On the Performance of Pedestrian Content Distribution

Vladimir Vukadinović, Ólafur Ragnar Helgason, and Gunnar Karlsson
ACCESS Linnaeus Center and School of Electrical Engineering
KTH, Royal Institute of Technology, Stockholm, Sweden
Email: {vvuk, olafurr, gk}@ee.kth.se

ABSTRACT—Mobile communication devices may be used for spreading multimedia data without support of an infrastructure. Such a scheme, where the data is carried by people walking around and relayed from device to device by means of short range radio, could potentially form a public content distribution system that spans vast urban areas. There are basically only three system parameters that can be determined in the design: the transmission range of the nodes, the setup time when nodes make a contact, and their storage capacity. The transport mechanism is the flow of people and it can be studied but not engineered. The question addressed in this paper is how well pedestrian content distribution may work. We answer this question by modeling the mobility of people moving around in a city, constrained by a given topology. The model is supplemented by simulation of similar or related scenarios for validation and extension. Our conclusion is that contents spread well with pedestrian speeds already at low arrival rates into a studied region. Our contributions are both the results on the feasibility of pedestrian content distribution and the queuing analytic model that captures the flow of people.

Keywords: delay-tolerant networks, disruptive networks, broadcast, multicast, content distribution, queuing analysis, mobility modeling.

1. INTRODUCTION

Ubiquitous wireless coverage has often been promoted for providing continuous connectivity in mobile communications. Such coverage is alas hard, or at least uneconomical, to provide in reality. The work

Dagstuhl Seminar Proceedings 09071
Delay and Disruption-Tolerant Networking (DTN) II
http://drops.dagstuhl.de/opus/volltexte/2009/2359

presented in this paper is based on the premise that continuous connectivity is not universally needed and that intermittent communication is useful for applications characterized by a low degree of interactivity, e.g., broadcasting, paging, messaging, and data collection. We explore the performance of a communication mode that relies on mobile nodes which communicate with one another when they are within radio range and which carry the data onwards through their own movements. Hence, the mobility patterns of nodes affect the speed, throughput and reliability of the data forwarding. The application area addressed herein is the distribution of multimedia contents and we are primarily concerned with pedestrian mobility. The contents may either be generated by the mobile nodes or they may be provided to the nodes through a sparse infrastructure (e.g. public WLANs).

The setup is as follows. The mobile nodes may communicate over short-range radio, such as Bluetooth or WiFi. We assume a simple model of the physical layer in which nodes connect if they are within a transmission range Δ of one another. Interference, fading, power control and other data-link functions are not considered. The mobile nodes—which could be devices such as mobile phones, media players and cameras—may cache contents both for their users and for other mobile nodes. When two or more nodes get within transmission range, they will connect according to the hand-shaking protocol of the data link and will then start to exchange data according to an application-level protocol. This protocol would be a simple query-response between nodes; the details of how it operates are not germane to this study. We assume that the contents are brought in by some of the mobile nodes, or that nodes receive the data when passing by an access point, depending on the case of study. Furthermore, the contents are provided in atomic units—mp3 files, still images, video clips, news items—that are meaningful to the application independently of one another.

We assume that the storage in the mobile nodes does not restrict the performance and we do not consider the issue of power efficiency of the system. Our concern is with the performance of the content distribution: the spread of the contents, as well as the rate and length distribution of contacts. We study in what manner the performance is affected by the mobility of the nodes and by the system parameters. The last study treats the case when contents are brought into an area by a fraction of the arriving nodes and how well it then spreads to other nodes. We use both analysis and simulation in the study. The queuing model shows explicitly how the mobility and system parameters affect the performance. Both the queuing model and the simulations capture topological restrictions of the mobility as well as the stochastic number of mobile nodes in a region of interest.

The contributions we report in this paper are the following. i) We have developed a detailed queuing-analytic model for people moving on a street and use it to study how well contents are distributed there. ii) The model for a single street may be used to build a network of streets to model larger areas. This is shown in principle and with a realistic example. iii) We report on the performance results for the system where the analytic results are compared to simulation results with good agreement. The simulation is also used to study the transient behavior of the system. We find that contents may become resident in an area under favorable conditions, even though the nodes sojourn there only temporarily.

The structure of the paper is as follows. We review related work in the next section and discuss similarities and differences of our contributions with those reported in the literature. The street model is introduced in Section 3 and the performance results for a single street are discussed in Section 4. The content distribution in a grid of streets is addressed in Section 5. We conclude our findings in Section 6.

2. RELATED WORK

We position our work with related works in two respects: the wireless content distribution, and mobility modeling.

2.1 Content distribution

There has been substantial work on peer-to-peer content distribution systems for the Internet. BitTorrent is a successful instance of such systems that post-facto has gained interest by the research community; see for instance [2]. It is based on a general family of gossip protocols [3]. Our system belongs to the general field of delay-tolerant networking [1]. The application of gossiping protocols to mobile communication has been proposed in, for instance, [4] – [6]. Multicast for delay-tolerant networks has been proposed in [8]. Our work assumes open user groups. The type of mobile content distribution system that we analyze is described in [9].

2.2 Mobility modeling

Mobility has most frequently been studied by simulation where a fixed number of nodes move on a convex area, often a square or a circle [10]. The random waypoint model is notable owing to its popularity; its stationary distribution of nodes is provided in closed form in [11]. The random-trip model is a general mobility model that allows perfect simulation and general topologies [12]. Mobility-assisted routing for mobility in two-dimensions has been studied in [5], [13]. The pocket switched experiment and analysis is primarily concerned with contact opportunities for use in opportunistic routing [19]. There is not any routing in the system we study and our measures of performance concerns the data flows and the probability of obtaining contents.

There is precedence for using queuing models for mobility: The Markovian highway PALM model is used for dimensioning cellular telephony for cars on a highway [14]. The one-dimensional topology is also considered for ad hoc networks in [16]. This model assumes a fixed number of nodes moving on a finite line with reflections at the end. A more general one-dimensional model is presented in [17]. It allows the selection of destination, speed and pause times to be correlated. Our work is based on the mobile infostation model in [15]. We extend it by considering the boundary effects for finite street segments, the generalization to a grid of streets, and a wider set of performance issues.

The mobility parameters of our model may be obtained from measurements in urban areas. However, we have not found a suitable dataset for this in the extensive CRAWDAD database [18]. The spatial information is often missing and the time-granularity of the measurements is too coarse for our scenarios; the measuring nodes typically search for contacts every two minutes.

The speed of the mobile nodes has large impact on the performance of the mobility-assisted content distribution. There is evidence that the walking speed of pedestrians is dependent on culture and the ‘pace of life' in different urban areas [20]. In sections 4 and 5 we study the performance effect of pedestrian walking speed. In the field of urban planning there have been studies on the walking speed and behavior of pedestrians for dimensioning traffic structures such as width of pavements, scheduling of traffic lights, maximum vehicle speed, as well as bus and train schedules [21] – [24]. We are however not aware of any models or data that have the fine granularity needed for comparison with our model.

