Epidemic spreading over social networks using large-scale biosensors: a Survey

João Andrade and Artur Arsenio*

Instituto Superior Técnico / Universidade Técnica de Lisboa
Taguspark campus, Room 2N3.09
Av. Prof. Dr. Aníbal Cavaco Silva, 2780-990 Porto Salvo, Portugal
{joao.andrade,Artur.arsenio}@ist.utl.pt

Abstract

Recent technological developments on mobile technologies allied with the growing computational capabilities of sensing enabled devices have given rise to mobile sensing systems that can target community level problems. These systems are capable of inferring intelligence from acquired raw sensed data, through the use of data mining and machine learning techniques. However, due to their recent advent, associated issues remain to be solved in a systematized way. Various areas can benefit from these initiatives, with public health systems having a major application gain. There has been interest in the use of social networks as a mean of epidemic prediction. Still, the integration between large-scale sensor networks and these initiatives, required to achieve seamless epidemic detection and prediction, is yet to be achieved. In this context, it is essential to review systems applied to epidemic prediction.

Corresponding author. Tel.: +351-21-4233552; fax: +351-214233268.
E-mail address: artur.arsenio@ist.utl.pt

Keywords: biosensors; epidemic spreading; vírus; social networks; large-scale sensing; survey
1. Introduction

Distributed systems have been used as a platform to allow the interaction between groups of individuals and a set of devices. As technology advances in sensing, computation, storage and communications become widespread, ubiquitous sensing devices will become a part of global distributed sensing systems [1][2].

Recently, the predominance of mobile phones equipped with sensors, the explosion in social networks and the deployment of sensor networks have created an enormous digital footprint that can be harnessed [3]. Furthermore, developments in sensor technology, communications and semantic processing, allow the coordination of a large network of devices and large dataset processing with intelligent data analysis [1].

The sensing of people constitutes a new application domain that broadens the traditional sensor network scope of environmental and infrastructure monitoring. People become the carriers of sensing devices and both producers and consumers of events [4]. As a consequence, the recent interest by the industry in open programming platforms and software distribution channels is accelerating the development of people-centric sensing applications and systems [4][1].

To take advantage of these emerging networks of mobile people-centric sensing devices, researchers arrived at the concept of Mobiscopes, i.e. taskable mobile sensing systems that are capable of high coverage. They represent a new type of infrastructure, where mobile sensors have the potential to logically belonging to more than one network, while being physically attached to their carriers [5]. By taking advantage of these systems, it will be possible to mine and run computations on enormous amounts of data from a very large number of users [1]. People-centric sensing enables therefore a different approach to sensing, learning, visualizing and data sharing, not only self-centered, but especially focused on the surrounding world. The traditional view on mesh sensor networks is combined with one where people (carrying sensors) turn opportunistic coverage into a reality [2]. These sensors can reach into regions whereas static sensors cannot, proving to be especially useful for applications that occasionally require sensing [5]. By employing these systems, one can aim to revolutionize the field of context-aware computing [3].

An alternative of a worldwide coverage of static sensors to develop people-centric systems is unfeasible in terms of monetary costs, management and permissions [6] [2]. Also, it is extremely challenging in static sensing models, due to band limits and issues that arise from covering a vast area, to satisfy the required density requirements [5]. Thanks to their mobility, mobile sensors overcome spatial coverage limitations [5][6]. Adoption issues might come up, as potential users are usually unaware of the benefits that arise from technological developments [1].

The modeling of behavior requires large amounts of accurately labeled training data [7]. These systems constitute an opportunity for machine learning systems, as relevant information can be obtained from large-scale sensory data and employed in statistical models [7] [1]. Great benefits can be taken from this unconstrained human data, in opposition to the traditional carefully setup experiments [7]. With these developments it is now possible to distribute and run experiments in a worldwide population rather than in a small laboratory controlled study [1].

