ARTICLE IN PRESS

Ad Hoc Networks xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Ad Hoc Networks

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

Impact of human mobility on wireless ad hoc networking in entertainment parks

Vladimir Vukadinovic *, Fabian Dreier, Stefan Mangold

Disney Research Zurich, Clausiusstrasse 49, CH-8092 Zurich, Switzerland

ARTICLE INFO

Article history:
Received 9 August 2011
Received in revised form 20 May 2012
Accepted 30 May 2012
Available online xxxx

Keywords:
Pedestrian mobility
GPS traces
Opportunistic communication
Delay tolerant networking
Store-carry-forward networks

ABSTRACT

Ad hoc networks of wireless devices carried by entertainment park visitors can support a variety of services. In such networks, communication links between the devices sporadically appear and disappear with the mobility of visitors. The network performance strongly depends on how often they encounter each other and for how long the contact opportunities last. In this paper, we study the mobility of visitors based on GPS traces collected in two entertainment parks. We demonstrate and discuss the implications of the observed mobility on the efficiency of opportunistic data forwarding. We show how hourly changes in the number and spatial distribution of the park visitors affect the delay of a broadcast application. Our results suggest that generic mobility models commonly used in wireless research are not appropriate to study this and similar scenarios: Targeted mobility models are needed in order to realistically capture non-stationarity of the number and spatial distribution of nodes. Therefore, we developed a mobility simulator for entertainment parks that can be used to scale up the evaluation scenarios to a large number of devices. The simulator implements an activity-based mobility model, where the mobility of park visitors is driven by the activities they wish to perform in the park. The simulator is calibrated based on the GPS traces and validated on several metrics that are relevant for the performance of wireless ad hoc networks.

1. Introduction

For many wireless services, a continuous connectivity and end-to-end paths are not necessary. Unlike in infrastructure-based networks that provide full wireless coverage, in the so-called ad hoc networks wireless devices communicate directly when within each other's range. This communication mode is useful when infrastructure-based communication is costly or unavailable. When devices are mobile (e.g. carried by people), the ad hoc communication may experience occasional disruptions as links between devices appear and disappear with changes in the distance between the devices. Network applications and protocols need to be delay and disruption tolerant to benefit from such intermittent connectivity [1,2]. This requires re-design of many protocols, especially routing protocols, since an end-to-end path between devices is not necessarily available throughout a communication session [3]. Therefore, messages are forwarded incrementally through the network in a store-carry-forward manner when contact opportunities arise. Understanding the mobility of people is crucial because mobility determines the rate and the duration of contact opportunities. Human mobility, however, is not easy to characterize. For example, working-day, shopping, and campus mobility will all result in different encounter patterns. For many practical applications, routing/forwarding algorithms must target specific mobility scenarios, even if this limits the scope of their applicability.

Future services offered to entertainment park visitors might rely on wireless technologies, devices, and applications [20]. Personalized location-based services, mobile

* Corresponding author. Tel.: +41 44 632 9491.
E-mail addresses: vvuk@disneyresearch.com (V. Vukadinovic), fdreier@disneyresearch.com (F. Dreier), stefan@disneyresearch.com (S. Mangold).

1570-8705/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.adhoc.2012.05.011

trans-reality games, park information services, social networking, and multimedia sharing would make a visit to a park more interactive and entertaining. However, it cannot be assumed that wireless data service is available throughout the park. Cellular 3G coverage is typically available, but many park visitors are foreign tourists who do not have data plans with local operators and are not willing to pay data rooming charges. Rolling out extensive Wi–Fi infrastructure to support wireless services in an entertainment park is not an easy task: The largest parks are comparable in size with big cities (e.g. the Walt Disney World Resort in Florida spans over $\sim$100 km^2, an area as large as San Francisco). Although Wi–Fi infrastructure and wireless services would provide some added value to the visitors, it is often not clear how would they increase the revenue (e.g. attendance) of theme parks. Besides, since theme parks usually offer very unique experiences, there is no push from competition to provide such services. Due to the lack of strong reasons that would justify costly and logistically complex full-scale infrastructure deployment, alternative solutions that enable gradual introduction and testing of wireless services at low-cost are preferred. Such lightweight solutions would help park management assess the needs and tech requirements for possible future deployments. For some theme park applications continuous connectivity provided by fixed wireless infrastructures is not needed: Spotty coverage might be tolerated if supported by opportunistic store-carry-forward type of communication among visitors. Examples include distribution of park information (waiting times at different attractions, schedules of street parades and other performances), mobile advertising, collaborative localization, participatory sensing, polling/surveying, and multimedia sharing. Some of the application scenarios are described in Section 2. The applications may run on smart phones brought by visitors, or on customized devices handed out to the visitors. The latter could be optimized for opportunistic communication and park-specific scenarios.

In this paper, we study the mobility of park visitors based on GPS traces collected in two entertainment parks in order to understand network requirements for opportunistic communication (minimum number and density of mobile devices and supporting infrastructure nodes). On an example of epidemic broadcasting, we analyze the impact of hourly changes in visitors' mobility and density on the speed of content dissemination [41]. Contact-related statistics, such as inter-any-contact time and mean square displacement, are extracted from the traces and their impact on the broadcasting performance is discussed. The number of traces in our dataset, even though larger than in most datasets used in related studies, is not sufficient for large-scale evaluation. Therefore, mobility models that can produce realistic node encounter patterns are needed. Simplistic and rather generic mobility models, which are often used in wireless research, assume constant number and a stationary, steady-state spatial distribution of nodes in an area. Targeted mobility models are needed in order to realistically capture non-stationarity of the number and spatial distribution of nodes. We present an activity-driven mobility model of park visitors, which we implemented in our ParkSim simulator [42]. The model is calibrated based on the GPS traces and other data obtained from the entertainment parks. The outputs of the simulator are synthetic mobility traces of park visitors, which can be used for trace-driven simulations of mobile ad hoc networks.

The remainder of this paper is organized as follows: Some examples of the entertainment park applications that may benefit from opportunistic communication are given in Section 2. GPS traces are described in Section 3. The performance of opportunistic broadcasting is studied in Section 4. Contact-related statistics are analyzed in Section 5. A mobility model derived based on the GPS traces is described in Section 6. Section 7 concludes the paper.

2. Application scenarios

Some of the application scenarios for opportunistic networking in entertainment parks are described in the following.

2.1. Mobile trans-reality games

Mobile trans-reality games often rely on wireless technologies. Some of them can be supported with a gossip-based communication among players. A simple example is Insectopia [4], a game where players with mobile phones roam Bluetooth-rich environments searching for and catching a multitude of different "insects". Insect types are represented by unique Bluetooth signatures of the devices. In scavenger hunt games, team members often exchange information needed to complete their mission. Kim Possible [5] is a Disney game played in the Epcot park where players take roles of secret agents equipped with communication devices. Some games however do not revolve around technology and dedicated communication devices (i.e. mobile phones). In those games, gossip-based protocols can be used in real-world to mimic the way game characters (e.g. toys) would communicate with each other in a fantasy-world. For example, in a game designed for young children, a task could be to guide a toy character through missions during which the radio-enabled toy is empowered (e.g. with skills and knowledge) through contacts with other toys and objects in the park.

2.2. Mobile advertising

Mobile advertising can be used in entertainment parks to inform visitors about special events (e.g. shows, street performances and fireworks) and shopping/dining opportunities. Advertisements may take the form of electronic tips and discount coupons that are distributed wirelessly from infrastructure nodes and forwarded epidemically from a device to a device. The advertisements may target the entire park population (flooding), or a sub-population based on visitor's personal profile (multicasting) or current location (geocasting). Long waiting times at popular rides, which are common during summer vacations and holiday weekends, are undesirable. Opportunistic communication can be used to inform visitors about waiting times at different rides so that they can organize their visit time in a best possible way. A network-enabled queue management application would allow visitors to request and obtain an electronic token for a ride on their mobile phones. The token would allow them to enter the ride at a prescribed time of the day without waiting.

