A Zero-Dimensional Mobility Model for Opportunistic Networking

Ljubica Pajevic and Gunnar Karlsson
Laboratory for Communication Networks and Linnaeus Center ACCESS
KTH, Electrical Engineering
Stockholm, Sweden
{ljubica, gk}@kth.se

***Abstract*—The demands for ubiquitous connectivity and high data rates have led to a proliferation of mobile devices such as smartphones, media tablets and netbooks. These devices, equipped with powerful communication and storage resources, are carried by humans during their daily activities and, thus, possess a high potential for opportunistic networking. In this networking mode, encounters between users define the "topology" of the network and can be exploited for content sharing. A large portion of user encounters occur in densely populated public urban areas such as city squares, subway stations, shopping malls, etc. In this paper, we propose an analytic model to study content dissemination characteristics inside smaller areas where the mobility of nodes does not affect their connectivity. Subsequently, we argue that this model is suitable for modeling larger areas as interconnections of these basic blocks.**

***Keywords*-Opportunistic networks, content distribution, queueing analysis, mobility modeling**

I. Introduction

The growing demand for all-in-one converged mobile devices and digital content has led to wide adoption of handheld devices such as smartphones and media tablets. These devices are equipped with powerful computing and communication resources, often providing support for various wireless technologies, including Bluetooth, Wi-Fi and 3G. On one hand, this brings the convenience of pervasive connectivity, but on the other hand it invites more user activity, leading to high data traffic. Mobile data traffic represents already an increasingly large fraction of Internet traffic (with the current share of 25%), predominantly as a result of traffic migration from fixed to mobile networks [1]. Opportunistic networking is a communication paradigm based on proximity of mobile nodes and their capability to store and forward contents. The mobile devices (nodes) carried by users define the "topology" of the network. Whenever two nodes get into transmission range of each other, they are able to establish a wireless connection and exchange contents. The advantage of this type of communication is that it does not require support of infrastructure since the links between nodes are formed ad hoc and use license-free spectrum. Various factors such as the nodes' movements, interference and usage of power-save mode may cause disruption of links. Hence, the existence of end-to-end paths between arbitrary nodes in the network cannot be assumed. Opportunistic networks are suitable for content-provider or user-generated data such as multimedia files, documents, or any other type of data that can tolerate modest delays and losses.

Estimates are that smartphones and portable devices will be the primary drivers of overall data traffic growth in the near future. Unlimited data plans for mobile users lead to extremely high mobile data consumption that is raising the risk of network congestion. Thus, when deploying new technologies (such as LTE) in order to meet increasing traffic demands, mobile operators should also consider offloading mobile data by using opportunistic contacts [2].

The content dissemination scheme addressed herein is based on a publish/subscribe model. Data is organized into information feeds to which users subscribe. Subscribers of the same feed exchange data opportunistically in a peer-to-peer manner.

The importance of mobility modeling stems from the fact that wireless system performance can be heavily influenced by user mobility. However, large traces of *realistic* movement patterns are difficult to obtain. Thus, performance is usually evaluated by using analytic and synthetic models. In this paper we propose an analytic queueing model to study how mobility and system parameters affect the performance. We focus on specific urban environments where we could assume zero mobility (e.g shop, bus station, inside buses and trains, stoplight at pedestrian crossing), which can be represented as a single "mobility block". Then, we propose the internal structure of larger urban areas such as public squares, shopping malls, subway stations and commercial buildings that can be approximated by networks of basic zero-dimensional blocks. Subsequently, vast areas can be modelled as interconnection of different blocks.

The rest of this paper is organized as follows. In the next section we further motivate our work and contrast it with related work. Section III contains the detailed description of our analytic model. In section IV we explore the model by means of simulation and in section V we propose an idea how this model can be exploited for more complex modeling. We include analysis of the real-world scenario dataset and comparison to our model in section VI. Finally, we conclude our work in section VII and give an outline of the future work on this topic.