3. THE STREET MODEL

We consider a scenario where nodes move on a two-way street segment and exchange data with other nodes in proximity. The street segment is such that the node arrivals and departures occur only at its endpoints, i.e. it is a segment of an actual street between two intersections. Nodes arrive at both endpoints, according to Poisson processes with parameters λ and λ'. The rationale for this assumption is that we consider people arriving independently from a large population. Temporal correlation among the arrivals tends to improve the performance of the system, as we show later in section 4. Since we are primarily concerned with the *achievable* performance, we consider Poisson arrivals to be representative scenario for our study.

The speed of the nodes are i.i.d. random variables with a probability density function $f_V(v)$ with the support [V_{min}, V_{max}]; $0 < V_{min} \leq V_{max} \leq \infty$. We assume that node encounters do not incur delay, i.e. nodes are bypassing each other freely, without lining up behind a node that moves slowly. The described scenario resembles the arrival and movement of pedestrians on a sidewalk that is wide enough to prevent collisions, but not wider than the transmission range of their mobile devices. We believe that this is realistic for low arrival rates of pedestrians to the street. From the viewpoint of the achievable performance, the continuous flow of people is the critical case since congestion/congregation points would incur longer contact times and thus facilitate the spread of the content. We describe a model that allows us to study the basic performance measures of a distribution system in this environment, such as:

- *Connectivity*: the percentage of time spent in the street segment when a node has at least one neighbor.
- *Contact rate*: the number of contacts per second that the node makes while being in the street segment.

- *Contact duration*: the life-time of the contacts that the node makes.

In order to obtain a tractable model, we impose the following assumptions and limitations:

- Contacts with nodes in other street segments are not possible, meaning that all connections break at the endpoints.
- Nodes do not change speed or direction while in a street segment, but they may do so upon entering a new segment.
- A node can be connected to multiple nodes within its transmission range.
- The amount of data exchanged between two nodes is proportional to the time they remain in contact.

Suppose that an observer node moves in the street segment at a speed v_o (Fig. 1). Let $L \geq 2\Delta$ be the street length and Δ the transmission range of a node. We distinguish between three types of contacts that the observer node may make: with slow nodes ($v < v_o$) that it overtakes, with fast nodes ($v > v_o$) that are overtaking, and with counter-directed nodes that are bypassing the observer node. Since node arrivals to the street are Poisson, it is easy to show that, for the each type, the number of contacts over an arbitrary time interval is also Poisson distributed, but with a time-dependent mean. Therefore, we model the observer node as an $M_t/G_t/\infty$ queue with three types of arrivals. The mean arrival rate to this observer queue is $\mu(v_o,t) = \mu_f(v_o,t) + \mu_s(v_o,t) + \mu_c(v_o,t)$, where μ_f, μ_s and μ_c are mean arrival rates for the fast, slow, and counter-directed nodes, respectively. If the speed distribution of nodes entering the street is uniform on [V_{min}, V_{max}], it can be shown that the mean arrival rates are given by the expressions in Table 1.

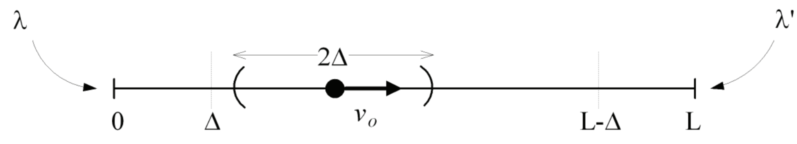

Fig. 1. An observer node equipped with a mobile device with a transmission range Δ travels along a street segment of length *L* at speed v_o.

Let $N(t)$ denote the total number of nodes in the observer queue at time *t* and $N_f(t)$, $N_s(t)$ and $N_c(t)$ the numbers of fast, slow, and counter-directed nodes, respectively:

$$\Pr\{N(t)=n\}=\Pr\{N_f(t)+N_s(t)+N_c(t)=n\} \tag{1}$$

Note that $\Pr\{N(0)=0\}\neq 0$ because all nodes within the first Δ meters of the street will be within the transmission range of the observer node at the moment when it enters the street. We assume that $N(0)=0$ to simplify the presentation; *the complete model includes the non-zero initial state of the observer queue*.

The probability of having *i* fast nodes in the observer queue can be written as:

$$\Pr\{N_f(v_o,t)=i\}=\sum_{j=i}^{\infty}\Pr\{N_f(v_o,t)=i \mid A_f(v_o,t)=j\}\Pr\{A_f(v_o,t)=j\} \tag{2}$$

where $A_f(t)$ is the arrival-counting process for fast nodes. Since the fast node arrivals constitute a non-homogeneous Poisson process with parameter $\mu_f(v_o,t)$, $A_f(t)$ is given by:

$$\Pr\{A_f(v_o,t)=j\}=\frac{m_f(v_o,t)^j}{j!}e^{-m_f(v_o,t)} \tag{3}$$

where $m_f(v_o,t)=\int_0^t \mu_f(v_o,\tau)d\tau$. The conditional probability for $N_f(v_o,t)$ given that $A_f(v_o,t)=j$ can be obtained as:

$$\Pr\{N_f(v_o,t)=i \mid A_f(v_o,t)=j\}=\binom{j}{i}q_f(v_o,t)^i\left(1-q_f(v_o,t)\right)^{j-i} \tag{4}$$

where $q_f(v_o,t)$ is the probability that a fast node, which has arrived at some time $0\le\tau\le t$, is still in service (within the transmission range) at time *t*. This probability depends on the service time, which we denote by $s_f(v_o,v,\tau)$, as:

$$q_f(v_o,t)=\int_0^t \Pr\{s_f(v_o,v,\tau)>t-\tau \mid \text{fast node arrival at } \tau\}\Pr\{\text{fast node arrival at } \tau\}d\tau\ . \tag{5}$$

Since the arrivals are Poisson, (5) becomes

$$q_f(v_o,t) = \frac{1}{m_f(v_o,t)} \int_0^t \Pr\left\{ s_f(v_o,v,\tau) > t-\tau \mid v > v_o \right\} \mu_f(v_o,\tau) d\tau \,. \quad (6)$$

On an infinitely long street (the highway model [15]) the service time would be $2\Delta/|v-v_o|$. However, on a finite street segment, the service time can be truncated because

- the observer node has just entered the street ($0<\tau<\Delta/v_o$),
- the observer node is just about to exit the street ($(L-\Delta)/v_o<\tau<L/v_o$), or
- the node or the observer node exit the street before $t=\tau+2\Delta/|v-v_o|$

Therefore, on a finite street, the service time $s_f(v_o,v,\tau)$ of a node depends on its speed v and the time τ when the contact had been established. It can be obtained by considering all possible ways in which a contact may end.

