Epidemic Modeling using Mobile Phones

Moritz Köhler
Institute for Pervasive Computing
ETH Zürich, 8092 Zürich
Switzerland
mkoehler@ethz.ch

Vlad M. Trifa
Institute for Pervasive Computing
ETH Zürich, 8092 Zürich
Switzerland
trifa@inf.ethz.ch

ABSTRACT
Infectious diseases are famous for their potential to turn into epidemics if the appropriate measures are not taken early enough, and the best strategy to counter epidemics is to rely on the early detection and identification of the infectious agents involved to take appropriate measures. In our work we want to investigate if it is possible to build disease spreading models using sensory data gathered by mobile phones from a large group of participants. We then want to use these or other models to return valuable information to users regarding their health status.

Categories and Subject Descriptors
H.4.3 [**Information Systems Applications**]: Communications Applications

Keywords
Mobile health care, participatory research, infectious disease modeling

1. INTRODUCTION
Outbreaks of infectious diseases can have a tremendous impact on a population if sufficient measures are not taken early enough. With the advent of computers, complex models of disease spreading have been created to help policy makers to prepare for future epidemics by simulating the impact and effectiveness of different control methods and novel technologies. However, many factors such as heterogeneity of populations and the spatial structure of disease spreading are difficult to model accurately, thus often limiting the usefulness of these models. The general trend toward miniaturization of efficient bio-sensors, with increased sensitivity and robustness, along with the large adoption of personal communication networks (e.g., mobile phones) is a promising context to foster the development of cost-effective mobile devices that can rapidly detect, identify and characterize potential threats. A tremendous advantage is that public health can be monitored at all times, thus such a system would not only be useful where diagnosis of presymptomatic or asymptomatic infections can lead to more rapid, successful and cost-effective treatment, but also facilitate preventive actions on the "healthy" population.

As timing is a critical issue, a system for real-time monitoring of the health of a population, where data can be collected quickly enough, is necessary to generate accurate predictions of the status of epidemics which could be used to optimize the allocation of available resources.

The advantage of providing mobile phone users with a cheap and practical tool for medical self-diagnosis, is not only to offer them on-the-fly feedback and suggestions about their health status, but also to offer results to public health agencies by correlating this information at a global level, over huge populations and across discrete and localized communities. Due to today's increase in international travel, movement of people and their individual interactions both need to be considered in order to define more precisely the transmission vectors of a disease.

2. RELATED WORK
Rasid et al. present in [6] a telemedicine processor for mobile phones. It allows for transmission of biomedical data to remote locations such as hospitals. Their system is currently being deployed and their goal is to transmit data readings to health specialists located in India [1]. In our work we go beyond the 1:n relationship (one doctor serves many users) to a n:n (many users serve many users) approach.

In regards to disease spreading behavior, Murray gives an excellent introduction into modeling the dynamics of infectious diseases [4] and their geographic spread [5].

3. SOLUTION APPROACH
Two different levels are to be considered. On the one hand the technologies that can be easily integrated into mobile phones, on the other hand how local information and interactions could be combined to obtain a real-time monitoring of health status at the population level.

3.1 Technology
Our system model is shown in Figure 1. The health status of each participant is estimated based on sensor data, for example by embedding temperature or blood pressure sensors attached to a mobile phone or to the head-set for detecting anomalies of vital functions. Voice signals can be analyzed

to detect coughing or find out if the person is hoarsely. In the near future, on-body sensors or implanted bio-sensor could connect to the mobile phone to transmit their sensor readings wirelessly.

3.2 Interaction

To move on to a global view, we propose to integrate interactions between individuals to generate a *Personal Contact Graph*. This graph models spacio-temporal relationships between participants, by applying machine learning or data mining algorithms on information such as location, phone call history, or communication range detection, e.g., by using Bluetooth discovery.

Together, the individual health status of each participant, and the global relationships between participants then can be used to model epidemics at a global scale. For example, given enough data, a spreading model could be created or existing models could be improved. On the other hand, given a certain epidemic model, it could be tested against the data, which could be used for disease classification.

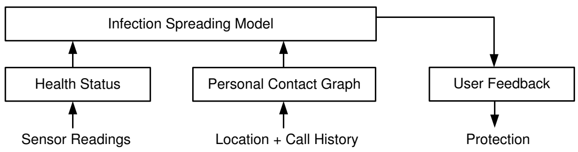

Figure 1: A general framework to enable modeling of infectious disease spreading and giving feedback to the user

In any case, the global system view can be used to refine each participant's individual health status assessment. This refined information then can be returned to the user in form of status reports, or action suggestions (e.g., *don't go in this area*, or *you should see a doctor because you have high blood pressure*).

3.3 Level of Participation

Participatory research is fairly new in Ubicomp in general and sensor networks in particular. From Ubicomp we can use a lot of results from Participatory Design and deployment of technology for the user [3, 2]. From these results, we know that acceptance of our technology will be crucial to the overall success of the research effort. For the level of engagement we envision empowerment of the participants, as they will receive important behavioral information upon which they can individually decide on how to react. However, we do not see a way for group decision making, i.e., each individual will be unaware of the consequences of his decision for the group.

3.4 Deployment strategy

Given that we can only expect a very limited amount of participants for the initial deployment, i.e., less than 20, we need a clear deployment strategy for our system, which we will discuss here.