978-1-4577-0351-5/11/$26.00 ©2011 IEEE

II. Related Work

The mobile peer-to-peer opportunistic system we consider in this paper is described in [3]. Our work addresses two topics: content dissemination and mobility modeling and we compare related work with respect to these.

There has been a significant research in delay-tolerant networking. Most of the studies address different routing schemes in closed systems and confined two-dimensional areas. In [4], authors have developed analytic model for pedestrian movement on a one-dimensional, street-like segment. They assume open system and one-hop forwarding scheme to study content distribution. Another study that addresses data dissemination in peer-to-peer mode is presented in [5]. Although the main asset of this study is that it carefully structures contact patterns for a large population of users over large time scale, social interactions among users in this case have not been considered.

Mobility has been frequently studied by simulations and using synthetic mobility models such as random-waypoint [6] or random walk [7], [8]. Synthetic models attempt to represent mobility behaviour of mobile nodes without use of realistic traces. Although easier to generate and use in simulations, these models often lack realism.

Realistic mobility modeling is essential for evaluating and improving mobile system performance. This requires direct capture of real-life human traces and contact opportunities. CRAWDAD [9] is a public archive of traces collected from mobile users in a specific contexts, such as a conference [10] or from a campus [11]. However, contact traces obtained in these cases have short duration, a small population of users, specific environments and low time-granularity. SLAW mobility model [12], represents a statistical model that simulates human mobility patterns, mimicking the way people move over the course of a day, a month or longer. This model was developed based on previously reported statistical features of human mobility and with respect to the GPS traces collected from various user groups at different outdoor sites.

III. Description of Zero-Dimensional Model

In this section we propose an analytic queueing model for content distribution. We denote the model as zero-dimensional, where "zero-dimension" should indicate that nodes in the system are, regardless of their mobility patterns, always connected; that is, internal mobility of the system does not affect connectivity. The setup is as follows. We consider an open system consisting of mobile nodes equipped with short range radios (e.g. WiFi or Bluetooth) and ample data storage. The mobile nodes arrive into the area according to some arrival process, reside inside for a particular amount of time and eventually leave (depart the system). We assume the nodes have the same transmission range Δ. The size of the area is sufficiently small in comparison to Δ; whenever two nodes collocate inside, they will be in communication range of each other and able to exchange data.

Some of the nodes entering the area carry the published data and transfer it to subscribers when contact is established. We are interested in analyzing the achievable content spreading. For sake of simplicity, we assume that all the nodes are subscribed to a single channel and interested in obtaining the same data element. Hence, the content spreading scheme is epidemic. Further, the size of the element is small in comparison to the data link capacity and the transfer time is hence negligible.

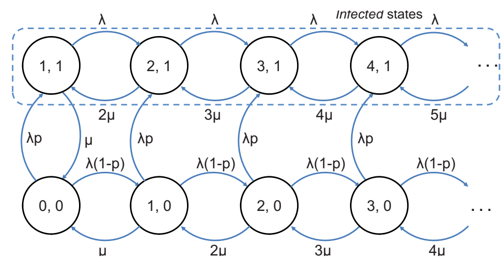


Figure 1. Markov chain for the zero-mobility model.

Denote by λ the arrival rate of the nodes and by $p \in [0, 1]$ the proportion of nodes that carry the content (in further text, we will also use the term *injection probability* for p). Assuming that arrivals occur according to a Poisson process and the node sojourn time (t_s) is exponentially distributed with mean value $\bar{t_s} = 1/\mu$, we can model the system with the Markov chain depicted in Fig. 1.