By leveraging the behavioral patterns related to individuals, groups and society, a new multidisciplinary field is created: Social and Community Intelligence (SCI) [3]. Real-time user contributed data is invaluable to address community-level problems and provide universal access to information, contributing to the emergence of innovative services [3][2][1], such as the prediction and tracking of epidemic outbreaks across populations [3]. Thus, technological benefits are shifted from a restricted group of scientists to the whole society [2].

Such systems can be applied to Healthcare, to facilitate both monitoring and sharing of automatically gathered health data [2]. Epidemics are a major public health concern and it has been shown impact can be reduced by early detection of the disease activity. For instance, it has been shown that the level of influenza-
like illness in regions of the US can be estimated with a reporting lag of one day, when compared to clinical methods whose results take a week to be published [3].

The advent of ubiquitous networks of mobile sensing devices constitute a paradigm shift, offering researchers challenges in network architecture, protocol design and data abstractions [5]. Results from mobile sensor networks, pervasive computing, machine learning and data mining can be exploited, however, new unaddressed challenges arise. These challenges range from growing volumes of multimodal sensor data, dynamic operating conditions and the increasing mobility of the sensing devices [1]. As most people possess sensing-enabled phones, the main obstacle in this area is not the lack of an infrastructure. Rather, the technical barriers are related to performing privacy and resource respecting inference, while supplying users and communities with useful feedback [1].

The main challenges in this area are as follows:

- Managing user participation (participatory versus opportunistic sensing) [1];
- Managing trust in users as not to compromise the whole system [1];
- Coping with sensing device mobility (e.g. lack of mobile sensors or the sensor is moving and jeopardizing sampling for a given context) [6] [2];
- Enabling devices to share sensory data while protecting user privacy [6] [1];
- Relating and optimizing diverse resources and application-relevant metrics to define data collection and dissemination methods [1] [5];
- Managing large amounts of generated data [1];
- Performing robust and accurate activity classification in a dynamic real-time environment [1] [5];
- Sensing system scaling from a personal to a population scale [1];

On the other hand, in an epidemic it is necessary to detect, monitor and foresee the evolution of disease spreading. To operate in such a scenario the system should know who is infected and which people have been in contact and where. Contact location, time and relationship with the subject are relevant metrics that affect the probability of disease propagation. Biosensors and intelligent networks allow the integration of these concerns into personal devices, while developments in data mining and modeling allow a more accurate analysis of this data potentially indicating new disease outbreaks and estimating their impact.

This paper aims therefore to provide a survey for previous research work on epidemic disease propagation prediction by applying intelligent analysis methods to the large-scale data sourced from the users and their environment. We will start by reviewing on section 2 important issues for large-scale wireless sensor networks. Afterwards we will review opportunistic communication systems capable of evaluating and predicting epidemic disease propagation throughout a community (section 3), looking at the integration of social network platforms as data sources, and on the usage of machine learning algorithms, taking into consideration information security (section 4). Section 5 discusses applications and presents the study conclusions.

2. Pervasive Computing

There is a tendency to augment devices with sensing, computing and communication functionalities, connecting them together to form a network, and make use of their collective capabilities [3]. This sensing network is made possible across various ranges, including: single individuals, groups with common interests and the entire population of a city [1].

Users become a key system component, enabling a variety of new application area such as personal, social and community sensing. Each of these scenarios has its own challenges on how to understand, visualize and share data with others [2]. In personal sensing, the focus is on monitoring the individual [2]. In these applications, information is generated for the sole consumption of the user and is generally not shared with
others [1]. In social sensing, information is shared within the group [2]. Individuals who participate in these applications have a commonality of interests, forming a group [1].

In community sensing (SCI), data is shared for the greater good of the community. Considering data source origin, SCI has its source in three fast growing research fields: mobile sensor based activity recognition, context inference in smart spaces and social network analysis. The key idea behind mobile sensor based activity recognition is to acquire the mathematical model behind human activities after a series of observations. It takes advantage of the prevalence of sensors that accompany users in their mobility patterns. Context inference in smart spaces relies on already deployed infrastructures of static sensors. Static sensors allow the detection of activities, enabling space context. Social network analysis has been studied by physicists and social scientists for a couple of decades and is a major source of information and relationships among a group of individuals. Aggregation of data from these sources constitutes an opportunity for the extraction of intelligence in a community. Applications only become useful once they have a large enough number of individuals participating. An infrastructure capable of integrating heterogeneous data sources is required [3], combining the resulting multimodal data and extracting behavioral patterns from it, through data mining and machine learning methods.

