that exploits the social relationships to reduce the rate at which a new disease spreads. Based on the people's encounter events, multiple community are derived and kernel structures are constructed. We believe those people within the same communities or kernel structures as the sick people have higher risks to get infected, and thus should be given disease alert messages or receive vaccine shots if available. We experimentally tested our strategy by using the MIT reality mining trace [1] through a trace-drive approach. The results showed that our strategy is effective when helping to reduce the rate of the new disease spreading, indicating the initial success of exploiting the social community information for healthcare related applications.

II. System Overview

We assume there are two types of messages to be distributed to people to prevent them from being infected: *vaccination* and *alert*. The person who receives the vaccination massage will go to get vaccine shots so that to immunize himself from the disease. Whereas the person who gets the alert message will be provided with precautions he can take to reduce his chances of being infected. This effect is captured by a lower infection probability in our work.

We consider a biological disease propagation model with three states[5]: *susceptible*, *infectious* and *immunized*.

Susceptible: A person at this state does not have the disease but he is vulnerable to the disease. Thus, the person can be further categorized into two sub-states: (i) he does not receive a disease alert message and stays at *Susceptible without alert message* or (ii) he receives a disease alert message and goes to *Susceptible with alert message*.

Infectious: A person in this state is the host of the disease virus and can infect other people who is within the close proximity with him.

Immunized: Those in this state do not have the disease and they have been immunized from being infected with the disease.
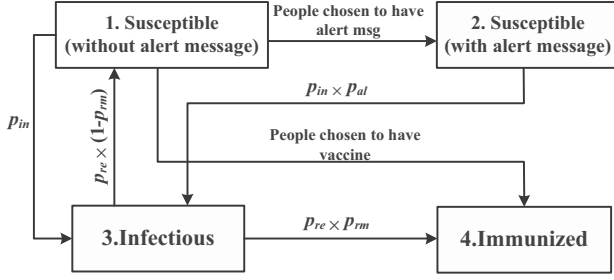
Figure 1: Disease propagation model.

Figure 1 shows the state transition diagram of an infectious disease. The probability of transition from the *Susceptible without alert message* state to the *Infectious* state is defined as p_{in}, whereas it is $p_{in} \times p_{al}$ from *Susceptible with alert message* to *Infectious*. Further, we define the probability of recovery from the disease as p_{re} after every recovery cycle. We note that the *Infectious* state can transfer to either the *Immunized* state with probability $p_{re} \times p_{im}$ or *Susceptible without alert message* with probability $p_{re} \times (1 - p_{im})$ after the person recovered from the disease.

We next present our proposed community-based method which utilizes the extracted communities from human contact graphs to reduce the rate of a disease propagation and an existing encounter-based method for comparison.

Community-based: The human contact graphs can be constructed at a server through encounter traces collected by mobile sensors, e.g., a mobile phone with bluetooth capability. People will be grouped into different communities based on the derived contact graphs. Those who have close relationships, such as family members and close friends, will be further mapped into special communities called kernel structures. Intuitively, the one who has a close relationship or frequently encounters with the sick people has the highest risk to be infected. Once an infectious disease is discovered in public, the community and kernel structure information will be downloaded to each person's device. And then the alert messages and vaccine massages are sent by the sick person to the people within its same community or kernel, who are supposed to have higher risk to get infected.

Encounter-based: We compare our scheme to the existing approach [2] of using human encounters to send alert messages and vaccination messages when a new disease is found in public places. In the encounter-based method, alert messages and vaccination messages will be randomly sent out by the sick person to those the sick person encountering. For comparison, we define the probability that a sick person sends out a vaccine message to those he encounters as p_{va}.

III. Community Based Method

In this section, we present the details on how to extract social community information from human contact traces collected by mobile devices and use it to reduce the rate of a disease propagation.

Contact Graph Construction: A human contact graph is a geometric representation of the relationship between people by counting the encounter events between them. In particular, during the time period $[T_i, T_{i+1}]$, a contact graph $G = (V, E)$ consists of a set of vertices V and a set of edges E. Each vertex $u \in V$ denotes a person, while each edge $e(u, v)$ denotes that person u has contacted person v for at least W times. The weight $t(u, v)$ denotes the closeness between person u and v during $[T_i, T_{i+1}]$, and is defined as the number of times that two persons have encountered with each other.

Community Extraction: The concepts of social community and kernel structure represent the close relationship or similar behavior patterns among people. To construct the contact graphs, we set $W = w_1$ and $W = w_2$ for extracting the communities and the kernel structures with $w_2 > w_1$, respectively.

In general, people may belong to different social communities during different time periods. Furthermore, the same social communities can reappear again and again in the daily life of the community members. Thus, in our proposed community extraction method, instead of directly extracting communities from the contact graphs, we first extract the communities for each non-overlapping time period. We then merge those communities with high similarity.