The states of the chain are $S_{i,j}$ where i, $i \geq 0$ denotes the number of nodes in the area and j, $j \in \{0, 1\}$ the presence of the contents inside: in states $S_{i,1}$ for $i \geq 1$ all the nodes in the area have obtained the contents in infinitesimally short time upon arrival. In state $S_{0,0}$ the system is empty. When the system is in the lower branch of the chain, i.e. in any of the states $S_{i,0}$, it transits to state $S_{i+1,1}$ with the rate λp, with the arrival of a node that carries the contents. Otherwise, the system transits to state $S_{i+1,0}$ with the rate $\lambda(1-p)$ or to state $S_{i-1,0}$ when a departure occurs. The system transits from state $S_{i,1}$ to state $S_{i+1,1}$ with the rate λ or to state $S_{i-1,1}$ with the rate $i\mu$. An arriving node obtains the contents if it finds the system in any of the infected states. Clearly, once the contents enter the system, it will reside inside as long as the system remains non-empty. The probability of the system being in state $S_{0,0}$ and the state equations are given by:

$$\pi_{0,0} = e^{-\frac{\lambda}{\mu}} \tag{1}$$

$$\begin{aligned} \pi_{k,0}\,(\lambda + k\mu) =& \pi_{k-1,0}\,\lambda(1-p) \\ &+ \pi_{k+1,0}\,(k+1)\mu, \;\; k \geq 1 \end{aligned} \tag{2}$$

$$\pi_{1,1}\,(\lambda + \mu) = \pi_{0,0}\,\lambda p + \pi_{2,1}\,2\mu \tag{3a}$$

$$\begin{aligned} \pi_{k,1}\,(\lambda + k\mu) =& \pi_{k-1,1}\,\lambda + \pi_{k-1,0}\,\lambda p \\ &+ \pi_{k+1,1}\,(k+1)\mu, \quad k \geq 2 \end{aligned} \tag{3b}$$

Despite the simple structure, the state probabilities can alas not be easily expressed in closed form. We will therefore study the system numerically.

IV. Exploring the model by simulations

In order to get an insight how the proposed content spreading scheme works, as a first approximation we presume the above scenario. We will then relax the assumptions on arrival process and sojourn time distribution by means of simulation.

The performance of the scheme is evaluated by using two basic metrics which we denote as:

– *Content distribution*: probability that an arbitrary node possesses the contents upon its departure from the system. This metric gives the proportion of nodes that carry the contents when they leave the system—i.e. the sum of nodes that brought the contents and nodes that obtained the contents from the others.

– *Contact duration*: the time that two nodes are connected to each other. Assuming that the transmission range covers the entire simulation area, so that all nodes residing in the area are connected.

A. Content Distribution

Fig. 2(a) shows the content distribution as a function of the arrival rate λ. In the simulations, the mean value of the sojourn time was set to 50 s and the percentage of nodes carrying the contents took values 1%, 0.5% and 0.1%. We also show the results from a simulation run where the mean sojourn time $\bar{t_s}$ was set to 30 s. In all the simulations, the estimated values are averaged over 20 trials and the number of nodes in each trial was 10^5. It can be concluded that the content distribution is very sensitive to the distribution of sojourn time and, consequently, the average number of nodes in the system. For example, when $\lambda = 0.17$ s^{-1}, there are 5 and 8.5 nodes in the system on average for $\bar{t_s} = 30$ s and $\bar{t_s} = 50$ s, respectively. In the first case the content distribution yields only 18% while in the latter it reaches almost 84%. As the sojourn time increases, the scheme becomes more efficient even when the proportion of nodes bringing the contents is small. Table I shows the arrival rate, average number of nodes ($\bar{N}$), the content distribution probability (P_{cd}) and the average time between two arrivals of the contents (T_c) in a scenario when $p = 0.1\%$ and $\bar{t_s} = 50$ s.