2.1. Participation

The need exists to define an individual’s role in the sensing system. Two modalities are considered: participatory and opportunistic [3].

In participatory sensing [8][9], individuals are incorporated in the decision making process over the sensed data [3]. They can decide which data to share, enjoying control over data privacy issues. In this approach the target is restricted to a group of users willing to participate in the system [3]. As a consequence, a participatory sensing application should have community appeal [2].

In opportunistic sensing [2][10][11], a system automatically takes advantage a device’s resources whenever its state (e.g. location or user activity) matches the context requirements of the application [3]. Opportunistic sensing becomes possible by the system’s ability to modify its state in response to a dynamic environment [1]. Sampling only occurs if requirements are met and it is fully automated, with individuals having no involvement in the data collection process [2]. A result of this is that the decision burden is shifted away from users and moved into the system, resulting in more resources being demanded in this decision-making process [3] [2]. This heavier resource demand should not noticeably impact the normal usage experience of the sensing devices [2]. This issue can be tackled if opportunistic sensing is considered a secondary, low priority operation on the sensing devices [2]. Nonetheless, as devices might only be able to meet sensing requirements for short and intermittent periods, a trade-off between availability and resource management should be considered [2].

2.2. Context

Context affects data sensing, while sensing devices with mobility can be used in unpredictable ways [1] [11]. Context is the metadata that describes the conditions to which sensors are exposed, affecting both data and sensors’ ability to perform sensing operations. In opportunistic sensing, context contributes to the evaluation of potential sensor candidates, indicating when sampling should be started and stopped [2].

Context is important for analyzing sampled data, especially when samples might be taken under suboptimal conditions [2]. In these environments, statistical models may fail to generalize. Also, sensors may be exposed to events for a too short time period, i.e. if the user is traveling too quickly or the sensor’s
sampling rate is too low [1]. Possible solutions are sharing sensors of available neighboring devices temporarily if they are best suited to sense events [2][1]. Devices exchange context information and data is selected from the device whose context most closely matches application requirements. A mobile sensor detects the event target using its sensors and forwards the task to its better suited neighbors. To recover a lost event source, the area to which the source is predicted to be in is estimated and the task is forwarded to sensors in the predicted area. Another approach is to use super-sampling, where data from nearby sensors is collectively used to lower the noise in an individual reading [1]. One challenge is determining a metric for context matching that provides samples with enough fidelity to respect application requirements [2].

The reliability of machine learning algorithms may decrease under the dynamic and unexpected conditions presented by mobile sensor use (e.g. different individuals execute the same activity differently). These problems can be overcome by gathering sufficient samples of the different usage scenarios, i.e. training data [1]. However, acquiring training data is costly and anticipating the different scenarios that might be encountered is not possible for all applications [1], compromising the scalability of large-scale learning models [11]. Existing solutions are based on borrowing model inputs, i.e. features, from nearby sensors and performing collaborative inference between the associated models. These models might have evolved based on different scenarios, so it is possible to discover new events that were not considered during application design [1]. Other approaches consider a combination of supervised and unsupervised learning techniques, where the learning method to apply depends on data classification stage [11].

2.3. Data

Data producers can be classified in terms of modality (e.g. mobile sensors, static sensors, web services), internet connectivity (e.g. constant, intermittent), privacy sensitivity, and resource capabilities (when data is processed locally). Information consumers are heterogeneous in terms of running environments (applications that run locally or remotely at community level), data needs (high-level information or raw sensor data). This heterogeneity leads to several challenges on data management [3].