Probability $q_f(v_o,t)$ in (6) can be obtained from $s_f(v_o,v,\tau)$ and the conditional speed distribution of fast nodes that arrive to the observer queue $f_V(v|V>v_o)$. In the case of the uniform speed distribution, it is given by:

$$f_V(v|V>v_o) = \frac{1-\frac{v_o}{v}}{V_{\max} - v_o\left(1+\ln\frac{V_{\max}}{v_o}\right)}\,, \quad (7)$$

for $v_o < v \le V_{\max}$. Finally, from (2), (3), and (4):

$$\Pr\left\{N_f(v_o,t)=i\right\} = \frac{\left(m_f(v_o,t)q_f(v_o,t)\right)^i}{i!} e^{-m_f(v_o,t)q_f(v_o,t)} \quad (8)$$

Hence, the number of fast nodes connected to the observer node is Poisson distributed with time-dependent mean $m_f(v_o,t)q_f(v_o,t)$. It is easy to show that the numbers of connections with slow and counter-directed nodes are also Poisson distributed with means $m_s(v_o,t)q_s(v_o,t)$ and $m_c(v_o,t)q_c(v_o,t)$, respectively. Therefore,

$$\Pr\left\{N(v_o,t)=n\right\}=\frac{\rho(v_o,t)^n}{n!}e^{-\rho(v_o,t)}\,, \tag{9}$$

where $\rho(v_o,t)=m_f(v_o,t)q_f(v_o,t)+m_s(v_o,t)q_s(v_o,t)+\;m_c(v_o,t)q_c(v_o,t)$. The probability that the observer node is connected (to at least one node) at time *t* is given by:

$$\Pr\left\{N(v_o,t)\neq 0\right\}=1-e^{-\rho(v_o,t)}\;. \tag{10}$$

Connectivity of the observer node is obtained by averaging the connection probability (10) over time and over all possible speeds v_o. For uniform speed distribution, this becomes

$$\frac{1}{(V_{\max}-V_{\min})}\frac{1}{L}\int_{V_{\min}}^{V_{\max}} v_o \int_0^{L/v_o}\left(1-e^{-\rho(v_o,t)}\right)dt\,dv_o\;. \tag{11}$$

For $L>>\Delta$, (11) simplifies to:

$$1-e^{-\frac{2\Delta(\lambda+\lambda')}{V_{\max}-V_{\min}}\log\frac{V_{\max}}{V_{\min}}}\;. \tag{12}$$

The average rate at which data disseminates via intermittent connections is given by the product of the connectivity and the data rate offered by the radio interface.

Average contact rate can be obtained from the node arrival rates given in Table 1:

$$\frac{1}{(V_{\max}-V_{\min})}\frac{1}{L}\int_{V_{\min}}^{V_{\max}} v_o \int_0^{L/v_o}\mu(v_o,t)\,dt\,dv_o\;. \tag{13}$$

Table 1. Mean arrival rates at the observer queue for fast, slow, and counter-directed nodes (← means "same as in the previous column").

μ \ t	$0<t\leq\frac{\Delta}{v_o}$	$\frac{\Delta}{v_o}\leq t<\frac{L-\Delta}{v_o}$	$\frac{L-\Delta}{v_o}\leq t<\frac{L}{v_o}$
$\mu_f(v_o,t)$	$\frac{\lambda}{V_{\max}-V_{\min}}\left(V_{\max}-v_o\right)$	$\frac{\lambda}{V_{\max}-V_{\min}}\left(V_{\max}-v_o\left(\ln\frac{V_{\max}}{v_o}+1\right)\right)$	←
$\mu_s(v_o,t)$	$\frac{\lambda}{V_{\max}-V_{\min}}v_o\ln\frac{v_o}{V_{\min}}$	$\frac{\lambda}{V_{\max}-V_{\min}}\left(V_{\min}+v_o\left(\ln\frac{v_o}{V_{\min}}-1\right)\right)$	0
$\mu_c(v_o,t)$	$\frac{\lambda'}{V_{\max}-V_{\min}}\left(V_{\max}-V_{\min}+v_o\ln\frac{V_{\max}}{V_{\min}}\right)$	←	←

The tail distribution of the contact durations, $\bar{F}_T(t) = P(T > t)$, is also of interest because it gives us the percentage of useful contact (we omit the derivation for brevity).

4. PERFORMANCE RESULTS

In this section, we show performance results from the street model and compare them with the results from the simulation implementation described next.

4.1 Simulations

We have built a simulation model of pedestrian content distribution using the Omnet++ [25] discrete event simulator and the mobility framework extension [26].

In the simulator, we attach a node generator to the street endpoints. Each generated node independently selects a random speed from an arbitrary but known probability distribution. The node then traverses the street at the constant speed. Whenever the distance between two nodes is less than or equal to the transmission range Δ they are connected and can communicate with each other. For each node, we collect a number of attributes such as the connectivity, contact rate, and contact duration. We introduce a warm-up period in our simulation runs to minimize the effect of the initial transient. The minimum length of the warm-up period is determined using the graphical procedure of Welch [27]. We conduct 10 runs where in each run we collect statistics from 1000 nodes.

4.2 Street model

We present performance results from the analytical and simulation models for a street of length L = 100 m, transmission range Δ = 10 m and symmetric Poisson arrival rates $\lambda = \lambda'$. We also provide simulation results for other arrival processes and inter-arrival time distributions with the same mean arrival rate λ: Poisson batch, Markov modulated Poisson process (MMPP), hyper-exponential, and power-law (Pareto).

Fig. 2 (left) shows the effect of the speed distribution on the average connectivity of nodes. We have selected three different distributions, all with the same mean of 1 m/s but with different variance: a constant speed, a uniform distribution with the support [0.1, 0.9] m/s, and an exponential distribution with

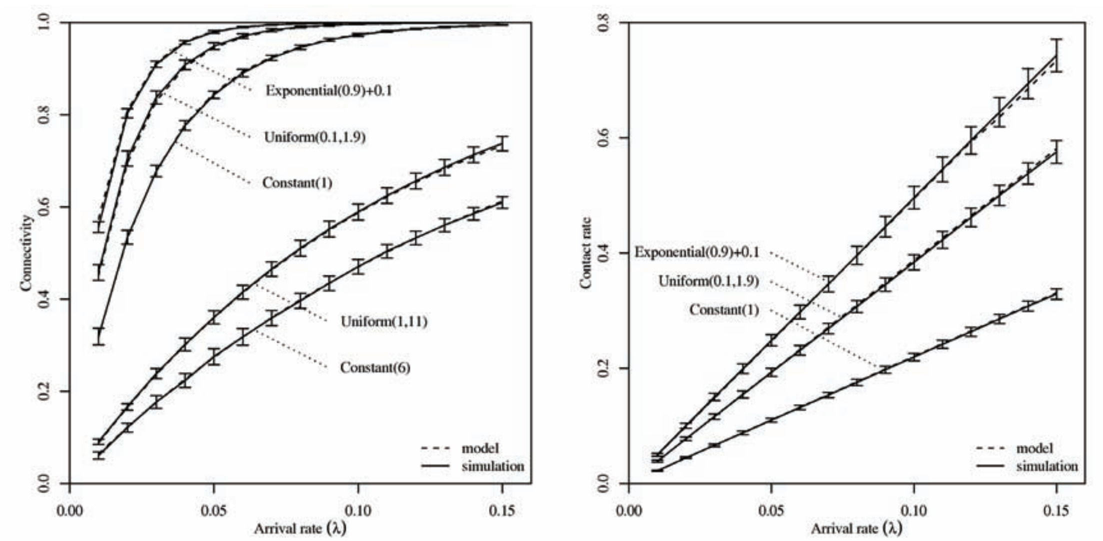

Fig. 2. Average connectivity (left) and average contact rate (right) as functions of the arrival rate in nodes per second. The pedestrian traffic is symmetric (λ = λ'), the length of the street is 100 m and the transmission range is 10 m.

mean 0.9 m/s that is shifted by 0.1 to have the same lower bound and mean as for the Uniform[0.1, 1.9] distribution. We see that the connectivity of the nodes increases with the arrival rate and with the variance of the speed distribution. The contact rate also increases with the variance of the speed distribution, as shown in Fig. 2 (right). It increases linearly with the arrival rate ($\mu(v_o, t)$ in (13) is a linear function of $\lambda = \lambda'$).

