Computer Communications 59 (2015) 1–11

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Large scale model for information dissemination with device to device communication using call details records

Rachit Agarwal [a], Vincent Gauthier [a,*], Monique Becker [a], Thouraya Toukabrigunes [a,b], Hossam Afifi [a]

[a] *Lab. CNRS SAMOVAR UMR 5157, Telecom SudParis/Institut Mines-Telecom, 9, Rue Charles Fourier, 91000 Evry, France*
[b] *Orange Lab, 38 Rue du Général Leclerc, 92130 Issy les Moulineaux, France*

ARTICLE INFO

Article history:
Received 10 July 2014
Received in revised form 7 November 2014
Accepted 24 December 2014
Available online 13 January 2015

Keywords:
Device to device communication
Human mobility
Information dissemination
Opportunistic networks

ABSTRACT

In a network of devices in close proximity such as *Device to Device* (*D2D*) communication network, we study the dissemination of public safety information at country scale. In order to provide a realistic model for the information dissemination, we extract a spatial distribution of the population of Ivory Coast from census data and determine migration pattern from the Call Detail Records (*CDR*) obtained during the *Data for Development* (*D4D*) challenge. We later apply epidemic model towards the information dissemination process based on the spatial properties of the user mobility extracted from the provided *CDR*. We then propose enhancements by adding latent states to the epidemic model in order to model more realistic user dynamics. Finally, we study dynamics of the evolution of the information spreading through the population.

1. Introduction

In communication networks, devices in each other's physical proximity can communicate. Information dissemination in such proximity based networks have been the focus of a lot of studies. Through this paper, we study the dissemination of public safety information at country scale. For the type of communication network, as a case study, we choose *Device to Device* (*D2D*) communication paradigms. This choice relates to the fact that the market of context-aware applications and location-based services has grown tremendously and operators have started to consider the deployment of *D2D* communications as an underlay to the cellular networks. *D2D* communication is defined as a short range communication between devices in physical proximity without any involvement of the network infrastructure. *D2D* has many advantages like, autonomous communication, improved performance and spectrum reuse, low energy consumption and reduced load on the infrastructure. Moreover, other benefits of *D2D* communication include direct communication between devices even when the traditional infrastructure is down, better connectivity in poor connectivity regions and increased average rate of successful message delivery. These above mentioned benefits of *D2D* motivate us to choose *D2D* communication paradigms as a Use Case. Nevertheless, our Use Case can also be applied to mobiles applications such as `Firechat` [1] that leverages direct connection between devices in close proximity to temporally connect users.

Moreover, in opportunistic networks, lot of studies have focus on the store-carry-and-forward paradigm and on the dynamics of the message delivery. The first attempt to model the process was represented via epidemic forwarding. Variations of the process includes the integration of *Social Network Analysis* measures to study the dissemination process [2–5]. Further, modeling network performance through Markov Chains in opportunistic networks has also been the focus of several studies [6–9]. However, none of the methods are scalable to the country scale since they rely on computing individual transmission probability. Dissemination methods based on mean field approach, however, are more scalable [10,11] but rely upon homogeneous contact dynamics that is far from being valid. Thus, recent studies [12–15] on dissemination process are based on heterogeneous flow within the *metapopulation*. This enabling us to study and create a heterogeneous population model. These studies offer new perspectives on such models and enable us to study more easily country wide dissemination process.

Since, our interest lies in studying the dissemination process on a countrywide scale, we consider the population of the device to be spatially structured into subpopulation of well mixed individuals where the diffusion of safety information takes place. We use the

---

* Corresponding author.
E-mail addresses: rachit.agarwal@telecom-sudparis.eu (R. Agarwal), vincent.gauthier@telecom-sudparis.eu (V. Gauthier), monique.becker@telecom-sudparis.eu (M. Becker), toukabrigunes@orange.com (T. Toukabrigunes), hossam.afifi@telecom-sudparis.eu (H. Afifi).

http://dx.doi.org/10.1016/j.comcom.2014.12.010
0140-3664/

official boundaries (termed as *subprefecture*) in *Ivory Coast* to define the subpopulation structure of our *metapopulation* model. In order to know the flow of people between the subprefectures, we use the Call Detail Records (*CDR*) provided by the Orange Labs during the Data for Development (*D4D*) Challenge [16]. For future reference, we refer the provided *CDR* as the *D4D* dataset. The analysis of the movement pattern of the people in Ivory Coast reveals the flow of people between subprefectures. *D4D* dataset enable us to monitor trajectories of people at unprecedented scale and derive key spatio-temporal patterns of the movement. Thus, the mobility is modeled as a network of interactions between each subpopulation (population within a subprefecture) where the connections correspond to the flow of people among them. Metapopulation[17,18] model has been studied towards diffusion of information combined with human mobility where mobility pattern show strong heterogeneity [19,20,12]. Moreover, the outcome of the diffusion of the public safety information is regulated by the coupling of the mobility process and the diffusion process in each subpopulation. Thus, it is of utmost importance that the mobility model correctly describes realistic pattern of mobility.