Different sensors consider the physical and virtual world with different levels of accuracy. Lack of correlation between data collected from distinct viewpoints and resolutions leads to an ineffective data merge and processing. A sensor may sense the same event under different conditions and classify it differently, yielding inconsistent results. Also, due to environmental differences, a group of sensors in the same location might sense the same event in time and infer different results with the same classification algorithm. For these reasons data needs to be mapped to a shared vocabulary, respecting the same metrics [3].

2.4. Sensing Architecture

Technologies such as Mobiscopes are still recent, leading to a lack of normalized architectures [1]. As these systems have no control over human mobility patterns, the coverage of spaces, events and human interactions becomes opportunistic [2][6]. In order to face mobility, decisions are taken in real-time [5].

Sensing devices enjoy a high degree of heterogeneity. Typically, sensed data has varying time-space resolutions and may become biased depending on the sensing context. Nonetheless, the heterogeneity in sensing data can be harnessed to increase system robustness by exploiting distinct views that may complement each other [5].

Sensing devices have typically resource limitations that require careful consideration as to where data processing takes place [2]. One approach is to persist data by employing local buffering capabilities [5]. However, for analyzing large amounts of data, local storage limitations may require to have persistency on
remote servers [2] [6]. Privacy issues also need to be taken into account, as it may be inappropriate to store sensitive data in a remote untrusted system.

Connectivity issues in the system affect sensing performance. In sensing networks, at a given time, a greater amount of data is gathered when compared to data that can be delivered. To circumvent these issues and avoid resource waste, data prioritization schemes [5], to be used when multiple nodes cover the same area, have been suggested. Opportunistic data diffusion schemes between sensing devices, with possible data aggregation, aim to improve connectivity and data quality despite data incongruences [5]. Since information needed by an application may only be available by integrating data from multiple sensing modalities, transmitted data must be compatible across heterogeneous networks [5].

Machine learning functionalities require a systemic view, considering the sensing devices’ resource constraints, communication costs to remote servers and the sampling rate required to detect and characterize interesting phenomena [2]. There is also a high correlation between data accesses and user location. Because of the time and space dynamic nature of sensor densities, system performance depends on the mobility patterns of the sensing devices. Uniform coverage for a given area is hard to achieve, as sensors tend to visit zones in a given area in a non-uniform fashion. And as interesting events might be rare, sparse data models need to be considered. For such cases data-mining techniques can be applied. Another approach is to have actuated sensing devices, i.e. sensors that are tasked to visit uncovered areas [5].

Some authors have provided a systematic architecture that can be used as a viewpoint to face these issues, consisting of five layers: pervasive sensing, data anonymization, hybrid learning, semantic inference, and application [3]. The pervasive sensing layer involves the gathering of data from the different data sources (mobile devices, static sensors, social web); The data anonymization layer anonymizes sensed data, offering different anonymization algorithms that can be applied according to the nature of the requirements; The hybrid learning layer applies machine-learning and data mining algorithms to convert low-level single-modality sensing data into high-level features or micro-context. Its focus is to mine data patterns and derive behavior and single space context, before multimodal intelligence is extracted; The semantic inference layer is needed when different micro-contexts need to be aggregated. Its objective is to match the inputted micro-contexts with an expected high-level activity; The application layer provides a set of accessible services that are sustained on the other layers. Applications may be installed directly on a mobile sensing device or on remote servers, communicating with the sensors.

3. Computational Epidemiology

Computational epidemiology consists on the development and use of computer models to understand the diffusion of disease through populations with regard to space and time [12].

Data used in these models should be representative in order to accurately predict and understand the propagation of diseases [13]. However, decisions have to be made with limited information. Consequently, an effective prediction is difficult, especially if initial data is not expressive enough [14]. Some systems obtain model data either through periodic online questionnaires [15], trusted web news sources [16] or by exploiting web search queries to monitor health-seeking behavior [17]. Social networks constitute a potential new data source as large-scale relevant user-related data can be acquired instantaneously and in real-time [13].