Although we focus on content distribution of pedestrians, the model itself does not make any specific assumptions thereof. In Fig. 2 (left) we have plotted the connectivity for constant speed of 6 m/s and a uniform distribution with the same mean: Uniform[1,11]. These speeds could represent vehicles (bicycles, cars, etc.) in an urban area. At higher speeds, the contacts become shorter which reflects in lower connectivity compared to pedestrian speeds. Therefore, a longer transmission range Δ is needed to achieve comparable performance for given arrival rates λ and λ'. Note that the connectivity scales with the product $(\lambda + \lambda')\Delta$ in (12), and therefore an increase of the transmission range has the same effect on the connectivity as an increase in the arrival rate.

The duration of a contact is a random variable T whose tail distribution $\bar{F}_T(t_{\min}) = P(T > t_{\min})$ is plotted in Fig. 3 for two different speed distributions. Here $t_{\min}$ is a minimum required contact duration, which includes both the conection setup and the actual data transfer. Contacts that are shorter than t_{min} are considered useless for content distribution. The minimum required contact duration is one of a few parameters that can be engineered in the system and it is therefore a key performance issue to minimize $t_{\min}$. The shape of the tail distribution for constant speed is due to the fact that contacts in the forward direction last for the whole length of the street and those with nodes in the opposite direction last for Δ/v s, when the node speed is v.

In order to verify whether the Poisson arrivals can be representative for a broader class of arrival processes, we run a set of simulations, for which the results are shown in Fig. 4. The connectivity and contact rates were measured for four different arrival processes with the same mean arrival rate λ: i) Poisson batch arrivals with a Uniform[1,7] batch size distribution and batch arrival rate λ/4, ii) two-phase hyperexponential inter-arrival time distribution with arrival rates 0.35λ and 5.7λ in the first, respectively, second phase, and selection probabilities 0.31 and 0.69, iii) Pareto inter-arrival time distribution with shape parameter 1/(1-λ), and iv) four-stage Markov modulated Poisson process with arrival rates [0.5, 0.8,

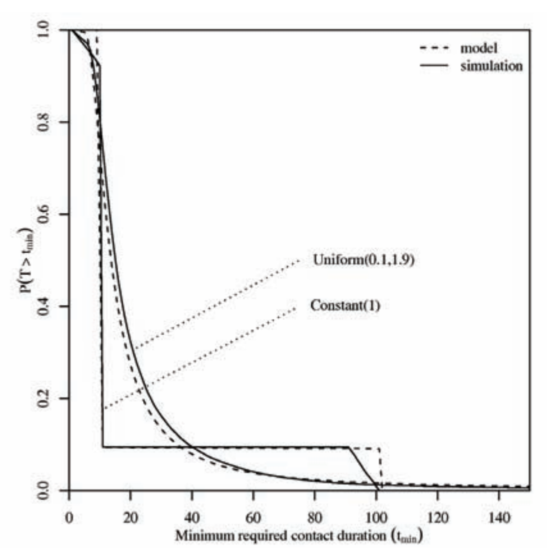

Fig. 3: Tail distribution of contact durations for Constant[1] and Uniform[0.1,1.9] speed distributions.

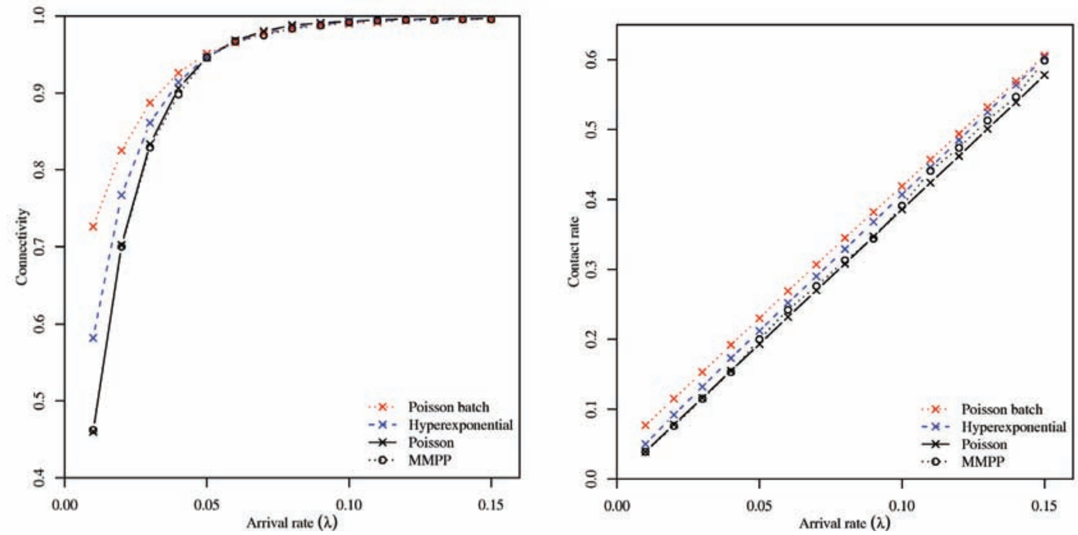

Fig. 4. Average connectivity (left) and average contact rate (right) for various arrival processes. The pedestrian traffic is symmetric (λ = λ'), the length of the street is 100 m and the transmission range is 10 m.

1.2, 1.5]λ and equal state probabilities with mean sojourn times 300s. In all simulations, nodal speeds take a Uniform[0.1,1.9] distribution. Results for the Pareto distribution are not shown in Fig. 4 because they cannot be distinguished from results for Poisson arrivals. From both the connectivity and contact rate, we see that the Poissonian assumption captures the qualitative behavior of the system, although the numerical results show some differences. The performance attained with the Poisson arrivals falls on the lower side (not a lower bound, strictly speaking) and therefore, we argue that Poissonian assumption may be used to obtain a measure of achievable performance for this system.

In summary, from the results presented in this section we see that the average connectivity and contact rate obtained from the analytical model are all within the 95% confidence intervals of the simulation results. The tail distributions for the contact duration also match closely. Therefore, we can rely on the model to study the pedestrian content distribution. We have shown this far how the performance measures of the system depend on the arrival rate (density) of the users in an area and the variance of their speed distribution.

5. CONTENT DISTRIBUTION

In this Section, we study the content distribution in a topology represented by a grid of streets, which is typical for an urban area. We start with the analysis of a single street segment, for which the results are later used to study the content distribution in a topology that represents a part of a downtown area in Stockholm.

