An analytical model for pedestrian content distribution in a grid of streets

Vladimir Vukadinović, Ólafur Ragnar Helgason, Gunnar Karlsson

Disney Research Zurich, Switzerland
KTH – Royal Institute of Technology, School of Electrical Engineering and Linnaeus Center ACCESS, Sweden

ARTICLE INFO

Article history:
Received 26 August 2011
Received in revised form 19 January 2012
Accepted 3 March 2012

Keywords:
Mobility modeling
Ad hoc network
Opportunistic networking
Delay-tolerant networking
Content distribution

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. The transport mechanism is the flow of people and it can be studied but not engineered. We study the efficiency of pedestrian content distribution 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. The results show that contents spread well with pedestrian speeds already at low arrival rates into a studied region. Our contributions are both the queuing analytic model that captures the flow of people and the results on the feasibility of pedestrian content distribution.

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 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 consider 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 content and we are primarily concerned with pedestrian mobility.

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, $\Delta$ of one another. Interference, fading, power control and other data-link issues 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 contents 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 (see [1] for a description). 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. The content may either be generated by the mobile nodes or they may be provided to the nodes through a sparse infrastructure (e.g. public WLANs). 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 content, as well as the rate and length distribution of contacts. We use both analysis and simulation in the study. Our analytical model shows explicitly how the mobility and system parameters affect the performance. Our simulations employ Legion Studio, an agent-based state-of-art pedestrian mobility simulator used in urban and traffic planning. It captures various aspects of the pedestrian mobility that are abstracted away in the model (details are provided in Section 4.1). The simulations are used to identify scenarios where our model provides a good approximation of reality, and to study scenarios that cannot be captured by the model. 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 can 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. This paper extends our previously published results [36] with more rigorous validation of the analytical model using Legion Studio and with new use case scenarios.

The structure of the paper is as follows. We review related work in Section 2. The street model is introduced in Section 3. 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.

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 of 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 [4]. The application of gossiping protocols to mobile communication has been proposed in, for instance, [5–7]. Multicast for delay-tolerant networks has been proposed in [8,9]. The type of mobile content distribution system that we analyze is described in [10,11].

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 [12]. The random waypoint model is notable owing to its popularity; its stationary distribution of nodes is provided in closed form in [13]. The random-trip model is a general mobility model that allows perfect simulation and general topologies [14]. Mobility-assisted routing for mobility in two-dimensions has been studied in [15]. The pocket switched experiment and analysis is primarily concerned with contact opportunities for use in opportunistic routing [16].

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 [17]. The one-dimensional topology is also considered for ad hoc networks in [18]. 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 [19]. It allows the selection of destination, speed and pause times to be correlated. Our work is based on the mobile infostation model in [20]. 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.

Pedestrian content distribution can be studied based on the mobility traces obtained from measurements in urban areas. However, we have not found suitable datasets for this in the extensive CRAWDAD and MobiLib databases [21,22]. The time and/or space granularity of the measurements is too coarse for our scenarios; we study extreme cases of intermittency, where contacts of the order of a few tens of seconds are used to exchange data chunks. Available traces are obtained in experiments where measuring nodes (e.g., PDAs/iMotes distributed to students or conference participants) typically search for contacts every two minutes. Traces with finer time resolution are available from experiments in infrastructure based WLANs, where measuring nodes record the IDs of reachable APs [23]. It is however impossible to extract accurate locations of nodes and estimate contact rates and durations from these traces. Most of the trace-based studies in the area of DTNs are focused on store-carry-forward routing for unicast [24–26]. Performance metrics in these works are inter-contact times for particular pairs of nodes and periods of re-appearance of a node at a particular location. Due to the granularity problem, rough approximations are used to estimate these metrics (e.g., two nodes are assumed to be in contact if they are associated with the same AP). Our focus is on multicast/broadcast applications. Therefore, we are interested in the rate of contacts that a node establishes with any node interested in spreading the content, which largely depends on the density of nodes in the studied area. Available traces are however related to very specific scenarios with given densities and their validity is difficult to generalize.

The speed of the mobile nodes has a 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 [27]. 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 [28–30].