As a consequence of their capability to estimate disease propagation, these models are powerful tools to evaluate the course of a disease in response to public health interventions [12][18]. The more is understood about infectious disease spreading, the more efficiently it is possible to deploy measures to counter outbreaks, such as vaccines [19][17][15]. The strength of an epidemic can be evaluated with resort to the generalized isoperimetric constant of the associated social contact graph, also known as Cheeger constant. When the ratio
between cures and infections is lower than this constant, an epidemic dies out quickly. Conversely, if it is higher an epidemic will die out slowly [20].

It is important to note that the end of an epidemic is caused by the decline in the number of infected individuals rather than an absolute lack of susceptible subjects. Thus, at the end of an epidemic, not all individuals have recovered.

**Model:** An epidemic model is a mathematical abstraction that describes the evolution of a transmittable disease in a population. It may be modeled as a computational network, which is itself modeled as a graph. A graph consists of a set of points called nodes or vertices. Interconnections between these nodes are named links or edges and, in this application, they represent a form of contact or relation. A node’s degree corresponds to the number of neighbors it has [20].

Various parameters impact model construction. Modeling challenges come from the large size, irregularity and dynamism of the underlying social contact network [12]. Scale-free networks are deemed appropriate to model real social networks due to their inherent large fluctuations between the number of connections in each vertex [20]. In these networks, the final size and persistence time of a given epidemic are highly sensitive to the multi-scale hierarchical structure of the considered population [19]. For instance, nodes that are in contact with a large number of other nodes are easily infected and constitute a bridge for the spreading of infections [19][18]. The epidemic threshold vanishes in these networks, which suggests that even weakly infectious virus can spread [19]. This social network graph may be analyzed using integrated statistical and machine learning methods to produce model input data, effectively representing a labeled social contact network [12].

**Mixing:** Under homogeneous mixing, individuals belonging to the population are neighbors with every other individual, making contact at random and not mixing into smaller subgroups. In this topology, the set of infected individuals has little meaning to the overall population dynamics and the relevant metric is the number of infected individuals [21]. In non-homogeneous mixing, the structure of the considered social network greatly influences disease proliferation as it conditions contact between individuals [12]. One way to accommodate asymmetric and variable contact is by weighting the links of the contact networks. The weighted links distribute the contact rate parameter over the graph. The weight value and distribution can have a significant effect on the epidemic resistance of the topology, offering a possibility to alter a graph without changing its topology. The introduction of weights gives rise to a new form of clustering, i.e. weight clusters. Such clusters can boost infectious agent spread through the network [20].

**Spatial Distribution:** Simple models assume uniform spatial distribution. More complex lattice-based models can cope with non-uniform spatial distributions [21].

**Age Structure:** A rectangular age structure assumes people live to reach the average life expectancy of the population. This model is suitable for developed countries. For other countries a triangular age structure is considered more appropriate [22].

**Genotype:** The genotype of the afflicted population constitutes its inherited genetic information and can determine its vulnerability against a given infectious agent and resistance towards another. The genotype of the infectious agent influences its behavior, infectivity, and resistance to public health measures and may contribute to the appearance of new substrains with different characteristics. The interaction between these two variables conditions epidemic dynamics [18].

**Transmission:** There are two directions in disease progression: within-host progression and between-host transmission. The start of within-host progression is triggered by between-host transmission. There is a latent period between the time an individual becomes infected and the time when the capability to infect others is acquired [12]. Between-host transmission can occur in different directions: horizontal and vertical. Horizontal disease transmission may be triggered through various forms of contact: direct contact; indirect contact (e.g. contact with a contaminated surface); droplet contact (e.g. sneezing), airborne contact (if the pathogen is
resilient enough to survive in the air); fecal-oral contact (e.g. contact with contaminated food or resources) [18]. Vertical disease transmission occurs from mother to child (e.g. in the case of AIDS and Hepatitis B).

**Epidemic Reaction:** The behavior of people is changed in response to the menacing nature of an epidemic. One future direction in computational epidemiology is to take it into account [15].