Table I

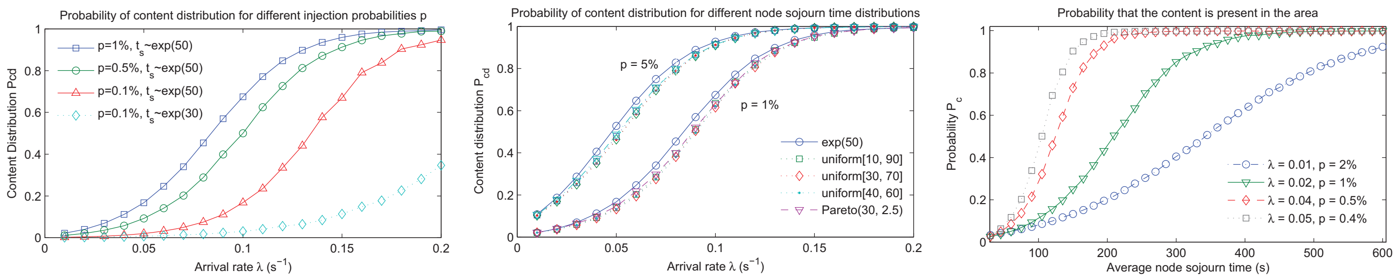
$\lambda[s^{-1}]$	0.14	0.15	0.16
$\bar{N}$	7	7.5	8
P_{cd} [%]	58	67	79
T_c [minutes]	119	111	104

Notice that for the average inter-arrival time of the contents $T_c = 104$ minutes and 8 nodes in the system, the probability of obtaining contents is 79%, which is a notable result bearing in mind that the average sojourn time is less than one minute.

Since we observe that content distribution works even for very low injection probabilities, we further investigate the impact of node sojourn time on this metric. Fig. 2(c) depicts the probability (P_c) that the system is *infected*, that is the probability the contents reside in the system as a function of the average sojourn time. The distribution of sojourn time is exponential and the mean value has support on interval [30, 600] s. The nodes arrive according to a Poisson process and the arrival of contents is a thinned Poisson process with probability p. Parameters λ and p in four simulation setups are tuned so that the contents arrival rate is the same in all cases and equals $\lambda p = 0.0002$ (the contents arrive every 83 minutes on average). As this measure indicates the proportion of time the contents reside in the area, we infer that the model encompasses the *virtual storage* effect– the contents remain despite absence of a storage node, e.g. an access point [13]. Having low inflow of nodes ($\lambda = 0.01$ s^{-1}), the probability P_c slowly increases with the increase of the sojourn time. The effect becomes more evident at higher node arrival rates. Intuitively, this is an obvious consequence of nodes queuing and the *all-peers-in-range* connectivity in the system.

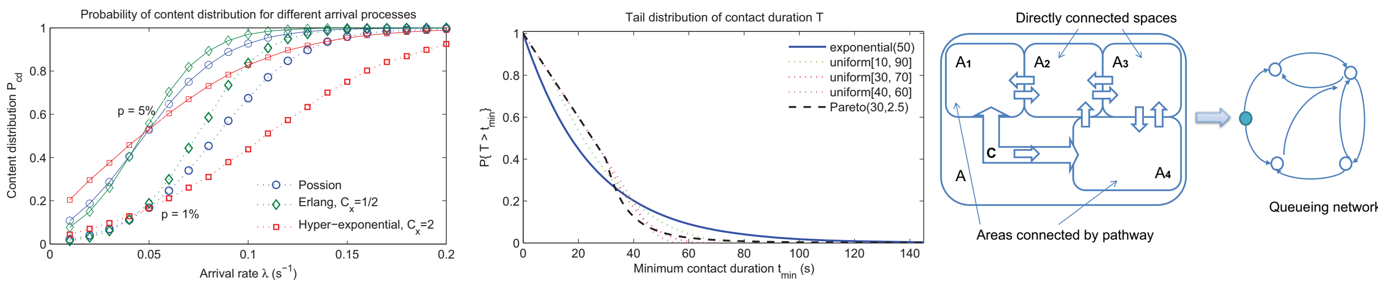
Then, we evaluate the impact of the node sojourn time distribution on content distribution. In the next set of simulations, nodes arrive according to a Poisson process with the rate $\lambda = 0.1$ s^{-1}. Plots in Fig. 2(b) show the content distribution for the following sojourn time distributions: exponential, three uniform distributions and Pareto distribution. The mean value for all distributions is 50 s. Uniform distributions have support on intervals: [10, 90] s, [30, 70] s and [40, 60] s and Pareto distribution has the minimum value 30 s and shape parameter $\alpha = 2.5$. The injection probabilities p for two sets of curves are 1% and 5%. The difference in performance is not significant and note that the exponential distribution achieves the highest performance, followed by Pareto and uniform distribution. This inference can be used to simplify the modeling of departure process and assume the inter-departure time is exponentially distributed.