3. Mobility modeling

To perform large-scale real-life experimental evaluations of pedestrian store-carry-forward networks is cumbersome and expensive, yet each experiment only captures a particular scenario and results are difficult to generalize. Therefore, it
is vital to mathematically model mobility. Our goal is to develop a library of analytical models that can be used to study the performance of pedestrian content distribution in some common case scenarios of urban mobility. The library would consist of simple models for abstract spaces: A “street model”, for example, would describe the rate and the duration of contact opportunities in one-dimensional topologies such as streets, sidewalks, or corridors. A “square model” would describe two-dimensional spaces, such as city squares, parks, parking lots, or airport halls. A “point model” would provide an abstraction of spots where people congregate: crosswalks, bus stops, pubs/cafés, or check-in terminals [31]. These simple models would then be used as building blocks to model larger areas.

In this paper, we describe the “street model” and how it can be used to build “grids of streets”. It allows us to study the feasibility of mobility-assisted content distribution in dynamic scenarios where nodes are in constant motion.

3.1. 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 \( \lambda \) and \( \lambda' \). 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 \( f(v) \) 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 contents. 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 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 \( \Delta \) the transmission range of a node. We distinguish between three types of contacts that the observer node makes: 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 passing the observer node. Since node arrivals to the street are Poisson, it is easy to show that, for 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 \( \mu_f, \mu_s \) and \( \mu_c \) 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.

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:

\[
\Pr\{N(t) = n\} = \Pr\{N_f(t) + N_s(t) + N_c(t) = n\}.
\]

Note that \( \Pr\{N(0) = 0\} < 1 \) because all nodes within the first \( \Delta \) 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.
Table 1
Mean arrival rates at the observer queue for fast, slow, and counter-directed nodes (← means “same as in the previous column”).

<table>
<thead>
<tr>
<th>μf(vo,t)</th>
<th>μs(vo,t)</th>
<th>μc(vo,t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 &lt; t ≤ 1/vo</td>
<td>1/vo-vmax(Vmax - vo)</td>
<td>1/vo-vmax(Vmax - vo)</td>
</tr>
<tr>
<td>1/vo ≤ t &lt; 1/vop</td>
<td>1/vo-vmax(Vmax - vo)(ln Vmax / vo + 1)</td>
<td>1/vo-vmax(Vmax - vo)(ln Vmax / vo - 1)</td>
</tr>
<tr>
<td>1/vop ≤ t &lt; 1/vo</td>
<td>1/vo-vmax(Vmax - vo + vo ln Vmax / vo)</td>
<td>1/vo-vmax(Vmax - vo - vo ln Vmax / vo)</td>
</tr>
</tbody>
</table>

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

\[ Pr[N_f(v_o, t) = i] = \sum_{j=i}^{\infty} Pr[N_f(v_o, t) = i \mid A_f(v_o, t) = j] Pr[A_f(v_o, t) = j] \]

(2)

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

\[ Pr[A_f(v_o, t) = j] = \frac{m_f(v_o, t)^j}{j!} e^{-m_f(v_o, t)} \]

(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:

\[ Pr[N_f(v_o, t) = i \mid A_f(v_o, t) = j] = \binom{j}{i} q_f(v_o, t)^i(1 - q_f(v_o, t))^{j-i} \]

(4)

where \( q_f(v_o, t) \) is the probability that a fast node, which has arrived at some time \( 0 \leq \tau < t \), 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:

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

(5)

Since the arrivals are Poisson, (5) becomes

\[ q_f(v_o, t) = \frac{1}{m_f(v_o, t)} \int_0^t Pr[s_f(v_o, v, \tau) > t - \tau \mid v > v_o] \times \mu_f(v_o, t) d\tau. \]

(6)

On an infinitely long street (the highway model [20]) 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 \( \tau = \Delta / |v - v_o| \).

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 \( \tau \) 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 \( f_V(v \mid V > v_o) \). In the case of the uniform speed distribution, it is given by:

\[ f_V(v \mid V > v_o) = \frac{1}{V_{\text{max}} - v_o} \left( 1 + \ln \frac{V_{\text{max}}}{v_o} \right). \]

(7)

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

\[ Pr[N_f(v_o, t) = i] = \frac{(m_f(v_o, t)q_f(v_o, t))^i}{i!} e^{-m_f(v_o, t)q_f(v_o, t)}. \]