However, a simple memoryless mobility model is far from being able to model recurring patterns such as commuting behavior or traveling behavior that are often observed in mobility data [19,21]. To include this effect we use a non-markovian diffusive process to cope with the recurring behavior [13,14]. In a realistic human population, a user is associated with a home location. Every time a user moves away from the home location, there is a probability to return to the home location from the new location. This return probability is also derived from the *D4D* dataset. More details on how the return probability is computed is presented in the Section 3.1.

The *D4D* dataset includes the information about the user mobility and the association of a user to a community, termed as subprefecture. Further, in order to study the model with realistic population structure, we use the census data provided in [22] to populate each subprefecture (cf. Section 3.1).

Thus, in this paper, we use insights from the dissemination in the metapopulation model and the *D4D* dataset to formulate our analytical model. Further, our model is divided into various stages. First, construct synthetic yet realistic population of the devices using census data [22] for the region of Ivory Coast. Then extract mobility information from the *D4D* dataset. Then find mobility steady state population in each community. Finally, apply Susceptible(S)-Infected(I)-Recovered(R) type epidemic modeling (cf. Section 4.2) to model the information dissemination process at the countrywide scale.

In order to reduce the complexity of the dissemination process at country scale, we use the reactive–diffusive equation with homogeneous assumption to describe the subpopulation at the mechanistic level. We use a variation of the *SIR* epidemic model where Susceptible (S), Infected (I) and Recovered (R) states correspond to the state of a device ‘not having the message’, ‘having the message and transmitting it’ and ‘having received the message but stop transmitting it due to message timeout (message time to live, *TTL*)’, respectively. A comprehensive survey about the epidemic model is provided in [23,24]. Recently, the authors of [25] showed that not only the community structure affects the dissemination process but the density of the communities also play an important role in the dissemination process. In metapopulation model it is often assumed that the population within a single subpopulation is well mixed. This is of course not generally true specially in a large subpopulation. To acknowledge this in a large subpopulation where some devices may never be in contact at a certain time, we include Latent states (E) in our model. Device in latent state account for devices that belong to a given subpopulation but are not participating in the diffusion process. To distinguish between devices that are Latent but Susceptible, Latent but Infected and Latent but Recovered, we subdivide E into three states: E_S, E_I and E_R. In our model only active devices in S, I and R states participate in the dissemination process. Thus, the novelty of the paper lies in the introduction of latent states and in the way we compute the number of interactions per device.

The remainder of this paper is organized as follows. In Section 2 we first provide an overview of the *D2D* scenario. In Section 3 we describe how the dataset provided by the *D4D* organizers is used to get useful information. A detailed description of the model is then provided in Section 4 which is followed by the results obtained in Section 5. The paper finally concludes with Section 6.

2. Loosely controlled *D2D* support for information dissemination

We assume the following *D2D* scenario: Using a Public Warning System (PWS), an *eNB* (evolved Node B, a radio interface in the Long Term Evolution (LTE) network) has to broadcast a national emergency warning which should reach the maximum number of people in the country. However, in such situations, the operator’s network is massively overloaded by a huge traffic of users trying to use their mobile devices at the same time. Also, the warning message may not reach a part of the population due to network coverage issues in rural areas. Using *D2D* communications, devices in the physical proximity can communicate directly with each other and exchange messages. This mechanism allows not only an extension to the broadcasting area but also enhances the traffic on the network infrastructure and avoids network saturation and waste of resources. A *D2D* communication has two phases: the Neighbor Discovery phase and the Communication phase for the data exchange. Both phases can be based on either a direct approach or a network-assisted approach [26–30].

In the direct *D2D* approach, devices discover other devices in their surrounding by exchanging presence beacons. Both the phases, the discovery and the communication phase, are done in an adhoc-like way without any assistance or control of the network infrastructure. The communicating devices then form a self-organized and a self-configurable network. This approach is flexible and highly scalable as it can adapt to an increasing number of connected devices and active *D2D* links. This allows offloading the traffic of the core operator network to local *D2D* communication. Besides, it is also a solution to avoid communication disruption when some *eNBs* are down in a disaster situation like Earthquake or when an *eNB* fails. The users could still use their devices and be connected to the operator network using direct communication.

In the centralized *D2D* approach, the operator network may fully control the discovery and communication phases (fully controlled *D2D*) or assist the devices during the whole *D2D* process (loosely controlled *D2D*) by enabling authentication and QoS mechanisms. Nevertheless, this approach is less scalable than the direct approach as it consists of performance and load balancing issues at the Radio Access Network (RAN) when dealing with a huge number of *D2D* connections. As depicted in Fig. 1 we shows that some *eNBs* can allow the direct communications between devices that are in physical proximity while some do not.

In addition, a device can move from one *eNB* to another. Moreover, the device can also move within an associated *eNB*. In this paper, we assume that all *eNBs* allow direct communication between devices that are in physical proximity and that a public warning message is generated in some area associated to an *eNB*. In the Section 4, we are interested in investigating how fast dissemination of the warning message can be achieved through a large population in a *D2D* environment where the devices are also mobile and links are intermittent. By fast dissemination we mean