2.3. Collaborative localization

Many of the entertainment park applications require knowledge of guests' current location in the park. For example, in mobile games, the game engine often relies on the knowledge of players' positions to control the way in which the game unfolds (e.g. location-specific instructions/clues are sent to players' devices). Geocasting of messages/advertisements and other types of location-based services also require mechanisms to localize visitors. Localization solutions can be power demanding. For example, frequent sampling of a GPS receiver would quickly drain the battery on most mobile phones. Besides, not necessarily all devices carried by visitors have the same localization capabilities. Allowing neighboring devices to share their location information via short range radio contacts would help reduce the energy cost and allow less capable devices to localize themselves more accurately. We refer to this type of localization as cooperative/collaborative localization [43]. It relies on opportunistic broadcasting of the location information and smart data fusion algorithms to refine location estimates based on neighbors' locations obtained through the broadcasts.

3. GPS traces

Lack of large-scale measurements of human mobility is a big challenge for wireless research communities. It is difficult to organize large-scale measurement campaigns because of financial costs, logistical hurdles, privacy concerns, and government regulations [6]. Most previous studies of human mobility/encounter patterns for opportunistic communication rely on datasets that are limited in terms of the number of devices and/or time duration. Our dataset (1800 GPS traces in total, out of which 1647 are used in the analysis) is significantly larger than most datasets used in similar studies. For example, Bluetooth datasets used in [7–9], which contain records of discovered peers, are obtained in experiments with at most 100 devices. GPS datasets used in [10–13] contain up to 200 mobility traces (one of the dataset in [11] contains 15 GPS traces collected in The Walt Disney World Resort in Florida). Wi–Fi datasets used in [9,14,15], which contain SSIDs of access points visible by Wi–Fi devices, are much larger (up to several thousand laptops and PDAs). However, it is difficult to infer contact from such datasets. Typically, two Wi–Fi devices are assumed to be in contact as long as they see the same access point. This is a vague indicator that they may actually be able to connect to each other using short-range radios. Furthermore, some of the Wi–Fi devices were not carried by their owners at all times (e.g. laptops). Hence, observed contacts do not necessarily characterize human mobility.

Fig. 1. The Epcot park consists of two major sections, Future World and World Showcase. Map data © OpenStreetMap contributors, CC BY-SA.

Our GPS traces were collected during a research study carried out in the Epcot Park, Florida and Disneyland Paris (DLP), France. The layout of the Epcot park is shown in Fig. 1. The park covers an area of $\sim$1 km^2 and receives close to 10 million visitors per years ($\sim$28,000 per day on average, significantly more on weekends and holidays). It consists of two sections, Future World and World Showcase, with approximately 20 themed sub-areas/attractions. The Future World, which is closer to the park entrance, is more popular of the two. Often visitors need to wait in queues to enter attractions located in this section. The World Showcase is centered around a lake. A number of restaurants and stores are located throughout the park. The layout of DLP is shown in Fig. 2. The complex covers an area of approximately the same size as Epcot ($\sim$1 km^2) and receives close to 15 million visitors per years ($\sim$40,000 per day on average, up to 70,000 on a busy day). It consists of two parks, Disneyland Park and Walt Disney Studios Park, the former being the larger and more popular of the two. Vast majority of visitors opt for a ticket that gives them access to both parks, but do not necessary visit both on the same day. As in the Epcot park, there are several

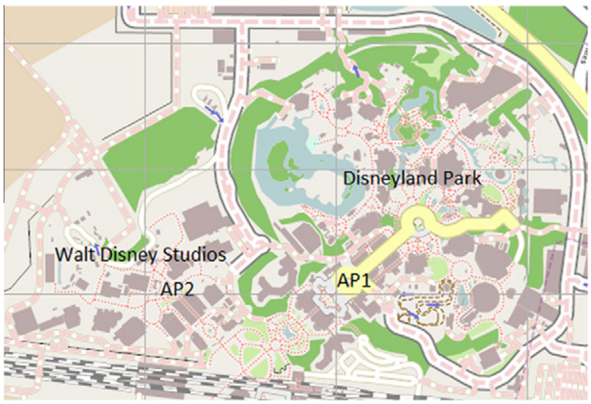
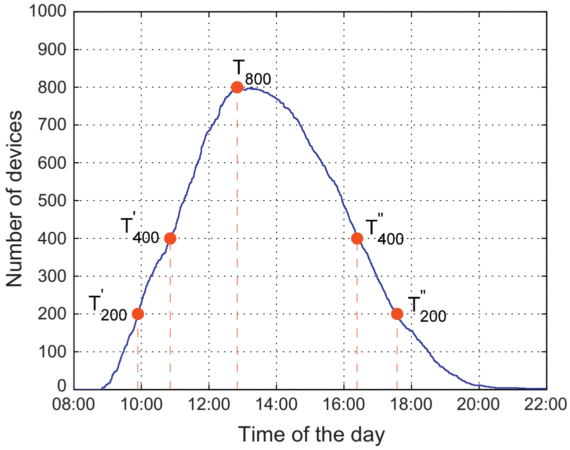


Fig. 2. The DLP consists of two parks, Disneyland Park and Walt Disney Studios Park. Map data © OpenStreetMap contributors, CC BY-SA.

Please cite this article in press as: V. Vukadinovic et al., Impact of human mobility on wireless ad hoc networking in entertainment parks, Ad Hoc Netw. (2012), http://dx.doi.org/10.1016/j.adhoc.2012.05.011



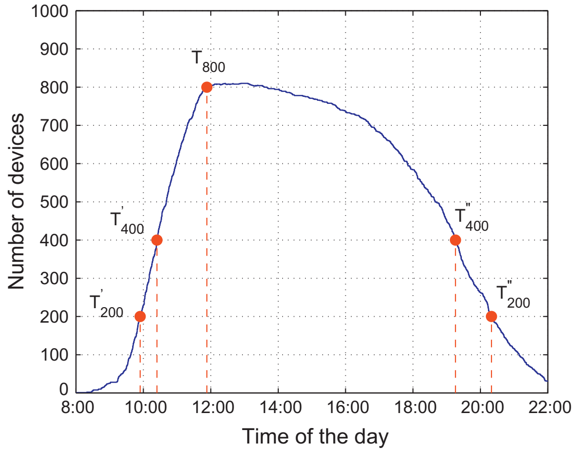


Fig. 3. Number of phones in Epcot (left) and DLP (right) at different times of the day. Times when the number reaches 200, 400, and 800 are indicated. Visits to Epcot are typically shorter than visits to DLP since Epcot is a part of larger park complex: visitors usually allocate only part of the day to Epcot.

attractions in DLP where long waiting queues are not unusual. There are approximately 40 restaurants and one hotel within the complex.

Over the course of 10 days (5 in Epcot and 5 in DLP), close to 200 smartphones were distributed each day to a total of 1800 randomly selected visitors (910 in Epcot and 890 in DLP). In case of groups/families only one of the members was selected. The phones were distributed at the entrance between 8 am and 1 pm, and collected when the visitors were exiting the parks. The phones ran an application that logged their GPS locations on average every 2 min when the satellite signals were available. In our study, we ignore the dates of the logs, as if all GPS traces were collected on the same day. This is needed to study networks where the number of devices is larger than the number of phones that were available for the experiment. The number of phones in the parks at different times of the (merged) day is shown in Fig. 3. Note that traces collected on days with high and low park attendance (e.g. parks are significantly more crowded on weekends than on business days) cannot be merged because mobility is affected by the crowdedness of the park. The traces that we merged were collected Monday through Friday. We checked if the waiting times at popular attractions were similar on the five days of experiments. For each attraction, we extracted the average visit times, which is the sum of the waiting time and ride time, from GPS traces. Then we calculated the waiting times on different days by subtracting the ride times from the visit times. The average waiting times for the Test Track in Epcot and the Rock'N Roller Coaster in Disney Resort Paris are shown in Table 1.