**Granularity:** A social network node can be defined to represent a single individual or group or a small location or country [20]. Aggregate models assume a population is partitioned into subpopulations with a predictable interaction structure within and between subpopulations. While, these models are useful for obtaining parameters, such as the total number of infections, they lack the capability to capture the complexity of human interactions that serve as a major infectious disease transmission mechanism and are incapable of providing causal explanations. The capability to provide specific details about the flow of disease spread may be required to provide insights to researchers investigating interventions against the epidemic. Also, as the granularity of subpopulations is considered to be high, parameters such as the base reproductive number and the contact rate are hard to observe [12].

Disaggregate models use a representation of individual agents with explicit interactions between them to model the disease spread across social networks, offering a much finer granularity [12].

**Mathematical Formulation:** Epidemiological models can be classified depending on their mathematical formulation [18].

On the **deterministic formulation**, individuals are assumed to be uniformly distributed in space and to mix at a certain rate, i.e. the contact rate. Deterministic epidemiological models are usually based on the Susceptible-Infected-Recovered (SIR) compartmental models. These are sustained on a set of differential equations, partitioning individuals across model-dependent compartments. Asymptotic behavior for resulting systems depends on parameter choices [18]. Nowadays, these models are not very viable as people enjoy a high degree of mobility and can easily travel abroad, carrying a disease with them. This results in complex interactions between individuals and originates complex social networks [15]. One example using this formulation is [23].

On the **stochastic formulation**, the probability distribution of potential outcomes in disease propagation is estimated by allowing input data to vary randomly over time. Systems may be modeled as stochastic discrete event systems (DES), in which the system state only changes upon the occurrence of an event. However, for large populations the use of parallel discrete event systems (PDES), may be considered to exploit resources provided by a set of machines [12]. Also, variants of finite state machines (FSM), called probabilistic timed transition systems (PTTs) may be used to represent within-host disease progression. In this approach, state transitions are probabilistic and timed and the considered states depend upon the considered implementation [12]. The theory of stochastic processes defines the asymptotic behavior for systems using this formulation [18], which is important to model epidemic processes. Examples of systems using it are [24], [25] and [12].

4. **Security**

Respecting the privacy of its users is a relevant concern in a mobile sensing system [1][3]. People are sensitive about how their data is captured and used, especially if it contains their location [1], speech [11], sensitive images [1], or personal records such as private health information. Interestingly, social network application users may take privacy as a less relevant concern [4].

Collected data may inadvertently reveal information about people. For instance, a connection between mobile sensors and observed parties may be implicit in their user’s relationships [5]. Revealing personal data can risk privacy and sharing community gathered data can reveal information on community behaviors [3].

People may fear that personal information will leak from the system. Even individuals that are not sensing targets may be vulnerable to accidental privacy breaches, if they are close to a sensing device [2].
Countermeasures pausing the collection of sensor data, are not suitable as they may cause a noticeable gap in the sensing data stream [1]. Revealing too much context can potentially compromise anonymity and location privacy. Conversely, the inability to associate data with its source can lead to the loss of context, reducing the system’s ability to generate useful information [5].

**Privacy:** Protection involves different variables, including identity (who wants data access), granularity (level of data revealed), and time (retention time of data) [3].

**Authentication:** Deals with validating the user to the system. The sheer amount of users in mobile sensing systems might pose impediments to cryptographics authentication. Nonetheless, there is the possibility of relying in the redundancy of sensor data to validate a source anonymously [5].

**User control:** Control over data sharing allows users to define their participation in the system, empowering the decision making process [3]. One approach is keeping sensitive relations from being exposed, either by local filtering or by providing users with an interface to review data before it is released [5]. In [11], the user has complete control in how information is presented in the different system interfaces.

**Anonymization:** Before data release and processing, different algorithms may be applied with the objective of not revealing the user identity [3]. Some approaches can help with these problems (e.g. cryptography, data and privacy-preserving statistics) [5][1]. Nevertheless, they may be insufficient [1]. In personal sensing, a solution is processing data locally [1][11]. In the context of community sensing, there is the risk of leaking personal and community information. Reconstruction attacks target innocuous-looking data and allow invasive information to be reverse-engineered [1]. A solution is for privacy to be based on group membership. Sensitive information is only shared within the groups in which users have existing trust relationships [2][4].