(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 nodes \( N_s(v_o, y) \) and counter-directed nodes \( N_c(v_o, t) \) 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. Based on the properties of Poisson distribution, the total number of nodes in the observer queue \( N(v_o, t) = N_f(v_o, t) + N_s(v_o, t) + N_c(v_o, t) \) is also Poisson distributed.

\[ Pr[N(v_o, t) = n] = \frac{\rho(v_o, t)^n}{n!} e^{-\rho(v_o, t)}. \]

(9)
where \( \rho(v_o, t) = m_f(v_o, t)q_f(v_o, t) + m_q(v_o, t)q_q(v_o, t) + m_e(v_o, t)q_e(v_o, t) \). Note that, while \( N_f(t) \), \( N_q(t) \), and \( N_e(t) \) in (1) are dependent random variables, \( N_f(v_0, t) \), \( N_q(v_0, t) \), and \( N_e(v_0, t) \) are conditionally independent given that the speed of the observer node is \( v_0 \) and, therefore (9) holds. The probability that the observer node is connected (to at least one node) at time \( t \) is given by:

\[
\Pr[N(v_o, t) \neq 0] = 1 - e^{-\rho(v_o,t)}.
\]  

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

\[
\frac{1}{(V_{\text{max}} - V_{\text{min}})} \frac{1}{L} \int_{V_{\text{min}}}^{V_{\text{max}}} v_o \int_0^{L/v_o} (1 - e^{-\rho(v_o,t)}) \, dt \, dv_o.
\]  

For \( L \gg \Delta \), (11) simplifies to:

\[
1 - e^{-\frac{2\lambda(\lambda+\lambda^{'})}{V_{\text{max}}-V_{\text{min}}} \log \frac{V_{\text{max}}}{V_{\text{min}}} }.
\]  

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

\[
\frac{1}{(V_{\text{max}} - V_{\text{min}})} \frac{1}{L} \int_{V_{\text{min}}}^{V_{\text{max}}} v_o \int_0^{L/v_o} \mu(v_o, t) \, dt \, dv_o.
\]

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 obtained from a simulator.

4.1. Simulations

We use Legion Studio [32] to simulate the pedestrian mobility and evaluate the impact of the simplifying assumptions made in the analytical model regarding lack of interactions of nodes. Legion Studio is an advanced pedestrian mobility simulator used in urban and traffic planning to design public spaces such as airports, subway stations, and sport stadiums. It aims to faithfully capture speed–distance relations that emerge when people navigate around obstacles and other pedestrians. These pedestrian-to-pedestrian and pedestrian-to-obstacle interactions are responsible for the formation of pedestrian crowds—capturing them is of paramount importance when dimensioning emergency exits, stairs, and escalators. Legion Studio is based on detailed models that describe the mechanics of movements and algorithms that pedestrians use to select routes. It captures the effects of platooning behind slower pedestrians, queuing at bottlenecks, and personal space requirements. It has been calibrated based on measurements performed in different locations and among people with different cultural backgrounds [33]. An arbitrary arrival pattern of pedestrians into the studied area can be specified as an input to the simulator, along with the desired speed distribution. The actual speed of a pedestrian, however, depends on the interactions with other pedestrians. Therefore, simulations may result in contact rates and contact durations that are different from those obtained with the analytical model, which assumes free unobstructed flow of people. We introduce a warm-up period in each simulation run to ensure that the steady-state spatial distribution of nodes has been reached. We conduct five runs for each scenario, and in each run we collect statistics from at least 1000 nodes.

4.2. Single street segment

We present performance results from the analytical model and Legion simulations for a street of length \( L = 100 \) m, transmission range, \( \Delta = 10 \) m and symmetric Poisson arrival rates \( \lambda = \lambda' \). We also provide simulation results for other arrival processes with the same mean arrival rate \( \lambda \): hyper-exponential and Erlang-4.

Fig. 2 (left) shows the effect of the speed distribution on the average connectivity of nodes. We have selected two different distributions with the same mean, but with different variances: constant speed of 1.3 m/s and uniformly distributed speed in [0.6, 2.0] m/s. The mean is chosen based on the empirical study of typical walking speeds presented in [27]. We see that the connectivity of the nodes increases with the arrival rate and with the variance of the speed distribution. Note that the connectivity scales with the product \( (\lambda + \lambda') \Delta \) in (12), and therefore an increase of the transmission range would have the same effect on the connectivity as an increase in the arrival rate. 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 \( \lambda = \lambda' \). The results show a good match between the model and simulations. The duration of a contact is a random variable \( T \) whose tail distribution \( \bar{F}_T(t_{\text{min}}) = P(T > t_{\text{min}}) \) is plotted in Fig. 3 for the two speed distributions. Here \( t_{\text{min}} \) is the time
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 ($\lambda = \lambda'$), the length of the street is 100 m and the transmission range is 10 m. 95% confidence intervals are indicated for simulation results.