Incentive model. Our incentive model will be twofold: First, we want to give users feedback on their current health status. Second, we aim to give them recommendations regarding their future action, e.g., "What is the probability to get the flue when you go to a particular sports event?". These two forms of feedback will be aligned in time, starting with feedback on health status, and then only in a later phase giving recommendations. The reason for this two phase incentive model, can be seen in Figure 1. Here, we want to feed our prediction models with data generated by users. This constitutes a critical mass problem, which we can only solve, once we have enough people participating. Due to the chasm of people available in the beginning and people necessary for our prediction models, we need to give them an initial *local* feedback which has value to every individual user, regardless of the size of the participating group.

Regarding local feedback, our idea is to embed a temperature sensor into a Bluetooth headsets, as such sensors are very common, small, inexpensive, and they would sense the temperature inside the ear directly, where it can be accurately estimated. This information then can be used for both, phase 1 local feedback and to feed the global prediction model.

Recruitment mechanism. In the beginning, in particular when prototyping hardware and algorithms, we aim to only work with people familiar with the matter, like PhD students from our lab. Therefore recruitment will be on a personal face-to-face basis. Later we hope to get funding from a mobile phone provider or producer in order to deliver subsidized mobile phones to volunteering participants. In the third phase we already hope to have a relatively large user base and having the network effect jump in which would help to make our system known by word of mouth.

Necessary size. The necessary size of the participating group heavily depends on its geographical distribution. According to [4] for the influenza virus a group of around 700 individuals is large enough to model the spreading behavior of the disease, if this group has limited geographic expansion[1].

3.5 Scenario

In our work we want to focus on influenza or similar virus diseases. The reason for this is that the disease occurs frequently, it spreads easily, it is well studied (i.e., models exist), symptoms are easily detectable by sensors, and it is rarely lethal, which can prevent us from legal problems in case the system does not work properly. Regarding the infection spreading model, [4] provides and excellent but still simple model for the influenza virus. We will study this and other model more in detail before implementing a first prototype.

In brief in such a model the population can be divided into three distinct classes: the susceptibles, S, who can catch the disease; the infectives, I, who have the disease and can transmit it; and the removed class, R, namely, those who have either had the disease, or are recovered, immune or isolated

[1] In concrete it is referred to a 1978 report with detailed statistics of a flu epidemic in a English boys' boarding school with a total of 763 boys.

until recovery. The progress of individuals is schematically represented by

$$S \rightarrow I \rightarrow R.$$

This model is called an SIR model. Although there are more detailed models, e.g., $SEIR$ models which introduce a fourth class E, in which the disease is latent, we belief that for our initial research the SIR model is sufficient. Furthermore adapting the model is a rather simple task once the sensor are mounted on mobile phones and communication with a back-end server is established.

4. DISCUSSION

We have motivated the idea of health monitoring and epidemic modeling using mobile phones and described one approach for this problem. But there remain several open issues we want to discuss briefly.

4.1 Technical issues

Due to the complex and dynamic environments in which users operate their mobile phones, sensory data can be noisy and bad readings have to be compensated in order to generate reliable estimations of the health status of an individual/population. The information collected with off-the-shelf sensors will remain only indicative and will never replace a doctor. Thus studies need to be conducted to assert if the reliability of such diagnosis tools is sufficient in order to be useful.

Especially in big cities with many people and with low localization accuracy, it will be problematic to find correct connections between people necessary for the *Personal Contact Graph*.

4.2 Legal issues

Since we will be dealing with sensitive health data, privacy issues need to be addressed. Our system will need to guarantee only authorized access to information. E.g., patients must only have access to their own data. In regards to direct local feedback such as body temperature, we do not see an urgent problem of privacy offenses.

However, the information we want to give back to participants has to be analyzed thoroughly under privacy concerns. This information is generated from data gathered from participating users. Therefore, we must ensure that it is trapdoor like, and no direct or indirect inference to any participant is possible.

4.3 General issues

It will be very important to find out, what kind of feedback users will appreciate, e.g., we need to evaluate if the mobile phone itself will be best used as a output channel or regular desktop computer based browsers. Also the level of user participation still is not totally defined.

Another very important issue will be the minimum size of the participating group. Since sensor readings only can result in a certain probability for the health status, we need a relatively large participant group of around 500-1.000 individuals (see section 3.4). Due to the involved costs and the market conditions, we have identified certain stakeholders we wish to involve at some point during the project. These include mobile phone producers, network operators, and government agencies.

5. REFERENCES

[1] Loughborough university takes mobile phone health monitoring to india. http://tinyurl.com/2q9ms7, 2007.

[2] G. D. Abowd. Evaluation of real deployments in ubicomp. Tutorials at the 5th International Conference on Pervasive Computing, 2007.

[3] E. H. et al. Ubicomp in opportunity spaces: Challenges for participatory design. In *Proceedings of the ninth Participatory Design Conference*, 2006.

[4] J. Murray. *Mathematical Biology*, volume I of *Biomathematics*, chapter 10. Springer, New York, 2002.

[5] J. Murray. *Mathematical Biology*, volume II of *Biomathematics*, chapter 13. Springer, New York, 2002.

[6] M. F. A. Rasid and B. Woodward. Bluetooth telemedicine processor for multichannel biomedical signal transmission via mobile cellular networks. *IEEE Transactions on Information Technology in Biomedicine*, 9(1):35–43, 2005.