Table 1
Average waiting times (in seconds) for two popular attractions.

	Test track	Rock'N roller coaster
Monday	955	613
Tuesday	779	377
Wednesday	888	739
Thursday	1032	660
Friday	1217	856

Although the waiting times vary, the table does not indicate that parks were substantially more crowded on some days of the experiment than on the others (e.g. it is not uncommon to have close to zero waiting on business days with bad weather and more than one hour waiting on holidays).

In addition to geo-coordinates, GPS accuracy was also logged during the trace collection. We discarded waypoints whose accuracy was worse than 25 m. We also discarded traces shorter than 2 h or containing less than 50 waypoints. Results presented in the following sections are based on the remaining 825 out of 910 Epcot traces and 822 out of 890 DLP traces. We interpolated the movements of visitors between the remaining waypoints assuming straight-line movements. The traces may contain gaps, which correspond to the periods when visitors were indoors (e.g. in a building where GPS signal is not available). During such periods, we assume that visitors move slowly from the spot where the last waypoint was recorded before they entered the building to the spot where the first waypoint was recorded after they exited the building.

4. Opportunistic broadcasting

Some of the entertainment park applications are broadcast in nature. For example, information about waiting times at different attractions/rides can be broadcasted to the visitors opportunistically. Since waiting times change slowly, this information is not time critical and delivery delays of up to a few tens of minutes can be tolerated. For some other services/information, shorter delivery delays might be required. Here we investigate how mobility and density of devices affects the speed of opportunistic broadcasting based on the GPS traces. We evaluate the time needed to distribute a message to a certain target percentage of park visitors (e.g. 98% is a tentative target for one of the applications described in Section 2). We consider both Epcot and DLP scenarios.

The scenario setup for the Epcot park is as follows: A single infrastructure node (e.g. info-station, access point), labeled as AP1 in Fig. 1, is located in the center of the Future World section of the park (later we consider adding AP2). This is one of the spots with the highest flux of visitors: almost all visitors pass by this spot when entering and leaving the park. The transmission range of the access point is assumed to be 50 m. At time T, the access point starts broadcasting a message to the visitors within the range. The message spreads epidemically among visitors as they encounter each other. Radio aspects (attenuation, interference, energy consumption) and protocol details (device and content discovery, connection setup delay, content caching) are ignored. The only assumption is that the transmission range of mobile devices is 10 m, unless stated otherwise. When a device without the message enters the transmission range of the access point or of another mobile device that possesses the message, it obtains the message instantaneously. The purpose of this simple scenario is to estimate the lower bound on the broadcast dissemination delay for observed mobility and density of the devices, irrespective of wireless technology constraints. This delay might be hard to achieve in practical systems. However, it provides an indication of how delay-tolerant an application should be to benefit from opportunistic communication and what number of devices is needed to meet certain delay constraints. A similar setup has been evaluated in [16] using a much smaller set of mobility traces collected in an office building in a university campus.

As described in Section 3, we merge five days of experiments in the Epcot park into a single day by ignoring dates in the GPS traces. The number of visitors with the phones at different times of the day is shown in Fig. 3 (left). The curve closely reflects the way in which the number of visitors in the park changes during a typical day. We assume that, if proprietary devices, such as electronic park guides, would be handed/rented out to the visitors, their number would follow a similar pattern. We assume that broadcasts are initiated at times when there are 200, 400, and 800 devices in the park. There were two moments when the number of devices reached 200, one in the morning and one in the afternoon, denoted by T'_{200} and T''_{200}, respectively. Similarly, there were 400 devices in the park at T'_{400} and T''_{400}. At the peak of the day, denoted by T_{800}, the number of devices reached 800. The spatial distributions of devices at those moments are shown in Fig. 4 (top). In the morning hours (T'_{200}, T'_{400}), there is an intensive inflow of visitors into the park, who tend to cram in the Future World section close to the entrance and the access point AP1. It takes several hours until visitors disperse. In the late afternoon (T''_{400}, and T''_{200}), there is an outflow of visitors. Hence, mobility patterns, which affect the efficiency of opportunistic broadcasting, depend strongly on the time of the day. The speed of content dissemination will be different at T'_{200} and T''_{200}, although the number of devices is the same. In the late afternoon, message is disseminated against the flow of crowd—visitors who obtain the content from the access point are likely to take the content out of the park very soon.

The scenario setup for the DLP is similar. We first consider a network supported by a single access point, labeled as AP1 in Fig. 2, which is located close to the entrance to the Disneyland Park. We then consider adding AP2 close to the entrance to the Walt Disney Studios. Message broadcasts are initiated at times when there are 200, 400, and 800 devices in the DLP; the times are indicated in Fig. 3 (right). It can be noticed from the widths of the bell-shaped curves in Fig. 3 that visits to DLP are on average longer than visits to Epcot, which is a part of larger park complex and

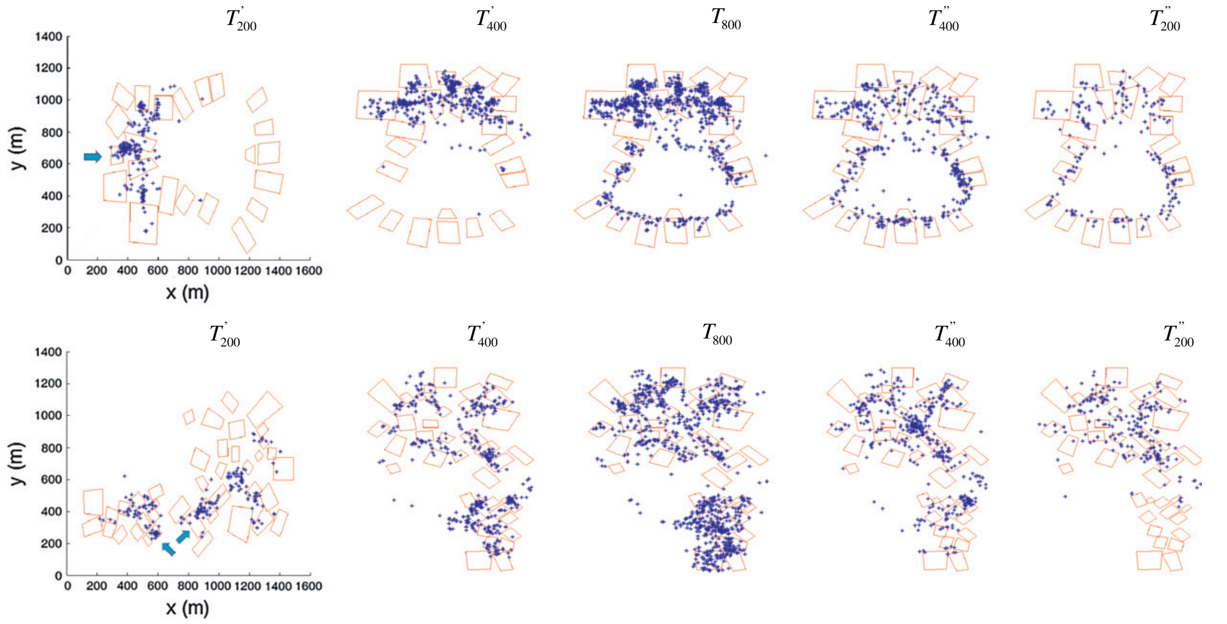


Fig. 4. Spatial distribution of visitors with the devices at T'_{200}, T'_{400}, T_{800}, T''_{400}, and T''_{200} (from left to right) in Epcot (top) and DRP (bottom). Park entrances are indicated in the first row of figures.