Fig. 3. Tail distribution of contact durations for Constant (1.3) and Uniform (0.6, 2.0) speed distributions.

required to transfer the smallest meaningful chunk of data, which includes the connection setup time—contacts shorter than $t_{\text{min}}$ are considered useless for content distribution. The shape of the tail distribution for constant speed is due to the fact that contacts in the forward direction last for the whole sojourn time in the street while most of the contacts in the opposite direction last for $\Delta/v$, when $v$ is the node speed.

The analytical model assumes Poisson arrivals. 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. In addition to the Poisson arrivals, the connectivity and contact rates were measured for two arrival processes with the same mean arrival rate $\lambda$: (i) two-phase hyper-exponential inter-arrivals with arrival rates $0.35\lambda$ and $5.7\lambda$ in the first, respectively, second phase, and selection probabilities 0.31 and 0.69 and (ii) Erlang-4 distributed inter-arrivals with the rate $4\lambda$ in each stage. The coefficients of variation for (i) and (ii) are 2 and 0.5, respectively. The hyper-exponential arrivals are “burstier” than the Poisson arrivals. Therefore, pedestrians often arrive in connected platoons, which is reflected in higher connectivity in Fig. 4 (left). To the contrary, the Erlang-4 arrivals are less bursty than Poisson arrivals and, therefore, result in lower connectivity. Arrivals and departures of pedestrians are bursty in many practical cases: at crosswalks, upon bus/train arrivals, at the ends of lectures/meetings, etc. In such cases, Poisson arrivals, which are uncorrelated, may be used as worse-case scenarios to estimate the achievable connectivities and contact rates. At higher arrival rates, the differences in connectivity become negligible and, therefore, independent of the arrival process, as shown in Fig. 4 (left).

In summary, from the results presented in this section we see that most of the connectivity and contact rate values obtained from the analytical model are within the 95% confidence intervals of the simulation results. The tail distributions for the contact duration also match closely. Therefore, we rely on the model to study the pedestrian content distribution in the following section. In addition to the street scenario, we have studied the impact of arrival processes and speed distributions on contact rates and durations for a number of other mobility scenarios in [34].
Fig. 4. Average connectivity (left) and average contact rate (right) for various arrival processes. The pedestrian traffic is symmetric ($\lambda = \lambda'$), the length of the street is 100 m and the transmission range is 10 m. All results shown are obtained from simulations.

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

5. Content distribution

We study the content distribution in a topology represented by a grid of streets, which is typical for urban areas. We start with the analysis of a single street segment—the analysis is later extended to a topology that represents a part of a downtown area in Stockholm.

5.1. Content distribution in a street segment

We consider a simple scenario where contents (e.g. a short text message) spread epidemically among nodes. We assume that every contact longer than $t_{\text{min}}$ results in successful spreading of the contents. Let $p_0$ and $q_0$ be the percentages of nodes that bring contents when they arrive to the near and far end-points 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 $p_0$ 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 ($p_0$, $q_0$) and outputs ($p_L$, $q_L$) of a street segment become outputs, respectively, inputs of neighboring segments. This allows us to study the content distributions in a grid of streets 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

\[
p(x + \Delta x) = p(x) + (1 - p(x))\theta(x),
\]

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 $\mu_f$, $\mu_s$, and $\mu_c$, respectively, the
probability of meeting at least one of the nodes with the content and establishing a contact that will last for at least $T > t_{\text{min}}$ seconds is given by:

$$\theta(v_o, x) = 1 - e^{-L/(v_o x/v_o) \bar{F}_T(t_{\text{min}}) p(x)} = e^{-L/(v_o x/v_o) \bar{F}_T(t_{\text{min}}) p(x) + \mu_s(v_o, x/v_o) \bar{F}_T(t_{\text{min}}) q(x)} \Delta x/v_o,$$