5.1 Content distribution in a street segment

Let p_0 and q_L be the percentages of nodes that bring contents when they arrive to the near and far ends of a street segment of length L (Fig. 5). Our objective is to determine the spatial distribution of the content $p(x)$ in the forward and $q(x)$ in the opposite direction, and particularly the percentages p_L and q_0 of nodes that will possess the content when they exit the street segment. In a grid of streets, which can be built by concatenating multiple street segments, inputs (p_0, q_L) and outputs (p_L, q_0) of a street segment become outputs, respectively, inputs of neighboring segments. This allows us to study the content distributions in a city area in Section 5.2 using a simple recursive algorithm.

We again use the notion of an observer node to find the content possession probabilities $p(x)$ and $q(x)$. The probability that a random observer, which moves in the forward direction, possesses the content at $x+\Delta x$ is

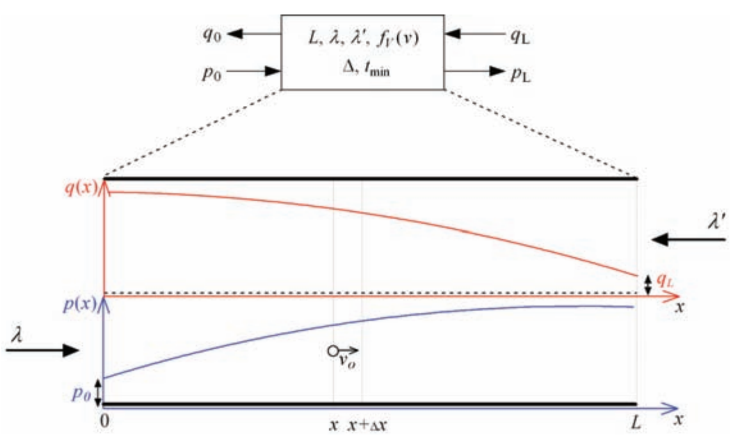

Fig. 5. Content distribution on a street segment. The pedestrians enter with average rates λ and λ'.

$$p(x+\Delta x) = p(x) + (1-p(x))\theta(x) , \qquad (14)$$

where $\theta(x)$ is the probability that it meets and connects to a node with the content on $[x, x+\Delta x]$. Since the arrivals of fast, slow, and counter-directed nodes to the observer node are Poisson distributed with means μ_f, μ_s, and μ_c, respectively, probability of meeting at least one of the nodes with the content and establishing a contact that will last for at least $T > t_{\min}$ seconds is given by:

$$\theta(v_o, x) = 1 - e^{-\left(\mu_f(v_o, x/v_o)\bar{F}_{T_f}(t_{\min})p(x) + \mu_s(v_o, x/v_o)\bar{F}_{T_s}(t_{\min})p(x) + \mu_c(v_o, x/v_o)\bar{F}_{T_c}(t_{\min})q(x)\right)\frac{\Delta x}{v_o}} , \qquad (15)$$

given that the observer node moves with the speed v_o. Since $1-e^{-\varepsilon} \approx \varepsilon$ for $\varepsilon << 1$:

$$\theta(v_o, x) = \left(\mu_f(v_o, x/v_o)\bar{F}_{T_f}(t_{\min})p(x) + \mu_s(v_o, x/v_o)\bar{F}_{T_s}(t_{\min})p(x) + \mu_c(v_o, x/v_o)\bar{F}_{T_c}(t_{\min})q(x)\right)\frac{\Delta x}{v_o} . \qquad (16)$$

Probability $\theta(x)$ is obtained by averaging (16) over the speed of the observer node v_o and it can be written:

$$\theta(x) = \left(p(x)a(x) + q(x)b(x)\right)\Delta x , \qquad (17)$$

where:

$$\begin{aligned} a(x) &= \frac{1}{V_{\max} - V_{\min}} \int_{V_{\min}}^{V_{\max}} \frac{1}{v_o}(\mu_f(v_o, x/v_o)\bar{F}_{T_f}(t_{\min}) + \\ &\quad + \mu_s(v_o, x/v_o)\bar{F}_{T_s}(t_{\min}))\, dv_o \\ b(x) &= \frac{1}{V_{\max} - V_{\min}} \int_{V_{\min}}^{V_{\max}} \frac{1}{v_o}\mu_c(v_o, x/v_o)\bar{F}_{T_c}(t_{\min})\, dv_o \end{aligned} \qquad (18)$$

assuming that v_o is uniformly distributed on $[V_{\min}, V_{\max}]$. After substituting (17) in (14) and letting $\Delta x \to 0$, the following differential equation is obtained:

$$p'(x) = (1-p(x))(p(x)a(x) + q(x)b(x)) , \qquad (19)$$

where $p'(x)$ denotes the derivative of $p(x)$ with respect to x. Similarly for $q(x)$:

$$q'(x) = -(1-q(x))(p(x)c(x) + q(x)d(x)) , \qquad (20)$$

where $c(x) = b(x)$ and $d(x) = a(x)$ if $\lambda = \lambda'$. When $L >> \Delta$, μ_f, μ_s, μ_c become independent of *x* (they are given by the middle column of Table 1. in that case) and therefore $a(x)$, $b(x)$, $c(x)$, and $d(x)$ become constants. Then the system of differential equations given by (19)-(20) can be easily solved with the initial conditions $p(0) = p_0$ and $q(L) = q_L$. This assumption is not essential, but it greatly simplifies the solution to (19)-(20). It slightly overestimates the possession probabilities $p(x)$ and $q(x)$ by ignoring the boundary effects in the first and the last Δ meters of the street.

We show the content dispersion in the forward direction $p(x)$ for various speed distributions and minimum required contact durations in Fig. 6 (left). Due to the symmetry, content dispersion in the opposite direction is simply $q(x) = p(L - x)$ and it is not shown in the figure. Our scenario assumes the following values of the street parameters: $\Delta = 10$ m, $\lambda = \lambda' = 0.02$, and $p_0 = q_L = 5$ %. It can be observed from Fig. 6 (left) that the content spreads more efficiently when the variance of the nodal speed distribution increases. As shown in Section IV, the larger speed variance results in a larger number of contacts and longer contact durations, which facilitate the content distribution. In the case of minimum contact duration $t_{\min} = 10$ s, the content spreads very efficiently in spite of the very low arrival rate (one arrival per 50 s on average) and low percentages of nodes bringing the content to the street ends