Particularly, we construct the contact graph $G_1, G_2, ..., G_D$, for each non-overlapping time period: $[T_0, T_1]$, $[T_1, T_2]$,...,$[T_{D-1}, T_D]$. Then, we divide each contact graph G_i into multiple communities by using the hierarchical clustering algorithm [3] and the modularity Q [4]. We then take a snapshot of the communities in each time period. Thus, a total of R snapshots are collected: $S_1, S_2, ..., S_R$. Each S_i contains a vertex set A_i. We assume that each A_i has been divided into k_i communities, which are represented as follows:

$$A_i = A_i^1, A_i^2, ..., A_i^{k_i} \qquad (1)$$

Then, we merge two extracted communities A_i^j and A_{i+1}^l if they satisfy:

$$\frac{|A_i^j \cap A_{i+1}^l|}{Max(|A_i^j|, |A_{i+1}^l|)} > \tau \qquad (2)$$

The threshold τ is set to 0.7 during our experiments. Finally, the communities and the kernel structures

are extracted as: $C = \{C_1, C_2, ..., C_Z\}$ and $K = \{K_1, K_2, ..., K_Q\}$ respectively.

Message Distribution: When the total number of vaccine is limited, distribution of alert and vaccination messages is an important factor to evaluate the effectiveness of our strategy in reducing the propagation rate of the disease. In our proposed method, when a new disease is discovered in public, those who are in the same kernel or community structures as the sick people should be identified as V_k and V_c, respectively. We assign each person in V_k or V_c a weight that reflects how high a priority we should give to this person to receive vaccine (for V_k) or alert (for V_c) messages. For each person, v_m, in V_k, we assign a weight which is equal to the total number of nodes in the kernel structures that this person belongs to as his weight. Similarly, for each person, v_m, in V_c, we assign a weight which is equal to the total number of nodes in the communities that this person belong to as his weight. Then, we give the vaccines to the top P people with the highest weights in V_k and send the alert messages to the top L people with the highest weights in V_c.

IV. Experimental Evaluation

To evaluate the effectiveness of our proposed community-based scheme, we used MIT reality trace [1], which contains human encounter events from 75 people by using Nokia 6600 smart phones. We used the traces of the Bluetooth devices in proximity to find out the encounter information between people. In particular, we used the first half of the trace (i.e., 10 days) as training data to extract the communities and kernel structures and the second half trace as the testing data to evaluate our approach.

Figure 2 presents the performance comparison between the community-based scheme and the encounter-based scheme in terms of the final ratio of the infected persons versus the initial ratio of the sick persons under different infection probability. We observed that the final infected ratio increases when we increase the infection ratio p_{in} for both community-based and encounter-based methods. The key observation is that the community-based method can achieve a lower final infection ratio than encounter-based method consistently. Specifically, for community-based method, when we increase the initial infected ratio from 0.03 to 0.3, the final infected ratio increases from 0 to 0.3 in scenario 1, whereas it is 0 to 0.4 in scenario 2. However, the final infected ratio is around 0.4 in scenario 1 and about 0.5 for Scenario 2 when using the encountered base approach. The experimental results show

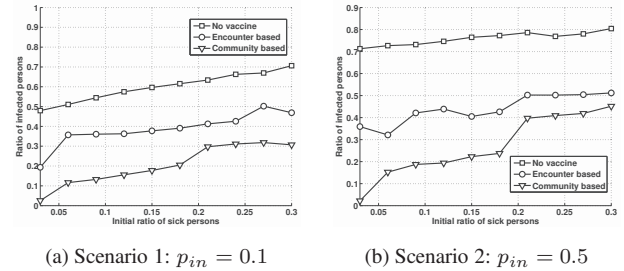(a) Scenario 1: $p_{in} = 0.1$ (b) Scenario 2: $p_{in} = 0.5$

Figure 2: Performance comparison between community-based and encounter-based schemes under different infection probability when $p_{al} = 0.7$, $p_{re} = 0.1$, $p_{im} = 0.2$ and $p_{va} = 0.3$.

that our proposed strategy is more effective in reducing the rate at which a disease spreads than the existing encounter-based method.

V. Conclusion

In this work, we explored the feasibility of utilizing the social community information extracted from human contact traces to more effectively reducing the rate at which infectious diseases spread in healthcare. The social communities and kernel structures are derived based on human encounter events. We believe that those people within the same communities or kernel structures as the sick people have higher risks to get infected, and thus should be given disease alert messages and even receive vaccine shots if available. We experimentally compared our community-based scheme with an existing encountered-based approach using the MIT reality trace. Our trace-driven simulation results showed that our scheme is highly effective, and thus showing a promising direction of utilizing social relationship information derived from mobile devices to support a broad range of healthcare related applications.

References

[1] N. Eagle and A. Pentland. Reality mining: Sensing complex social systems. *Personal and Ubiquitous Computing*, 10(4), 2005.

[2] S. Huang. Probabilistic model checking of disease spread and prevention. In *Scholarly Paper for the Degree of Masters in University of Maryland*.

[3] J.Scott. *Social Network Analysis: A Handbook*. Sage Publication Ltd, 2000.

[4] L. Tang, X. Wang, and H. Liu. Uncovering groups via heterogeneous interaction analysis. In *ICDM'09*, Dec, 2009.

[5] H. Yuan, G. Chen, J. Wu, and H. Xiong. Towards controlling virus propagation in information system with point-to-group information sharing. *Decision Support Systems*, 48(1):57–68, 2009.