In contrary to node sojourn time distribution, the arrival process affects content distribution, as the curves in Fig.3(a) show. The mean arrival rate λ is the same for all the processes: a Poisson process, an Erlang arrival process consisting of 4 stages, each with the rate $\lambda/4$, and a two-stage hyper-exponential process with parameters 0.35λ and 5.7λ for the rates, and 0.31 and 0.69 selection probabilities for the first and the second phase, respectively. The coefficient of variation for the Erlang distribution is 1/2 and it is 2 for the hyper-exponential distribution. We look into two series of curves for P_{cd}: for the lower series, probability $p = 1\%$ and for the higher series $p = 5\%$. We can see that when $p = 1\%$, the burstiness of the hyper-exponential arrival process has negative effect on the distribution since many arrivals occur in a short time, but the proportion of content-carriers is low.



(a) (b) (c)

Figure 2. Content distribution as a function of arrival rate and (a) different injection probabilities p. The node sojourn time is exponentially distributed. (b) node sojourn time distributions. (c) Probability the content is in the area. The content arrival rate is $\lambda p = 0.0002$. The arrival process is Poisson for (a)-(c).



(a) (b) (c)

Figure 3. (a) Content distribution for various arrival processes. The node sojourn time is exponentially distributed with mean value 50 s. (b) Tail distribution function of contact duration for different sojourn time distributions. (c) Two-dimensional area consisting of zero-dimensional building blocks.