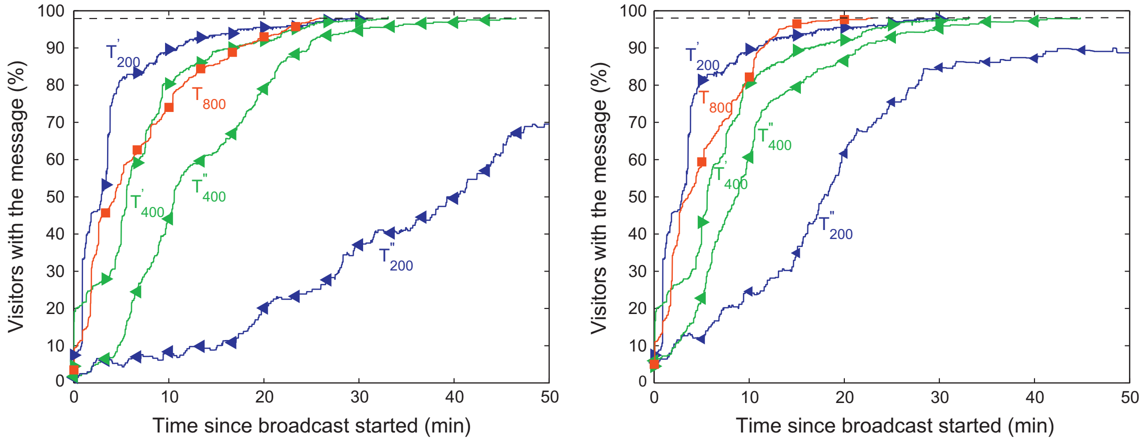


Fig. 5. Epcot scenario: Percentage of visitors with the message as a function of time since a broadcast started. Broadcasts start at T'_{200}, T'_{400}, T_{800}, T''_{400}, and T''_{200}. Left: With one access point (AP1). Right: With two access points (AP1 and AP2). The curves do not increase monotonically since visitors are entering and exiting the park.

visitors usually allocate only part of the day to it. The spatial distributions of devices at the times when the broadcasts are initiated are shown in Fig. 4 (bottom). Visitors tend to distribute evenly between the Disneyland Park and Walt Disney Studios, except in the evening. The latter closes at 19:00, when visitors either move to the Disneyland Park, or leave the DLP.

4.1. Performance results: Epcot scenario

The broadcast performance results for the Epcot park scenario with the single AP are shown in Fig. 5 (left). The curves in the figure show the percentage of visitors in the park that possess the message as a function of time elapsed since the start of the broadcast (at T'_{200}, T'_{400}, T_{800}, T''_{400}, or T''_{200}). We stop the simulation when the possession reaches 98% and we record the elapsed time ($\Delta T_{98\%}$) in Table 2. In the table, we also show the percentage of visitors that would have received the message by $\Delta T_{98\%}$ if peer-to-peer (P2P) forwarding was not used (hence, connecting to the access point was the only way to obtain the message). We report the following observations:

The time needed to broadly disseminate the message is in the order of tens of minutes. For example, it took 20 and 26 min, respectively, to deliver the message to 95% and 98% of the devices at T_{800}. At a typical peak hour there are 10,000–15,000 visitors in the Epcot. Therefore, the scenario with 800 devices assumes that 5–8% of the visitors have the devices and run the application, which is a significant number (especially considering that a big proportion of visitors are toddlers). When the range of the devices increases from 10 m to 20 m (at the expense of increased energy consumption), the time to deliver the message to 98% of devices decreases to 15 min, which is may be prohibitively long for some applications. To further reduce the dissemination time, more access points are needed and/or the number of devices should be larger.

Apart from the number of devices, dissemination time depends strongly on the spatial distribution of the devices and their residual times in the park after they receive the message. This is obvious when comparing the results for broadcasts initiated at T'_{200} and T''_{200} (former happens in the morning hours while the latter is in the late afternoon). At T'_{200} almost all visitors are located in the Future World section of the park. A share of 98% of them obtained the message within 30 min. Contacts with the access point accounted for 53% of delivered messages. To the contrary, at T''_{200}, visitors are spread throughout the park. It took 72 min to achieve 98% possession. Contacts with the access point accounted for only 13% of delivered messages. This illustrates the variety of performances that could be expected with the same number of devices, but at different times of the day. Furthermore, results in Fig. 5 (left) show that the message disseminates faster at T'_{200} than at T'_{400} and T_{800}. Hence, a larger number of devices does not guarantee

Table 2
Epcot scenario: Time $\Delta T_{98\%}$ needed to distribute the message to 98% of the devices and the percentage of devices that would receive the message by $\Delta T_{98\%}$ without P2P forwarding.

T	AP1			AP1 & AP2		
	$\Delta T_{98\%}$ (s)	AP + P2P (%)	AP only (%)	$\Delta T_{98\%}$ (s)	AP + P2P (%)	AP only (%)
T'_{200}	1855	98	53	1855	98	53
T'_{400}	1984	98	33	1984	98	33
T_{800}	1556	98	16	1380	98	20
T''_{400}	2790	98	21	2690	98	39
T''_{200}	4320	98	13	4240	98	36

better performance due to changes in spatial distribution and residual visit times.

We next study the effect of adding the second infrastructure node (AP2 in Fig. 1) on the speed of dissemination. Placing the node in the World Showcase may help reduce the broadcast delay in the afternoon hours, when many visitors are located in this section of the park, as shown in Fig. 4 (top). As expected, the results in Fig. 5 (right) and in the last three columns of Table 2 show that AP2 does not contribute to the message spreading at T'_{200} and T'_{400}. At T_{800} and T''_{400}, the message disseminates somewhat faster compared to the previous setup, as indicate by the slope of the curves in Fig. 5 (left) and Fig. 5 (right). A significant speed-up is achieved at T''_{200} when visitors from the back of the park, where AP2 is located, start to spread the message as they move across the park towards the exit. However, the addition of AP2 has very little effect on the time needed to reach 98% of devices, regardless of the time of the day. It is hard to deliver the message to the last few percent of visitors since they may be isolated from the rest (e.g. sitting on a boat in the middle of the lake). Another reason is constant inflow/outflow of visitors to the park. We evaluated the effect of adding two more APs at the border between the Future World and the World Showcase sections of the park. The additional APs helped speed up the dissemination, but the "last few percent" problem remained. An alternative to increasing the infrastructure coverage (either by adding more APs or by increasing their range) is to enable mobile devices to adapt their range according to the rate of encounters, current dissemination level, and remaining battery power.

4.2. Performance results: DLP scenario

The broadcast performance results for the DLP scenario with the single AP are shown in Fig. 6 (left) and in the first three rows of Table 3. In the morning hours (T'_{200}, T'_{400}, and T_{800}), message dissemination is initially slower than in the Epcot, as indicated by the slope of the curves. The contribution of the AP1 is also smaller (30% in DLP vs. 53% in Epcot at T'_{200}). The reason is that AP1 is in the Disneyland Park, while many people visit the Walt Disney Studios first and do not pass anywhere close to AP1. The time to disseminate the message to 98% of the visitors ($\Delta T_{98\%}$ in Table 3) is however shorter than in the Epcot. Visitors in the DLP are more mobile and, therefore, the problem of delivering the message to the last few percent of visitors is not as apparent as in the Epcot. For example, it took 540 s to reach 80% of visitors in the Epcot at T'_{400}, which is significantly faster than 940 s in the DLP. However, it took another 1424 s to achieve the 98% target in the Epcot, while only 430 s in the DLP. This illustrates how differences in spatial distributions and mobility levels affect the performance in the two scenarios. In the morning hours, Epcot visitors are clustered close to the AP1 and not particularly mobile: The message dissemination is achieved mostly