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

$$\theta(v_o, x) = \left(\mu_f(v_o, x/v_o) \bar{F}_T(t_{\text{min}}) p(x) + \mu_s(v_o, x/v_o) \bar{F}_T(t_{\text{min}}) p(x) + \mu_c(v_o, x/v_o) \bar{F}_T(t_{\text{min}}) q(x)\right) \frac{\Delta x}{v_o}.$$

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

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

where:

$$a(x) = \frac{1}{V_{\text{max}} - V_{\text{min}}} \int_{V_{\text{min}}}^{V_{\text{max}}} \frac{1}{v_o} \left(\mu_f(v_o, x/v_o) \bar{F}_T(t_{\text{min}}) + \mu_s(v_o, x/v_o) \bar{F}_T(t_{\text{min}})\right) dv_o$$

and

$$b(x) = \frac{1}{V_{\text{max}} - V_{\text{min}}} \int_{V_{\text{min}}}^{V_{\text{max}}} \frac{1}{v_o} \mu_c(v_o, x/v_o) \bar{F}_T(t_{\text{min}}) dv_o$$

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

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

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

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

where $c(x) = b(x)$ and $d(x) = a(x)$ if $L = \lambda'$. When $L \gg \Delta$, $\mu_f$, $\mu_s$, $\mu_c$ 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 $q(0) = q_1$. This assumption that $L \gg \Delta$ 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 $\Delta$ meters of the street.

The content dispersion in the forward direction $p(x)$ for different arrival rates $\lambda = \lambda'$ and minimum required contact durations $t_{\text{min}}$ is shown in Fig. 6. 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, $V \sim$ Uniform (0.6, 2.0) m/s, and $p_0 = q_1 = 5\%$. The content spreads more efficiently for higher arrival rates and lower $t_{\text{min}}$. The content distribution can be hindered by the requirements on the minimum contact duration (e.g., long connection setup time), as shown for $t_{\text{min}} = 15$ s. It is therefore important to minimize the setup time to utilize short contacts. The results indicate that the model provides a fairly good estimate of the content dispersion.

In Fig. 6 (left), we assume that all nodes in the street participate in the contents distribution and we focus on low arrival rates of the nodes, which are critical for the efficiency of the content distribution. At low arrival rates, node interactions...
are negligible—pedestrians move at their desired walking speed, unobstructed by other pedestrians. This agrees with the assumption made in the analytical model. Therefore, it is not surprising that the model is a close match to the simulations. In reality, however, only some pedestrians would distribute the content, others might not be interested in the content, or may not carry any communication devices at all. We refer to them as “bywalkers” and denote their arrival rates to the near and far ends of the street by $\lambda_{bw}$ and $\lambda'_{bw}$, respectively. The total arrival rates into the street $\lambda + \lambda_{bw}$ and $\lambda' + \lambda'_{bw}$ might be such that pedestrian interactions cannot be neglected. In such cases, platooning might significantly affect the walking speeds and, therefore, the rate and the duration of contact opportunities.

The content dispersions for three different arrival rates of the bywalkers $\lambda_{bw} = \lambda'_{bw} (0.1, 0.2, \text{and } 0.3 \text{ s}^{-1}$), $\lambda = \lambda' = 0.05 \text{ s}^{-1}$, and $t_{min} = 10$ s are shown in Fig. 6 (right). The results are obtained by simulations in Legion. The dispersion without bywalkers is plotted in the same figure for comparison. The results suggest that, in dense scenarios, such as busy pedestrian streets and crowded corridors, pedestrian interactions have a significant impact on the content dispersion: The contents spread more efficiently because platooning results in longer contacts. The analytical model may provide a worst-case estimate of the achievable performance in such scenarios.

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 part of Stockholm’s downtown area 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 $p_{in}$ percent of the nodes that arrive to the first street (hence, the content arrival rate is $\lambda p_{in}$); it is marked in Fig. 7 (right). The source of the contents 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. The dispersion is calculated recursively: First, the node arrival intensities at the end-points of street segments are calculated based on routing probabilities in the grid. Then, based on the initial content distribution $p_{in}$, the spatial distributions of the content in the segments $p_i(x)$ and $q_i(x)$ are calculated recursively until they converge to steady-state distributions, which determine the dispersion of the content in the grid.