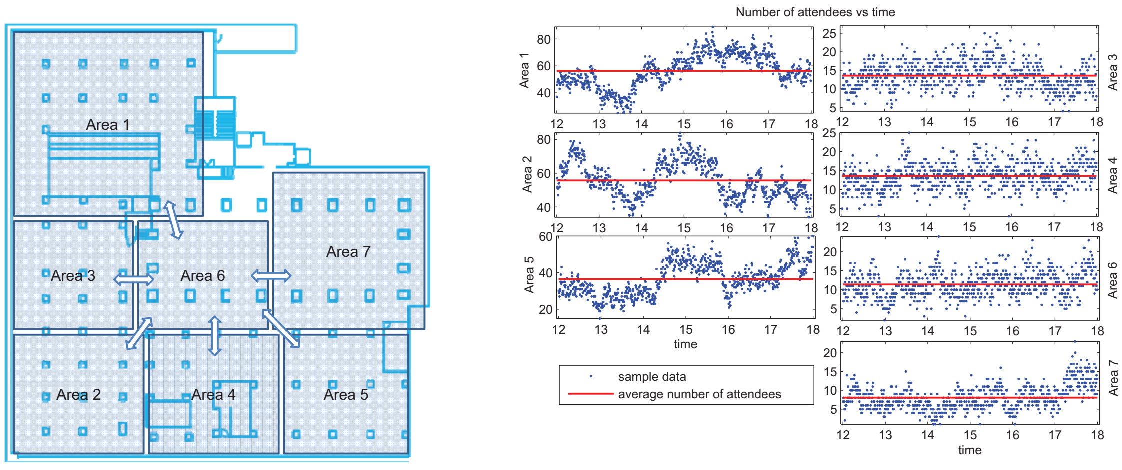
B. Contact Duration

An opportunistic content distribution system depends on the characteristics of human mobility and behaviour that can be studied, but not tuned for optimizing the performance. Our epidemic scheme presumes data transfer in zero-time. In reality, a transfer consists of the time it takes the nodes to discover the neighbour, to establish a connection and the actual time needed to transfer the data. Therefore, one of the essential issues in system design is to minimize the connection setup time to make even contacts with shorter duration useful. In scenarios where many of shorter contacts occur, this optimization can boost the dissemination significantly. Imposing the requirement for the minimum time t_{min} for connection establishment and data transfer, we consider only contacts that last longer than t_{min} to be useful. As an illustration, suppose that the data of interest is a media file, e.g. an *mp3* song (typically of size around 5 MB) which can be sent over 802.11 in a few seconds. Fig. 3(b) illustrates the tail distribution of a random variable T, given by $P\{T > t_{min}\}$, for exponential (plotted as solid line), uniform (dotted line) and Pareto distribution (dashed line). The mean value for all distributions is 50 s. For a required $t_{min} = 10$ s, the lowest probability is associated with exponential distribution, 67%. While there are still many contacts longer than t_{min}, the assumption of zero-transfer time needs to be reconsidered, since we cannot account for all contacts as useful.

V. A ZERO-DIMENSIONAL MODEL AS A BUILDING BLOCK

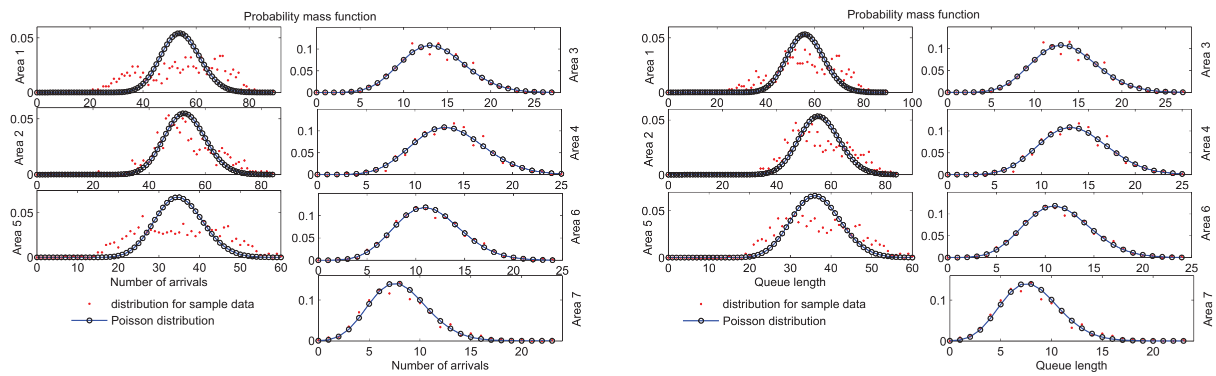
The idea of modeling larger areas from zero-dimensional blocks is illustrated in Fig. 3(c). Consider the area A to be, for example a floor in a building and assume areas A_1, A_2, A_3 and A_4 represent offices in that building. Since the offices are of relatively small dimensions and we assume free flow of pedestrians, we can model those areas as zero-dimensional blocks. C represents an area where few contacts occur or those contacts are not considered useful (e.g a corridor). The entire area can be modelled as a queueing network, where the white nodes in the network represent the offices and the dark one represents the corridor. Mobile users can be tracked on their path as they visit the network nodes and we evaluate content dissemination from their traces. The dark node does not have any impact on dissemination, but introduces waiting time, which should also be taken into account. The advantages with this model in comparison to other spatial models are:

– *Simplicity*—rather than observing the continuous space and capturing movement and connectivity characteristics of the nodes in the system, this model introduces discrete space analysis.
– *Computational tractability*—concatenating smaller areas, represented by zero-dimensional model, larger spaces can be built up from the basic building blocks. The entire model can be represented as a queuing network.



(a) (b)

Figure 4. (a) Floor plan partitioned into 7 areas. (b) Number of attendees in an area vs time (scatter plot) and average number of attendees (solid line).