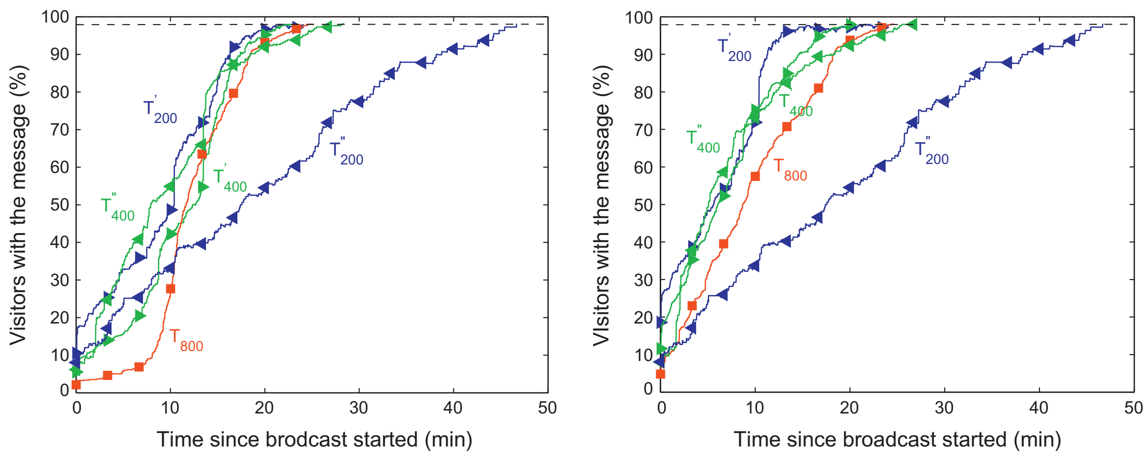


Fig. 6. DLP scenario: Percentage of visitors with the message as a function of time since a broadcast started. Broadcasts start at T'_{200}, T'_{400}, T_{800}, T''_{400}, and T''_{200}. Left: With one access point (AP1). Right: With two access points (AP1 and AP2).

Table 3
DLP scenario: Time $\Delta T_{98\%}$ needed to distribute the message to 98% of the devices and the percentage of devices that would receive the message by $\Delta T_{98\%}$ without P2P forwarding.

T	AP1			AP1 & AP2		
	$\Delta T_{98\%}$ (s)	AP + P2P (%)	AP only (%)	$\Delta T_{98\%}$ (s)	AP + P2P (%)	AP only (%)
T'_{200}	1445	98	30	1435	98	49
T'_{400}	1375	98	21	1205	98	36
T_{800}	1510	98	8	1475	98	19
T''_{400}	1700	98	15	1618	98	18
T''_{200}	2800	98	31	2800	98	31

through the infrastructure. At the same time, DLP visitors are dispersed and highly mobile: The message dissemination is achieved mostly through the mobility. In the evening at T_{800} all DLP visitors are in the Disneyland Park since Walt Disney Studios are closed at that time. The message dissemination is significantly slower than in the morning hours since the message disseminates against the flow of people who are leaving the park.

We also studied the effect of adding the second infrastructure node in the Walt Disney Studios (AP2 in Fig. 2). AP2 helped increase the dissemination speed in morning hours, since visitors entering the Walt Disney Studios are likely to pass next to it. This is evident from the slope of the curves in Fig. 6 (left) and Fig. 6 (right). However, as in the Epcot scenario, the time to reach 98% of visitors was not affected much by the additional AP since last of the visitors are reached through peer-to-peer encounters rather than infrastructure. In the evening hours, AP2 has no effect on the delay since all visitors are in the Disneyland Park.

5. Encounter statistics

In Sections 5.1 and 5.2, we analyze several contact-related statistics that are relevant for the performance of broadcast message dissemination (inter-any-contact time, mean square displacement, number of neighbors, and rate of new contacts) and correlate them with the results of the previous section. Additional statistics (inter-contact time and contact duration) are analyzed in Section 5.3: These statistics are not directly related to the performance of the application studied in the previous section, but they might be highly relevant for other examples of opportunistic networking in theme parks.

5.1. Encounter statistics: Epcot traces

Inter-any-contact time (IACT) is the time elapsed between starts of two successive contacts of a device with other devices. IACT determines the frequency of contact opportunities and, therefore, it affects the speed of opportunistic broadcasting. It strongly depends on the device density (i.e. time of the day). We observed IACTs in 30-min intervals following T'_{200}, T'_{400}, T_{800}, T''_{400}, and T''_{200}. Their complementary cumulative distribution functions (CCDFs) are shown in Fig. 7 (left). The curve labeled as T'_{200} represents CCDF of IACTs observed in $[T'_{200}, T'_{200} + 30$ min], for example. With distribution fitting, we found that the distribution of IACTs is accurately described by the gamma distribution with the shape parameter between 0.6 and 0.7 depending on the time of the day (95% confidence intervals for the MLE of the shape parameter were within ± 7% of the MLE at T'_{200} and T''_{200} and within ± 1.5% at T_{800}). This is consistent with the results presented in [17], but not with the power-law distribution observed in [18]. The Bluetooth sighting traces analyzed in [18] did not capture all contacts since neighbors were searched for every 120 s. Although the resolution of our traces is the same, shorter contacts can be detected because positions of visitors can be interpolated between GPS samples. Besides, the traces in [18] were collected in a conference environment with a lower degree of mobility compared to entertainment parks. Results in Fig. 7 (left) show that the average IACT corresponds well to the device density illustrated in Fig. 4 (top): It decreases with the number of devices and, for the same number of devices, it is shorter in the morning (e.g. at T'_{200}) than in the afternoon (e.g. at T''_{200}).

Beside the density, the number of contact opportunities depends on the level of mobility, which can be measured by the mean square displacement (MSD). Displacement measures how far away a mobile node is from its starting position after some time t. Let $p_t \in \mathbb{R}^2$ be the position of a node at time τ (e.g. in an x–y coordinate system). Mean square displacement after time t is given by $MSD(t) \triangleq \mathbb{E}\{P_{\tau+t} - P_{\tau}^2\}$. $MSD(t)$ increases with t, such that $MSD(t) \sim t^{\gamma}$. The exponent γ indicates the speed of diffusion. For Brownian motion $\gamma = 1$. When $\gamma > 1$, the mobility is superdiffusive. For example, when a node moves on a straight line $MSD(t) \sim t^2$, hence $\gamma = 2$. Nodes whose mobility exhibits stronger diffusion will cover larger area compared to nodes with weaker diffusion. As a consequence, they will encounter more new nodes. The speed of diffusion makes huge impact on the performance of forwarding algorithms [10].

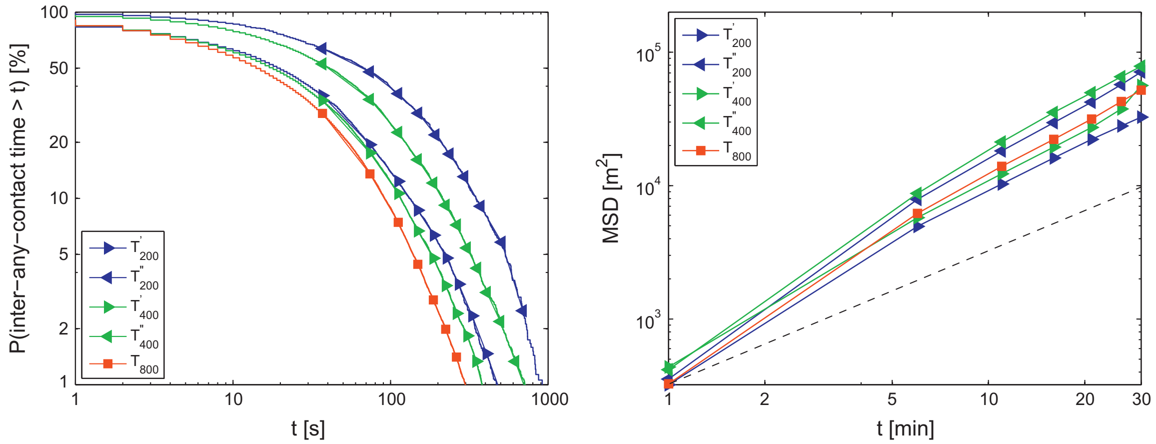


Fig. 7. Epcot traces: Left: CCDF of inter-any-contact times. Right: Mean square displacement (slope of the dashed line is $\gamma = 1$).