To illustrate the spreading of the content, we assume the following parameters in our model: uniformly distributed nodal speeds in [0.6, 2.0] m/s, $\Delta = 10$ m, and $p_{in} = 5\%$. Beside the variance of the speed distribution, the minimum useful 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. 8 (left). The maximum standard deviation in the simulation results is 0.015 for $t_{min} = 10$ s and $\lambda = 0.01 \text{ s}^{-1}$; it is significantly lower in all other scenarios. The discrepancy between the analytical and simulation results is mainly due to the modeling assumption that contacts break at street end-points, which results in a lower number of useful contacts. The tail distribution of the contact duration $F_T(t_{min})$ affects the dispersion through the coefficients of differential equations (18) and (19), and it is a function of $\Delta/t_{min}$. Therefore, the transmission range $\Delta$ and the minimum useful contact duration $t_{min}$ should be optimized to improve the performance. This optimization must address issues such as the energy consumption in mobile terminals, radio interference, time-efficiency of neighbor discovery, and segmentation of the contents into atomic...
Fig. 8. Effect of $t_{\text{min}}$ on content dispersion for various arrival rates and $p_{\text{in}} = 5\%$ (left) and content dispersion for various percentages $p_{\text{in}}$ of nodes that arrive with the content and $t_{\text{min}} = 20$ s (right) ($V \sim \text{Uniform}(0.6, 2.0)$ m/s and $\Delta = 10$ m).

Fig. 9. Modified network of street segments with a square in center of the area.

units (chunks) of optimal size. A study in [35] evaluates the impact of beaconing interval on the energy consumption and discovery delay.

In Fig. 8 (right) we show that the percentage of nodes that arrive with the contents $p_{\text{in}}$ does not have a significant impact on the content dispersion. Therefore, new contents can be seeded by a small number of nodes and still spread with a high efficiency. This is important for the distribution of user-generated contents. The result also suggests that the fixed infrastructure, which provides contents not generated by nodes themselves, can be rather sparse.

5.3. Content distribution in a modified topology

The analytical model presented in this paper is restricted to topologies composed of street segments. However, urban areas where pedestrians move often contain congregation points (bus stops, crosswalks, places of interest) and public spaces (squares, parks) where people pause or move differently from what is assumed in the model. In this section, we study the impact of topological changes on the efficiency of the content distribution on an example. We consider a topology obtained by converting the central part of the area shown in Fig. 7 to a public square, as shown in Fig. 9. Pedestrians arrive to the square from six adjoining streets. Upon arriving to the edge of the square, a pedestrian chooses one of 20 focal points that are uniformly distributed over the square area. The time that the node spends in the chosen focal point is uniformly distributed in $[0, T_{\text{max}}]$, where $T_{\text{max}}$ is selected so that the average sojourn time in the modified topology is the same as in the original topology. It turns out that $T_{\text{max}} = 60$ s results in the average sojourn time of 343 s, which is almost identical to the sojourn time of 344 s observed in the original topology. After the pause, the pedestrian randomly chooses one of the six adjoining streets and leaves the square. In Fig. 10, we compare content dispersions in the original and in the modified topologies for
Fig. 10. 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 and $t_{min} = 20 \text{s}$ (right) ($V \sim \text{Uniform}(0.6, 2.0) \text{m/s}$ and $\Delta = 10 \text{m}$).

various $\lambda$, $t_{\text{min}}$, and $p_{\text{in}}$. The results show that the impact of the modifications on the content distribution is negligible. This indicates that, in some cases, details of a topology can be abstracted away. Some parts of it can be replaced by structures that are easier to describe analytically.

6. Conclusion

We have considered infrastructureless mobility-assisted content dissemination to an arbitrarily large group of pedestrian nodes with short-range radio connectivity. We developed an analytical model to study the connectivity properties of pedestrian mobility in a street segment. The model is then used as a building block to model larger areas. The results obtained from the model 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. Our results indicate that the system and mobility parameters that directly affect the tail distribution of contact durations, such as the connection setup time and speed variance, have large impact on the content dispersion, while some details of the topology can be abstracted away. The analytical results are compared to simulation results with good agreement in scenarios that assume free flow of people. We will continue to develop the model to describe the content dissemination in open-space areas and spots where people congregate or swarm.

Acknowledgment

The authors wish to thank Jing Jin for performing simulations in Legion.

References