(a) (b)

Figure 5. Probability mass function per time slot (30 s) for: (a) number of arrivals, (b) queue length.

VI. Mapping the Model to Real-world Scenarios

Following the idea in the previous section, we study a specific mobility scenario, showing that it can be mapped to a queueing network model.

The *Attendee Meta-Data* set [14] contains RFID tracking data, collected from the HOPE (Hackers On Planet Earth) conference held in July 2008 in New York. Conference attendees received RFID badges that could identify and track them uniquely across the conference space. Badges would send out the beacons to RFID readers roughly every 30 s, locating the attendees in one of 21 zones on two floors of the conference hotel. The traces were collected from 1280 participants during 3 conference days; we choose the time window from 12 p.m. to 18 p.m. during the second day of the conference, when the highest number of the participants were active and traceable (1113 attendees). We focus our attention on one of the floors which is characterized by higher mobility. Zones located on this floor were, for instance: demonstration area, information desk, exhibition area, vendor and network operators stands; the other floor zones comprised mostly lecture rooms.

Fig. 4(a) depicts the floor-plan consisting of seven different areas. Initially, this floor was partitioned into 15 zones; we introduce a simplification of the model by merging adjacent zones and forming larger areas. Area 1 is different from the others, since this is where the escalators are located and all attendees that move from one floor to the other will visit it. The number of nodes in an area versus time is plotted in Fig. 4(b). Fluctuations during 6 hour period are obvious for areas 1, 2 and 5, while stationarity might be assumed for areas 3, 4, 6 and 7. The distributions of sojourn times show strong power-law decay with the mean values 62.0 s, 50.8 s, 53.6 s and 61.2 s for areas 3, 4, 6 and 7, respectively, and 233 s, 222 s and 147 s for the other three areas. We find that the arrival rates to areas range from 0.13 s^{-1} (area 7) to 0.27 s^{-1} (area 4) and $\lambda > 0.2$ s^{-1} for the rest. With regards to these results, we refer to Fig. 2(a) where we showed that the probability of content distribution is almost equal to 1 when $\bar{t_s} > 50$ s, $\lambda > 0.2$ s^{-1}, which is the case for areas 3, 4, 6 and 7. Further, as longer average sojourn times were estimated for the corner areas, referring to Fig. 2(c), we can also assume high dissemination in these areas.

We estimated the probability mass function for the number of arrivals and the queue lengths during a time slot of 30 s (the time between two consecutive timestamps). Fig. 5 shows the corresponding PDFs for all of the areas, where the samples from the traces are represented with dots. The solid lines are Poisson distributions with mean values equal to those estimated from the traces. Hence, we observe that the PDFs corresponding to areas 3, 4, 6 and 7 show a good match with a Poisson distribution. Our future work will address the question whether the discrepancies between the model and real traces (found in areas 1, 2 and 5) affect the system performance.

VII. Conclusion

In this work we have addressed mobile peer-to-peer content distribution in an open system where pedestrians exchange contents over short range radios in an opportunistic manner. We proposed an analytic queueing model to study the performance of content dissemination scheme in smaller areas where high connectivity of nodes can be assumed. Our findings are the following:

– The epidemic dissemination is found to be feasible even in critical cases for the system, when the inflow of nodes carrying the contents is relatively low. We also observe that the system provides *virtual storage* for the contents, i.e. contents resides in the system over long proportion of time, without support of a stationary node that could assist the distribution.
– On one hand, the scheme is not sensitive to node sojourn time distribution, which simplifies modeling of the system, since we can presume exponential distribution and exploit its memoryless property. On the other hand, the impact of nodes' arrival process is significant and becomes more pronounced for lower injection probabilities.
– We found that the assumptions of arrival process in our model show a reasonable approximation of those extracted from the real mobility patterns.

In addition, we propose a method for modeling and evaluating content distribution in larger spaces by building networks out of the basic zero-dimensional model.