We calculated the $MSD(t)$ of each park visitor in a 30-min interval following T (i.e. T'_{200}, T'_{400}, T_{800}, T''_{400}, and T''_{200}) by varying τ from T to $T+30$ min-t. Fig. 7 (right) shows the average $MSD(t)$ at different times of the day on a log–log scale. The initial slopes of the curves $\gamma > 1$ indicate that park visitors exhibit superdiffusive behavior over an interval of $\sim$10 min. The figure shows that the MSD is larger in the afternoon when visitors tend to move faster between the attractions to make the best use of the time before the park closes. The larger MSD, however, did not result in faster message dissemination, as shown in Fig. 5. Higher mobility in the afternoon leads to a wider dispersion of visitors and, therefore, fewer neighbors/contacts. This is obvious when comparing the average number of neighbors (devices within the range of 10 m) and contacts per minute at T'_{400} and T''_{400} or T'_{200} and T''_{200} in Table 4. The table also shows that the percentage of new contacts is rather high. Hence, there are few repeated contact with the same devices within the 30-min interval after T, which is consistent with the superdiffusive behavior observed in Fig. 7 (right). Note however that the lack of GPS data from indoor locations, where repeated contacts are likely to occur, may have affected the percentage of new contacts in Table 4.

5.2. Encounter statistics: DLP traces

The CCDFs of inter-any-contact times in the DLP are shown in Fig. 8 (left). The average IACTs correspond well to the device densities at different times of the day illustrated in Fig. 4 (bottom). As in the Epcot traces, the shape of the distribution of IACTs is best described by the gamma distribution. A major quantitative difference compared to the IACTs in the Epcot traces is observed at T''_{400}: IACTs in the DLP tend to be significantly shorter. Since T''_{400} in DLP traces corresponds to 19:15 h, as shown in Fig. 3 (right), the most probable reasons for the difference are: (1) The Walt Disney Studios closes at 19:00 h; the flow of people out of that park increases the number of encounters. (2) A street parade starts at 19:00 h in the Disneyland Park; many visitors gather at the center of the park to watch the parade, which increases the density and the number of potential encounters. Shorter IACTs at T''_{400} helped to achieve faster message dissemination in the DLP, as shown in Figs. 5 and 6. The average number of neighbors is given in the Table 5. This number is significantly smaller than in Epcot at T'_{200}, T'_{400}, and T_{800}, which is expected from the spatial distributions shown in Fig. 4. The number of neighbors is approximately the same at T''_{200}, and significantly larger at T''_{400} due to (1) and (2). This illustrates the importance of localized events that can only be captured in detailed targeted mobility models. The second row in Table 5 shows the number of contacts per minute. Except at T''_{400}, the frequency of contacts is approximately the same as in Epcot, in spite of the larger number of neighbors in Epcot. This indicates that DLP visitors are more mobile. The MSD shown in Fig. 8 (right) confirms this. Although the MSD is bigger in the DLP traces, the slope of the MSD curves is approximately the same as in the Epcot traces. Hence, the superdiffusive properties of visitors' mobility in the two scenarios are similar. The number of new contacts as a percentage of all contacts is however significantly smaller in the DLP traces, as shown in the bottom row of Table 5. This may appear contradictory to the fact that MSD is bigger, which suggests that DLP visitors are more mobile, and therefore, encounter more new visitors.

Table 4
Epcot traces: Average number of neighbors and contacts.

T	T'_{200}	T'_{400}	T_{800}	T''_{400}	T''_{200}
# of neighbors	1.82	2.40	2.49	0.69	0.31
# of contacts [min^{-1}]	0.47	0.61	0.77	0.29	0.16
% of new contacts	85.4	86.8	88.0	90.7	89.7

Table 5
DLP traces: Number of neighbors and contacts.

T	T'_{200}	T'_{400}	T_{800}	T''_{400}	T''_{200}
# of neighbors	1.09	1.16	1.50	1.08	0.33
# of contacts [min^{-1}]	0.55	0.64	0.81	0.60	0.18
% of new contacts	60.0	65.6	75.8	61.7	77.8

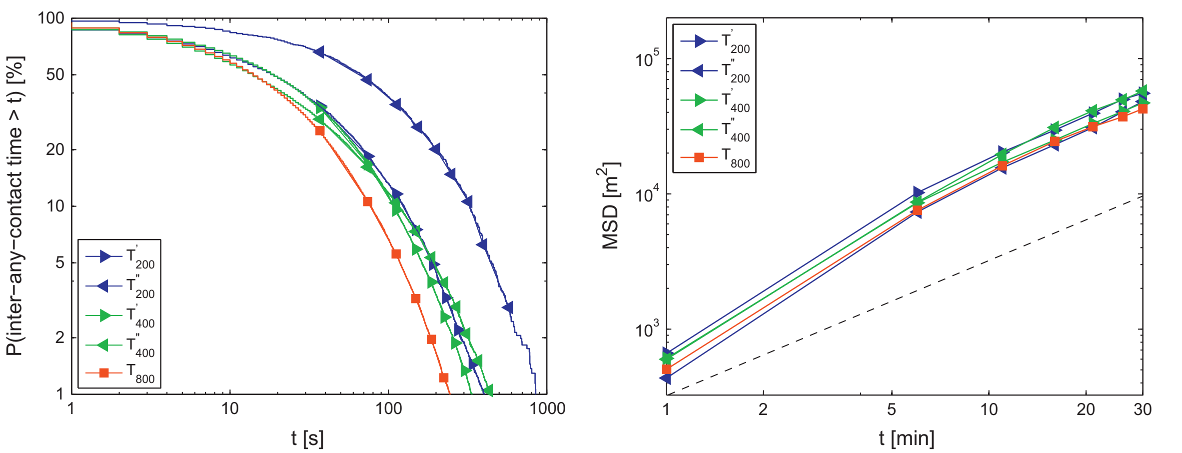


Fig. 8. DLP traces: Left: CCDF of inter-any-contact times. Right: Mean square displacement (slope of the dashed line is $\gamma = 1$).

Please cite this article in press as: V. Vukadinovic et al., Impact of human mobility on wireless ad hoc networking in entertainment parks, Ad Hoc Netw. (2012), http://dx.doi.org/10.1016/j.adhoc.2012.05.011

It could be that coordinated movements of groups of visitors, which would result in repeated contacts of the same pairs of devices, are more common in the DLP. We have not investigated if this is actually the case.

5.3. Other statistics

We analyze inter-contact times (ICTs) and contact durations (CDs) in Epcot and DRP traces, discuss their relevance for the performance of opportunistic networking, and compare their statistical distributions with some previous findings.

Inter-contact time (ICT) is the time elapsed between starts of two successive contacts of the *same pair* of devices. ICT is relatively independent (but not entirely independent) of the number of visitors in a park compared to the inter-any-contact time (IACT). The distribution of ICTs can be used to derive the residual time until a new contact occurs and, therefore, the time until any of the relays encounters the destination of the message. Hence, the distribution of ICTs is closely related to the end-to-end delay of (unicast) forwarding algorithms—longer ICTs lead to longer delays. ICTs have been studied extensively in the literature [9,13,15]. In [9], the authors show that the distribution of ICTs aggregated over all pair of devices exhibits a heavy tail such as one of a power law. This is at odds with the exponential decay implied by most mobility models that are widely used in wireless studies. The authors pose a hypothesis that the mean packet delay of any opportunistic routing scheme is infinite if the power law exponent of ICTs is smaller than or equal to one. Results in [13] indicate that the tail distribution of ICTs actually exhibits a dichotomy—it follows power-law decay only up to some characteristic time, beyond which the decay is exponential. In Fig. 9 (left) we confirm the same tendency in our GPS traces. The figure shows the complementary cumulative distribution function (CCDF) of ICTs aggregated over all pairs of visitors that encountered each other at least twice during a day. The shape of the curve indicates that the tail indeed follows a power low decay up to a characteristic time of ~100 min. The fast drop beyond the characteristic time indicates exponentially decaying tail. The slope of this decay is related to the space boundaries (i.e. size of the park) and the length of the traces (i.e. park visit durations).