**Trust:** Ensuring both data sources are valid and that information is accurate, should be a system concern. Also, correct system usage should be promoted to prevent abuses. When mining social and community behaviors, anonymous data is needed [3]. Data correctness must be verified without violating privacy [5]. In opportunistic sensing schemes user trust may become a barrier to wide-scale adoption [2]. These issues may be addressed by providing sensing device users with a notion of anonymity through k-anonymous tasking [2].

5. Conclusions

Millions of people regularly participate within online social networks. In [4] the use of phone sensors to automatically classify events in their lives was investigated. These classifications can be selectively shared using online social networks, replacing manual actions that are now performed daily [4][1].

In [11] a lightweight and scalable hierarchical audio classification system, designed with resource limited mobile phones in mind, while remaining capable of recognizing a broad set of events was provided. In opposition with off-line audio context recognition systems, classification was performed online at a lower computational cost, while yielding comparable results.

Conventional ways of evaluating environmental impact rely on aggregated statistical data that applies to a community [1]. In [8] a personalized environmental impact approach is described, allowing the tracking of human actions and their impact towards urban problem exposure and contribution. In [9] continuous physical activity data is captured and related to personal health goals in the form of user feedback [1]. These applications have been proven to be effective in impacting the way health is assessed, helping people improve behavioral patterns [1].

This paper reviewed several techniques relevant for the prediction of epidemiology on the scope of large scale sensing devices and exploiting social networks information. The review pinpointed several current issues that must be addressed by the forthcoming systems. Although the work here presented was targeted on human epidemiology, such systems might as well bring very useful insight on computer virus propagation on computer networks.
Acknowledgements

Authors supported by Harvard-Portugal Collaborative Research Award HMSP-CT/SAU-ICT/0064/2009: Improving perinatal decision-making: development of complexity-based dynamical measures and novel acquisition systems.

References

[1] Lane, N.D., Miluzzo, E., Lu, H., Peebles, D., Choudhury, T., Campbell, A.T., College, D.: Adhoc And Sensor Networks: A Survey of Mobile Phone Sensing. IEEE Communications Magazine (2010) 140–150

[2] Campbell, A., Eisenman, S., Lane, N., Miluzzo, E., Peterson, R., Lu, H., Zheng, X., Musolesi, M., Fodor, K., Ahn, G.: The rise of people-centric sensing. Internet Computing, IEEE 12 (2008) 12–21

[3] Zhang, D., Guo, B., Li, B., Yu, Z.: Extracting social and community intelligence from digital footprints: an emerging research area. Ubiquitous Intelligence and Computing (2010) 4–18

[4] Miluzzo, E., Lane, N., Fodor, K., Peterson, R., Lu, H., Musolesi, M., Eisenman, S., Zheng, X., Campbell, A.: Sensing meets mobile social networks: the design, implementation and evaluation of the cenceme application. In: Proceedings of the 6th ACM conference on Embedded network sensor systems, ACM (2008) 337–350

[5] Abdelzaher, T., Anokya, W., Bodha, P., Burke, J., Estrin, D., Guibas, L., Kansal, A., Madden, S., Reich, J.: Mobiscopes for human spaces. Pervasive Computing, IEEE 6 (2007) 20–29

[6] Kephart, J.O.: How topology affects population dynamics. C Langton ed. Artificial Life III Studies in the Sciences of Complexity (1994) 447–463

[7] Abbas, K., Mikler, A., Ramezani, A., Menezes, S.: Computational epidemiology: Bayesian disease surveillance. In: Proc. of the International Conference on Bioinformatics and its Applications, FL, USA, (2004) 1–12