Our future work will focus on studying characteristics of content distribution in such two-dimensional areas. We will compare these to two-dimensional models in which there is internal, non-negligible mobility which introduces contact opportunities, but also limits the connectivity of nodes inside. We will conduct simulations to compare the discrete space and the continuous space models and evaluate under what conditions the two-dimensional model may be approximated by networks of zero-dimensional models.

Further, we will investigate whether our model can capture other (and larger) datasets and different mobility scenarios. We will also study the idea of modeling vast spaces as queueing networks.

References

[1] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2010-2015. [online]. available: http://www.cisco.com [accessed Feb, 2011.]

[2] V. Vukadinović and G. Karlsson, "Spectral efficiency of mobility-assisted podcasting in cellular networks." *Proc. of the Second International Workshop on Mobile Opportunistic Networking* (MobiOpp '10). ACM, New York, NY, USA, 51-57.

[3] G. Karlsson, V. Lenders and M. May, "Delay-tolerant broadcasting." *IEEE Trans. Broadcasting* vol. 53, no.2, pp. 369-381. 2007.

[4] V. Vukadinović, Ó. R. Helgason and G. Karlsson, "A mobility model for pedestrian content distribution." *Proc. Simutools, SCENES Workshop*, Rome, Italy, March 2009.

[5] V. Srinivasan, M.Motani and W. T. Ooi, "Analysis and implications of student contact patterns derived from campus schedules" in *Proc. MobiCom '06*, Los Angeles, 2006.

[6] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks." in *Mobile Computing*. Kluwer Academic Publishers, pp. 153-181. 1996.

[7] E. Royer, P. Melliar-Smith and L. Moser, "An analysis of the optimum nodes density for ad hoc mobile networks." *Proc. IEEE ICC*, Helsinki, Finland. 2001.

[8] J.-Y. Le Boudec, M. Vojnovic, "The Random Trip Model: Stability, Stationary Regime, and Perfect Simulation." *IEEE/ACM Transactions on Networking*, vol.14, no.6, pp.1153-1166, Dec. 2006

[9] CRAWDAD: A community resource for achieving wireless data at Dartmouth. [online]. available: http://crawdad.cs.dartmouth.edu. [accessed Feb, 2011.]

[10] P. Hui, A. Chaintreau, J. Scott, R. Gass and J. Crowcroft, "Pocket switched networks and human mobility in conference environments." *WDTN '05: Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, Philadelphia, Pennsylvania, USA, 2005.

[11] M. Kim, D. Kotz and S. Kim, "Extracting a mobility model from real user traces" *Proc. IEEE Infocom*, 2006.

[12] K. Lee, S. Hong, S.J. Kim, I. Rhee and S.Chong, "SLAW: A mobility model for human walks." *Proc. Infocom*, Rio de Janeiro, Brazil, 2009.

[13] E. Hyytiä, J. Virtamo, P. Lassila, J. Kangasharju and Jörg Ott, "When Does Content Float? Characterizing Availability of Anchored Information in Opportunistic Content Sharing." *Proc. Infocom*, Shanghai, China, 2011.

[14] http://amd.hope.net [accessed April, 2011.]

[15] P.-U. Tournoux, J. Leguay, F. Benbadis, V. Conan, M.D.de Amorim, J. Whitbeck, "The accordion phenomenon: analysis, characterization, and impact on DTN routing." *Proc. Infocom*, Rio de Janeiro, Brazil, 2009.

[16] J. Lee, J.C. Hou, "Modeling steady-state and transient behaviors of user mobility: Formulation, analysis, and application." *Proceedings of ACM MobiHoc*, May 2006.

[17] J. Yoon, B.D. Noble, M. Liu, M. Kim, "Building realistic mobility models from coarse-grained traces." *Proceedings of ACM MobiSys*, June 2006.

[18] P. Scully and R. Skehill, "Demonstration of aggregate mobility models in MATLAB." *Irish Signals and Systems Conference*. pp.271 - 276, Dublin. 2006.