Contact duration (CD) is the time two devices remain in contact assuming certain transmission range. Actual transmission link duration is shorter than the corresponding CD because devices need to discover each other and set up the connection. Many short contacts arise when park visitors are bypassing each other. Longer contacts occur when visitors are static and collocated (e.g. while they are waiting in queues or sitting in restaurants). It is therefore not surprising that the distribution of CDs shown in Fig. 9 (right) appears to be heavy-tailed. Contacts need to be longer than the device discovery delay to be useful for content distribution. Hence, the tail of CCDF is of particular interest. Using distribution fitting in Matlab, we found that CDs in DLP traces can be accurately described by the generalized Pareto distribution with shape parameter 0.28 for the transmission range of 10 m and 0.30 m for 50 m. This confirms that the distribution of CDs is indeed heavy-tailed. Interestingly, CDs in Epcot traces could not be matched to the generalized Pareto distribution with the same accuracy. The shape of the tail indicates that their distribution is heterogeneous. We do not have a straightforward explanation for this difference: CDs are in a complex way dependent on the park layout and activities performed by visitors. As expected, our results show that longer transmission range significantly increases the probability of long contacts.

The presented statistics describe encounters of individual nodes. In many practical scenarios, however, clustering of nodes (e.g. due to social grouping) and rate at which clusters split and merge also plays a significant role in content forwarding. A model that translates the split and merge rates to the stationary cluster size distribution is described in [19]. The distribution indicates to what extent a scenario provides partial multi-hop routes that can be used to complement opportunistic forwarding between clusters. It may have important implications for some of the entertainment park application scenarios. Unfortunately, much of the social clustering information is lost in our GPS traces

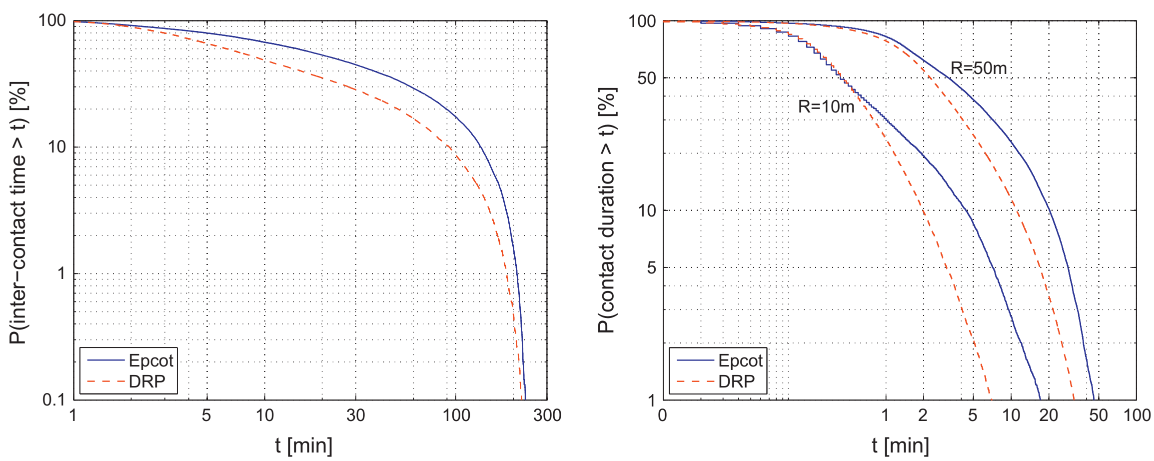


Fig. 9. Left: CCDF of inter-contact times aggregated over all pairs of nodes. The impact of the transmission range R is insignificant for studied values ($10\ m \leqslant R \leqslant 50\ m$). Right: CCDF of contact durations for various transmission ranges R.

since, in case of groups, only one of the members was given a GPS device.

6. Mobility simulator

There is already a number of pedestrian mobility models and simulation tools, which are used for research in various fields, including wireless networking. Some of them are mentioned in the following. So why do we need yet another mobility simulator? Available tools are either simplistic and do not correspond to any specific real-world mobility scenario, or they target some common scenarios, such as urban working-day mobility. For practical applications, models must target specific mobility scenarios, even if this limits the scope of their applicability. We created a model that targets entertainment park mobility and captures many features that are specific for such parks. Entertainment park mobility could be of significant interest to the research community because of the variety of wireless applications that can be deployed there.

6.1. Related mobility models

It is impossible to capture all details of human mobility in a model. Therefore, models typically focus on aspects that are considered relevant for their intended use (e.g. research in the areas of transportation, urban planning, social sciences, or in our case wireless networking). According to [22], human mobility can be described at three levels: strategic, tactical, and operational.

- At the *strategic* level, people decide on their destinations based on activities that they want to perform, as well as the order in which the activities are performed. It is extremely difficult to model human behavior at this level. Existing models focus on particular mobility scenarios (e.g. working day mobility, airport mobility). They are often empirical and based on observations and surveys.
- At the *tactical* level, people decide which routes to take between destinations. In reality, route choices are often results of some complex utility maximization which includes many route parameters (travel time, distance, attractiveness, safety, etc.). Yet most route-choice models used in simulations, including ours, are based on simple shortest path algorithms.
- At the *operational* level, people chose their walking speed and direction to avoid collisions with obstacles and other pedestrians. Mobility models used in the research of wireless networks typically assume free, unobstructed flow of people. However, in ad hoc networks, interaction among pedestrians may affect the rate and the duration of contact opportunities used for data transfer.

Mobility at the strategic level is often referred to as *macro-mobility*. The tactical and operational levels constitute *micro-mobility*. Many of the available models focus on micro-mobility and lack proper representation of macro-mobility (choice of destinations driven by person's activities) or vice versa. Our model includes both macro- and micro-mobility. It is still computationally simple enough to simulate the mobility of tens of thousands of people in real-time. Here we provide a brief overview of existing mobility models and simulation tools with a special focus on those commonly used in wireless networking research.

The Random Waypoint (RWP) [23] is a popular model that describes mobility patterns of independent, non-interacting nodes in an open area. Each node moves along a zigzag line from one waypoint to the next. RWP model is elementary and it is easy to argue about the paths being unnatural. It owns its popularity to simplicity and mathematical tractability. The RWP has been used as a basis for many other models. Based on empirical observations, [11] proposes Levy Walk model, which assumes power-law distributions of trip lengths between waypoints and pause times in the waypoints. Constrained Random Waypoint models, such as the one proposed in [24], include geographical restrictions—nodes are restricted to travel between waypoints using pathways. A survey of random walk models is provided in [25]. The macro-mobility in these models is simplistic; destinations and their visit order are chosen randomly without any strategic planning. This type of mobility hardly represents any realistic scenario. On the operation level, free unobstructed movement of people is typically assumed. The effect of these assumptions in the performance evaluation of wireless ad hoc networks has been studied in [26].

Similar to our model, some models have been derived based on observed mobility and geographical data. For example, the Weighted Waypoint model [27] defines a set of destinations, such as restaurants and classrooms on a university campus. Based on empirical distributions of pause times at each destination and transition probabilities between the destinations, the model constructs a Markov chain of nodes' movements. The UDel mobility model [28] aims to capture a typical day-cycle of a working person based on data collected by the US Bureau of Labor Statistics. In the model, the mobility of people is driven by activities performed on a realistic city map: schedules and durations of activities are calibrated based on the data. Similarly, the model in [29] uses data from the US National Household Travel Survey. A drawback of using statistics aggregated from various locations is that models become rather generic. Other examples of empirical models are the Working Day Movement model implemented in the ONE simulator [30] and the model implemented in the CanuMobiSim simulator [31]. Most empirical models employ strategic destination planning based on activity scheduling. Scheduled activities are mapped to destinations where they can be performed. Activity-based user modeling belongs to the mature field of travel demand modeling in transportation research. Surveys of activity-driven approaches are provided in [32,33].