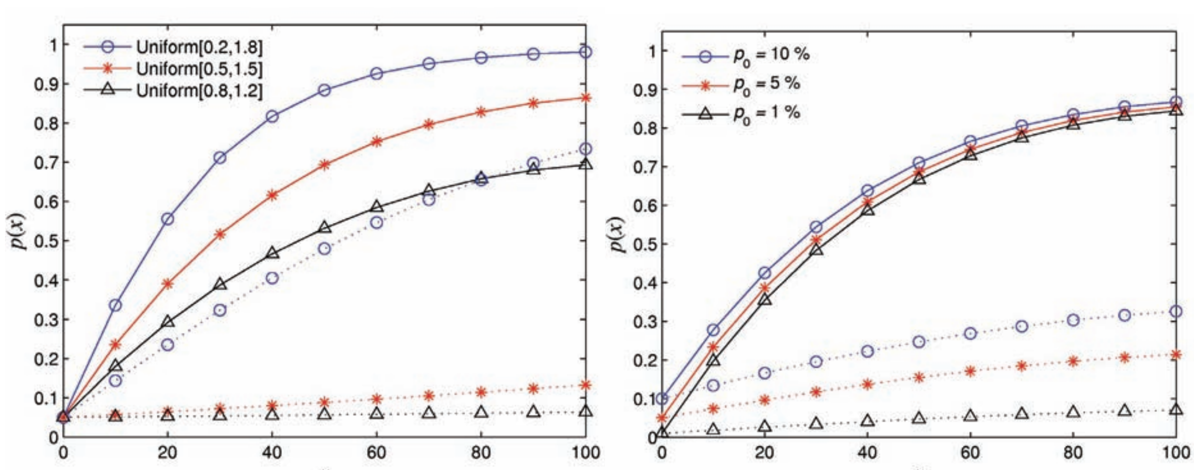

Fig. 6. Left: content dispersion for $t_{\min} = 10$ s (solid) and $t_{\min} = 15$ s (dotted) for various speed distributions. $\lambda = \lambda' = 0.02$ s^{-1}. Right: content dispersion for $\lambda = \lambda' = 0.02$ s^{-1} (solid) and $\lambda = \lambda' = 0.01$ s^{-1} (dotted) for various percentages p_0 of forward nodes that arrive with the content and uniform speed distribution on [0.5, 1.5].

($p_0 = q_L = 5\ \%$). However, the content distribution can be hampered by the requirements on the minimum contact duration, as in the case when $t_{\min} = 15\ \text{s}$, which is also shown in Fig. 6 (left). It is therefore important to minimize the connection setup time in order to reduce the required contact duration. Some work has been done on improving the time-efficiency of service discovery in Bluetooth [28].

In Fig. 6 (right), we vary the percentage p_0 of forward nodes that arrive with the content. We assume that nodes traveling in the opposite direction do not bring the content to the street ($q_L = 0$). The scenario assumes the following street parameters: $\Delta = 10\ \text{m}$, $t_{\min} = 10\ \text{s}$, and $V \sim \text{Uniform}[0.5,1.5]$. We may notice that, if the arrival rates of nodes are sufficiently large ($\lambda = \lambda' = 0.02$), the content dispersion depends very little on the arrival rate of the nodes with the content. The reason is that the content becomes resident in the street as long as there are new nodes to which the content can be passed before a node disappears from the street. For $\lambda = 0.02$ and $p_0 = 1\ \%$, the content arrives to the street every 83 minutes on average, yet every node that passes through this street will obtain the content with the probability of 85 %. Remark that the 83 minutes widely exceeds the sojourn time of the nodes in the street. Even though this analytic model does not provide the means to analyze non-recurrent content arrivals, such as when a single node injects the content to the street, these results indicate that the content will be still present in the street long after the node departs, even in that case. This effect becomes more pronounced as the length of the street segment increases. When the arrival rates decrease, the content dispersion becomes less efficient and the ability of the street to act as a virtual storage disappears, which is illustrated in Fig. 6 (right) for the case when $\lambda = \lambda' = 0.01$. The critical arrival rate, for which the dispersion performance becomes significantly affected by p_0, depends on the variance of the nodal speed distribution.

5.2 Content distribution in an urban area

Based on the street model described in the previous section, we evaluate the efficiency of mobility-assisted content distribution on a topology shown in Fig. 7. The equivalent grid consists of 29 street segments whose lengths vary between 20 m and 200 m. There are 12 passages that connect this area to the outside world: we assume that the arrival rates to the passages are $\lambda_i = \lambda,\ i = 1,\ldots,12$. Upon arriving at an intersection, nodes continue to move on the same street (if possible) with probability 0.5 or turn to other adjoining streets with equal probabilities (the alternative of choosing among all the routes with equal probability extends the sojourn times of the nodes in the area and hence shows better performance; it does not otherwise affect the results). Nodes with the content constitute p_{in} percents of the nodes that arrive to the first street (hence, the content arrival rate is λp_{in}); it is marked in Fig. 7 (right). The source of the content could be, for example, an access point located close to the first street segment. The performance metric of interest is *dispersion*, which represents the percentage of nodes in the area that possess the content in the steady state.

To illustrate the spatial spreading of the content, we assume the following parameters in our model: uniformly distributed nodal speeds in [0.5, 1.5], $\Delta = 10$ m, $t_{\min} = 20$ s, $\lambda = 0.05$, and $p_{in} = 5$ %. The dispersion for this case is 77.5 %, and the spatial distribution of the content is shown in Fig. 7 (right). The content distribution is fairly efficient: there are no indications that the percentage of users with the content depends on the distance from the source.

Besides the variance of the speed distribution, the minimum required contact duration $t_{\min}$ has the largest impact on the system performance. Its effect on the content dispersion for various arrival rates is shown in

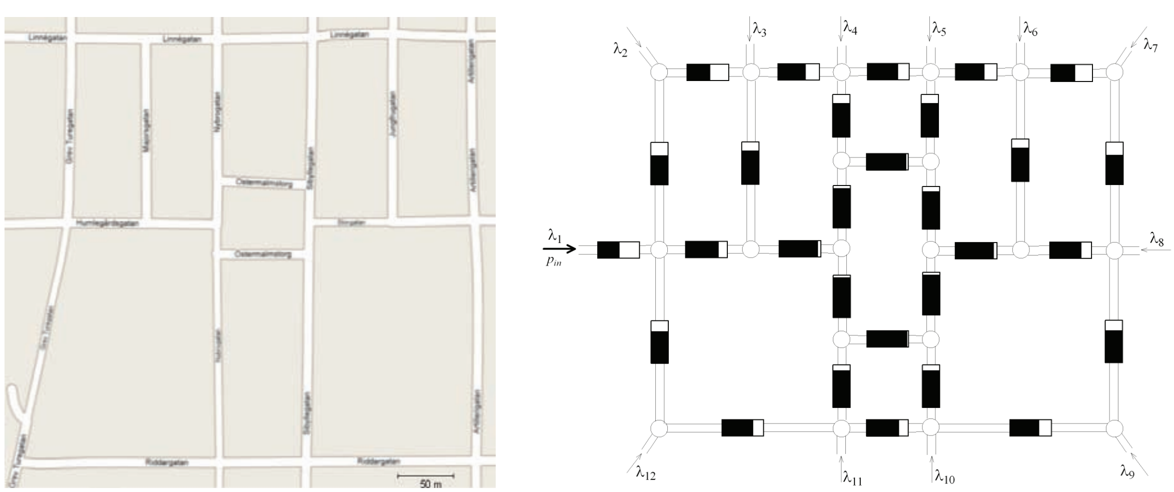

Fig. 7. A part of Stockholm's downtown area (left) and the spatial distribution of the content in a corresponding network of street segments (right). Darkened areas denote the percentages of nodes with the content.