[8] Barros, L., Bassanezi, R., Oliveira, R., Leite, M.: A disease evolution model with uncertain parameters. In: IFSIA World Congress and 20th NAFIPS International Conference, 2001. Joint 9th. Volume 3., IEEE (2001) 1626–1630

[9] Bisset, K., Chen, J., Feng, X., Kumar, V., Marathe, M.: EpiFast: a fast algorithm for large scale realistic epidemic simulations on distributed memory systems. In: Proceedings of the 23rd international conference on Supercomputing, ACM (2009) 430–439

[10] Kansal, A., Goraczko, M.: Building a sensor network of mobile phones. on Information processing in sensor (2007) 547–548

[11] Barrett, C.L., Bisset, K.R., Eubank, S.G., Feng, X., Marathe, M.V.: EpiSimdemics: An efficient algorithm for simulating the spread of infectious disease over large realistic social networks. In: International Conference for High Performance Computing, Networking, Storage and Analysis. (2008) 1–12

[12] Lu, H., Fan, W., Lane, N., Choudhury, T., Campbell, A.: SoundSense: scalable sound sensing for people-centric applications on mobile phones. In: Proceedings of 7th international conference on Mobile systems, applications, and services, ACM (2009) 165–178

[13] Kapadia, A., Triandopoulos, N., Cornelius, C.: AnonySense: Opportunistic and privacy-preserving context collection. Pervasive (2008) 1–18

[14] Lopes, L., Zamite, J., Tavares, B., Couto, F., Silva, F., Silva, M.: Automated social network epidemic data collector. In: INForum informatics symposium. Lisbon. (2009)

[15] Gorder, P.F.: Computational Epidemiology. Computing in Science & Engineering 12 (2010) 4–6

[16] Bisset, K.R., Eubank, S.G., Feng, X., Marathe, M.V.: EpiSimdemics: An efficient algorithm for simulating the spread of infectious disease over large realistic social networks. In: International Conference for High Performance Computing, Networking, Storage and Analysis. (2008) 1–12

[17] Kansal, A., Goraczko, M.: Building a sensor network of mobile phones. on Information processing in sensor (2007) 547–548

[18] Barrett, C.L., Bisset, K.R., Eubank, S.G., Feng, X., Marathe, M.V.: EpiSimdemics: An efficient algorithm for simulating the spread of infectious disease over large realistic social networks. In: International Conference for High Performance Computing, Networking, Storage and Analysis. (2008) 1–12

[19] Lopes, L., Zamite, J., Tavares, B., Couto, F., Silva, F., Silva, M.: Automated social network epidemic data collector. In: INForum informatics symposium. Lisbon. (2009)

[20] Gorder, P.F.: Computational Epidemiology. Computing in Science & Engineering 12 (2010) 4–6

[21] Boor, S.V., Muehlen, M., Rebelo, A.: Gripenet: an internet-based system to monitor influenza-like illness uniformly across Europe. (2007) 1–14

[22] Brownstein, J., Freifeld, C., Others: HealthMap: the development of automated real-time internet surveillance for epidemic intelligence. Euro Surveill 12 (2007) E071129

[23] Barrett, C.L., Bisset, K.R., Eubank, S.G., Feng, X., Marathe, M.V.: EpiSimdemics: An efficient algorithm for simulating the spread of infectious disease over large realistic social networks. In: International Conference for High Performance Computing, Networking, Storage and Analysis. (2008) 1–12

[24] Barrett, C.L., Bisset, K.R., Eubank, S.G., Feng, X., Marathe, M.V.: EpiSimdemics: An efficient algorithm for simulating the spread of infectious disease over large realistic social networks. In: International Conference for High Performance Computing, Networking, Storage and Analysis. (2008) 1–12

[25] Bisset, K., Chen, J., Feng, X., Kumar, V., Marathe, M.: EpiFast: a fast algorithm for large scale realistic epidemic simulations on distributed memory systems. In: Proceedings of the 23rd international conference on Supercomputing, ACM (2009) 430–439

[26] Eubank, S.: Scalable, efficient epidemic simulation. In ACM symposium on Applied computing (2002) 139–145