Fig. 8 (left). It confirms that critical cases are when the arrival rate is very low. The tail distribution of the contact duration $\bar{F}_T(t_{\min})$ affects the dispersion through the coefficients of differential equations (18) and (19), and it is a function of $\Delta / t_{\min}$. Since extending the range of mobile devices Δ is coupled with many problems, such as the increased interference and power consumption, minimum required contact duration $t_{\min}$ is the system parameter that is most likely to be engineered to improve the performance. In that respect, several issues can be addressed, including the segmentation of the contents into atomic units of optimal size and, as mentioned before, the time-efficiency of service discovery.

Finally, in Fig. 8 (right) we show that the percentage of nodes that arrive with the contents p_{in} does not have significant impact on the content dispersion, except in the case of very low arrival rates. Therefore, the requirements on the fixed infrastructure, which may be needed to provide new contents, are minimal: one access point may serve a relatively large geographical area. We conducted a set of simulations to verify the results for $p_{in} = 10\%$. Results summarized in Table 2. indicate that the model gives a fairly good estimate of the dispersion. The difference comes mainly from the simplifying assumption made in the Section 5.1 that the length of each street segment is considerably larger than the transmission range, which is not the case in the topology under consideration.

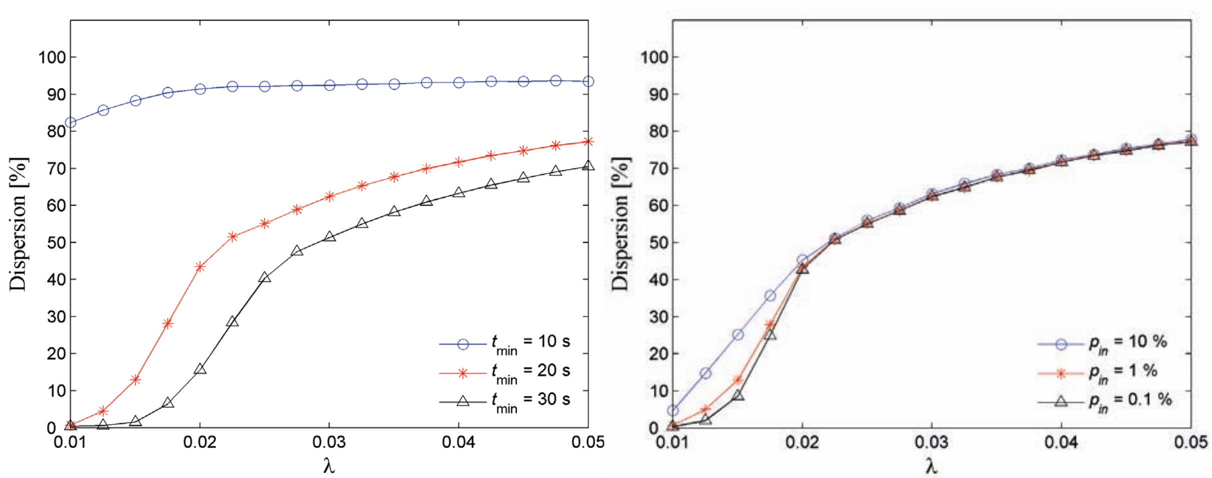

Fig. 8. Effect of $t_{\min}$ on content dispersion for various arrival rates and p_{in}=5% (left) and content dispersion for various percentages p_{in} of nodes that arrive with the content (right)
($V \sim U[0.5,1.5]$, $\lambda = 0.02\,s^{-1}$, $\Delta = 10$ m, $t_{\min} = 20$ s).

Table 2. Content dispersion for $p_{in} = 10\%$ ($V \sim U[0.5, 1.5]$, $\lambda = 0.02\,\mathbf{s^{-1}}$, $\Delta = 10$ m , $t_{\min} = 20$ s) obtained from the model and simulations (95 % confidence intervals are indicated).

Disp. [%] \ λ	0.01	0.02	0.03	0.04	0.05
Model	5.1	42.7	62.6	71.8	76.5
Simulations	5.2 [3.1-7.3]	39.0 [36.4-41.5]	58.0 [56.6-59.5]	68.3 [67.4-69.3]	73.7 [72.8-74.5]

5.3 Transient behavior

The results presented so far are for the steady state. Considering the complexity of analyzing the transient behavior analytically, we resort to simulations. Our simulator is described in Section 4.1. The scenario under study is similar to that of the previous section: We consider the topology shown in Fig. 7. We assume that the node transmission range is Δ = 10m and the contact setup time is $t_{\min} = 20$ s. In each street, a node chooses its speed from a uniform distribution with support [0.5, 1.5]. We assume that there is a single access point, located at the intersection closest to the first street segment, which provides a content that is of interest to all nodes. Nodes start downloading the content at t = 0 (after a simulation warm-up period), either directly from the access point or from peers that already have it. We consider five different arrival rates λ and for each arrival rate we conduct 100 simulation runs. We collect the time-series of the fraction of nodes carrying the content (Fig. 9). The results confirm that the spreading of the content is strongly dependent on the density of the nodes in the area. For λ = 0.05, approximately 75% of the nodes are carrying the content in steady state. It takes approximately 1200 s to reach the steady state. For λ = 0.01 the average fraction of nodes carrying the content is much lower (~40%), and it takes close to 2000 s to reach the steady state.

It is interesting to study the effect of switching off the access point. In Fig. 9, we switch off the access point at t = 2000 s (when the system has reached its steady state for all arrival rates). We note that when the arrival rate is low (λ = 0.01) the content disappears from the area because the density of nodes in the area is not high enough to facilitate the ad hoc spreading. At higher arrival rates, once the steady-state is

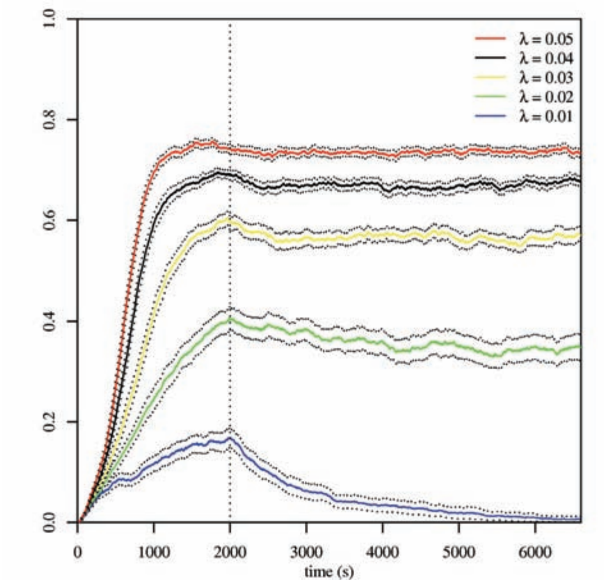

Fig. 9. Dispersion as a function of time. The access point is switched off at time 2000 s. Dotted lines indicate 95% confidence intervals.

reached, the spreading is not any more dependent on the infrastructure support: the content stays in the area as long as there are new nodes to which it can be passed. In other words, the spreading process exhibits a virtual storage effect: the content resides in the area even though there is no infrastructure support and nodes are coming and going.

6. CONCLUSION

We have considered mobility-assisted content dissemination to an arbitrarily large group of pedestrian nodes with short-range radio connectivity. The service provides content distribution even if ubiquitous coverage is not feasible, and it offers a new ad hoc distribution mode for contents that originate from the mobile nodes. We developed a detailed analytical model to study the connectivity properties of street mobility at low arrival rates of pedestrian; the critical case for the system. The analytic results are compared to simulation results with good agreement. The model is extended to a network of streets to model larger areas. We focus on areas characterized by low user density, where people usually do not congregate or swarm. The reported results give an insight into various system parameters and how they affect the content distribution. We have found that the content spreads with high efficiency in a large number of common-case scenarios. We will hence continue to evaluate the system by considering data on the flows of people in our city, which we hope to obtain from a local public authority and by studying the availability of different types of contents and the corresponding caching policy that determines what data to request in a contact. Finally, we intend to include energy consumption as a performance measure.

REFERENCES

[1] Delay Tolerant Networking Research Group. [Online]. Available: http://www.dtnrg.org/wiki. [Acessed June 30, 2007].

[2] M. Izal, G. Urvoy-Keller, E. Biersack, P. Felber, A. A. Hamra, and L. Garćes-Erice, "Dissecting BitTorrent: Five Months in a Torrent's Lifetime," *Proc. Passive and Active Measurements Conference*, Antibes Juan-les-Pins, France, April 2004.

[3] S.M. Hedetniemi, S.T. Hedetniemi and A. Liestman, "A Survey of Gossiping and Broadcasting in Communication Networks," *Networks*, vol. 18, No. 4, pp. 319-349, 1988.

[4] A. Vahdat and D. Becker, "Epidemic Routing for Partially Connected Ad Hoc Networks," Tech. rep. CS-200006, Duke University, NC, April 2000.

[5] M. Papadopouli and H. Schulzrinne, "Effects of power conservation, wireless coverage and cooperation on data dissemination among mobile devices," *Proc. ACM SIGMOBILE MobiHoc*, Long Beach, California, 2001.

[6] J. Chesterfield and P. Rodriguez, "DeltaCast: efficient file reconciliation in wireless broadcast systems," *Proc. ACM Mobisys*, Seattle, WA, 2005.

[7] A. El Fawal, J.-Y. Le Boudec, K. Salamatian, "Self-Limiting Epidemic Forwarding," Tech. rep. LCA-REPORT-2006-126, EPFL, Switzerland, July 2006.

[8] W. Zhao, M. Ammar, E. Zegura, "Multicasting in Delay Tolerant Networks: Semantic Models and Routing Algorithms," *Proc. ACM Sigcomm Workshop on Delay Tolerant Networking*, Philadelphia, PA, Aug. 2005.

[9] S. Jung, U. Lee, A. Chang, D-K Cho and M. Gerla, "BlueTorrent: Cooperative Content Sharing for Bluetooth Users," *Proc. IEEE Pervasive Computing and Comm* (*PerCom '07*), March 2007, pp. 47-56.

[10] T. Camp, J. Boleng, and V. Davies, "A Survey of Mobility Models for Ad Hoc Network Research," *Wireless Communications and Mobile Computing*, vol. 2, no. 5, pp. 483-502, Aug. 2002.

[11] E. Hyytiä, P. Lassila and J. Virtamo, "Spatial Node Distribution of the Random Waypoint Mobility Model with Applications," *IEEE Trans. Mobile Computing*, vol. 5, no. 6, pp. 680-694, 2006.

[12] J.-Y. Le Boudec and M. Vojnovic, "Perfect Simulation and Stationarity of a Class of Mobility Models," *Proc. IEEE INFOCOM*, Miami, FL, 2005.

[13] Th. Spyropoulos, K. Psounis, C. S. Raghavendra, "Mobility models: Performance analysis of mobility-assisted routing," *Proc. ACM MobiHoc*, Florence, Italy, May 2006.

[14] K. K. Leung, W. A. Massey and W. Whitt, "Traffic Models for Wireless Communication Networks," *IEEE Journal on Selected Areas in Communication*, vol. 12, No. 8, 1994, pp. 1353-1364.

[15] W. H. Yuen, R. D. Yates and C-W Sung, "Effect of Node Mobility on Highway Mobile Infostation Networks," *Proc. ACM MSWiM*, San Diego, USA, Sep. 2003.

[16] R. Groenevelt, E. Altman and P. Nain, "Relaying in mobile ad hoc networks: The Brownian motion mobility model," *Springer/ACM Wireless Networks*, vol. 12, no. 5, Oct. 2006.

[17] D. N. Alparslan and K. Sohraby, "A Generalized Random Mobility Model for Wireless Ad Hoc Networks and Its Analysis: One-Dimensional Case," *IEEE/ACM Trans. Networking*, vol. 15, no. 3, June 2007.

[18] J. Yeo, D. Kotz, and T. Henderson, "CRAWDAD: A Community Resource for Archiving Wireless Data at Dartmouth" *ACM Sigcomm Comp. Comm. Review*, vol. 36, no. 2, pp. 21-22, April 2006.

[19] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, Richard Gass and James Scott, "Impact of Human Mobility on the Design of Opportunistic Forwarding Algorithms," *Proc. IEEE INFOCOM*, Barcelona, Spain, April 2006.

[20] R. Wiseman, "The Pace of Life." [Online]. Available: http://www.richardwiseman.com/quirkology/pace_home.htm. [Acessed: June 30, 2007].

[21] S. P. Hoogendoorn and P. H. L. Bovy, "Pedestrian Travel Behaviour Modeling," *Networks and Spatial Economics*, vol. 5, no. 2, June 2005.

[22] K. Fitzpatrick, M. A. Brewer and S. Turner, "Another Look at Pedestrian Walking Speed," *Transportation Research Record: Journal of the Transportation Research Board*, no. 1982, pp. 21-29, 2006.

[23] J. Y. S. Lee and W. H. K. Lam, "The Variation of Walking Speeds on a Uni-directional Walkway and on a Bi-directional Stairway," *Transportation Research Record: Journal of the Transportation Research Board,* no. 1982, pp. 122-131, 2006.

[24] X. Ma and I. Andréasson, "Predicting the effect of various ISA penetration grades on pedestrian safety by simulation," *Accident Analysis & Prevention*, vol. 37, no. 6, 2005.

[25] A. Varga, "The OMNeT++ discrete event simulation system," *Proc. European Simulation Multiconference*, Prague, Czech Republic, 2001.

[26] W. Drytkiewicz, S. Sroka, V. Handziski, A. Köpke and H. Karl, "A mobility framework for OMNeT++," 3rd International OMNeT++ Workshop, January 2003.

[27] A. M. Law and W. D. Kelton. Simulation Modeling and Analysis. 3rd edt., McGraw-Hill, 2000.

[28] Sedov; S. Preuss, M. Haase, C. Cap and D. Timmermann, "Time and Energy Efficient Service Discovery in Bluetooth," Proc. 57th IEEE Vehicular Technology Conference (VTC), pp. 418-422, Jeju, Korea, April 2003.