Elaborate efforts to capture human walking behavior on the operation level have been made in the fields of urban and transportation planning and traffic engineering. The social force model (SFM) described in [34] assumes that each pedestrian is driven by two types of forces: social and physical. The social forces reflect the intentions of pedestrians not to collide with other people and obstacles; in response to these forces pedestrians accelerate or decelerate. Several commercial multi-agent simulators use the SFM to generate mobility; the most notable are VISSIM [35], Legion Studio [36], and SimWalk [37]. Besides the SFM, cellular automata models (CAM) provide convenient way to capture human walking behavior [38]. In the CAM models, walking space is divided into small cells, which can generate potential fields that represent the local effect of obstacles and other pedestrians on walking direction and speed. PedGo [39] is an example of a simulator based on the CAM. VISSIM, Legion Studio, SimWalk, and PedGo are often used in urban and traffic planning to design public spaces such as airports, subway stations, and sport stadiums. They aim to capture speed-distance relations that emerge when pedestrians navigate obstacles and other pedestrians. These relations are responsible for the formation of pedestrian crowds—capturing them is of paramount importance in emergency evacuation scenarios. These models are often called microscopic since they focus on walking behavior rather than on strategic macro-mobility. Therefore, defining large-scale mobility scenarios using these tools requires significant efforts from users. Furthermore, they are often computationally intensive, which limits the number of pedestrians that can be simulated.

6.2. Park representation

The first step to simulate the mobility of entertainment park visitors is to create a representation of park's spatial layout in the simulator. ParkSim does not implement its own layout editor where a park could be drawn. Instead, the park layout is specified in the OpenStreetMap (OSM) format [40]. Therefore, any OSM editor can be used. An advantage of using the OSM is that large parts of major entertainment parks are already mapped in details, as shown in Fig. 1 and Fig. 2. The OSM maps are easily parsed by the simulator because they use XML syntax. We distinguish between two types of areas in parks: walking areas and activity areas. Visitors use walking areas to move between activity areas. The outlines of the areas need to be specified in the OSM map editor before the park map is parsed by the simulator.

6.2.1. Walking areas

Most walking areas are already specified as such in OSM maps. We distinguish between two types of walking areas: walkways and plazas. In the OSM, a walkway is represented by a series of waypoints. In ParkSim, this series is split up into segments where each segment contains only two consecutive waypoints. The shape of segment is then computed as a rectangle whose height is equal to the distance between the waypoints and the width is set according to the assumed width of the walkway. Fig. 10 (left) illustrates the geometry of a walkway. A plaza is represented as a polygon with an arbitrary number of edges. A number of connection points located at the edges of the polygon connect the plaza to the adjoining walkways, as shown in Fig. 10 (right). For simplicity, we assume that visitors use a shortest path when they move between two connection points (i.e. two adjoining walkways) across a plaza. One drawback of this is that a shortest path may fall outside of a concave polygon. This can be avoided by splitting up such plaza into multiple smaller convex plazas.

6.2.2. Activity areas

Activity areas are locations where park visitors perform typical park-related activities. We distinguish between four different types of activity areas: attractions, rides, restaurants, and event areas. The activity areas need to be specified as polygons in the OSM editor before being parsed by the simulator. The characteristics and parameters that can be configured for the activity areas are:

- An attraction is an indoor or outdoor location that visitors can visit. The number of visitors that an attraction can accommodate is limited by the available space, personal space requirements, and safety regulations. However, an attraction is usually able to accommodate all its visitors except when a park is extremely crowded.

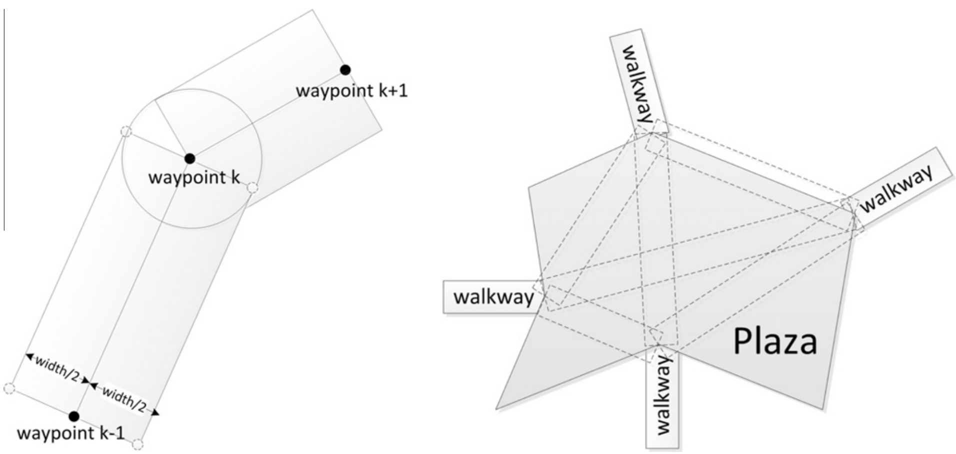


Fig. 10. Left: Geometric representations of a walkway. Right: A plaza with adjoining walkways.

There is a single entrance/exit point to each attraction from adjoining walking areas. The probability distribution of visit durations and mobility model by which visitors move inside an attraction can be specified in the simulator.

- A ride is a special type of an attraction whose capacity is limited to the extent that an entrance policy must be specified. There is a queuing space at the entrance to the ride where visitors can wait to enter if the ride is currently full. A reservation scheme described in Section 4.1 allows visitors to take a so-called *fastpass* ticket for the ride to avoid long queues. The capacity and ride's duration can be specified in the simulator.
- A restaurant is an eating area, which has a specified capacity, but no queue at the entrance. When in a restaurant, visitors are static. The capacity and the probability distribution of visit durations can be specified in the simulator.
- An event area is an area where visitors gather to watch some popular park events, such as street performances and fireworks. An event area encompasses some of the walking areas on which visitors may stand to watch the event. The starting time, duration, and popularity of an event can be specified in the simulator.

6.3. Mobility model

Visitors arrive to the park entrance according to an empirical arrival rate distribution. The arrival rate depends on the time of the day. The total number of arrivals is a parameter of choice (parks tend to be more crowded on weekends and holidays). Once they pass the entrance gate, their mobility is driven by the model implemented in the ParkSim simulator. The model describes mobility at two levels: Macro mobility determines how visitors select activity areas to visit and how they prioritize between different activities. Micro mobility determines how they move between and inside activity areas and how they avoid colliding with each other in walking areas and in queue lines. The total time that a visitor spends in the park is drawn from an empirical visit time distribution. At the end of his visit, the visitor walks towards the exit.

6.3.1. Macro mobility

A visitor can be in one of the following three states: *walking* (visitor is moving in a walking area), *visiting* (visitor is visiting an activity area), or *queuing* (visitor is waiting to enter a ride). Initially, a visitor is in the walking state and appears at the park entrance. He then chooses to visit one of the activity areas that are specified as possible initial destinations. The initial destinations are typically located close to the park entrance. The visitor walks towards the initial destination (activity area) using the shortest path provided by the walking areas.

At the destination, the visitor can either (i) enter the activity area if the area is not full; his state changes to *visiting*, (ii) if the activity area is a ride, the visitor can join the queue; his state changes to *queuing*, (iii) decide to visit the area at some later time; his state remains unchanged as he walks toward the next destination. When a visitor enters an activity area, he randomly chooses visit duration from an area-dependent visit duration distribution. In case of rides and events, the visit duration is deterministic. While inside an activity area, the visitor moves according to an area-specific micro-mobility model. At the end of the visit, the visitor chooses his next activity area based an activity matrix. The activity matrix contains probabilities with which visitors chose to visit other attractions and rides in the park, given the last visited attraction or ride. The probabilities are derived from GPS traces of park visitors. Currently, the activity matrix does not account for possible differences between visitors of different age and/or sex. Visits to the restaurants and event areas are not driven by the activity matrix, but by visitors' hunger and by a timetable of park events, respectively. When a visitor selects his next destination, his state changes to *walking*. The state transition diagram is shown in Fig. 11 (left).

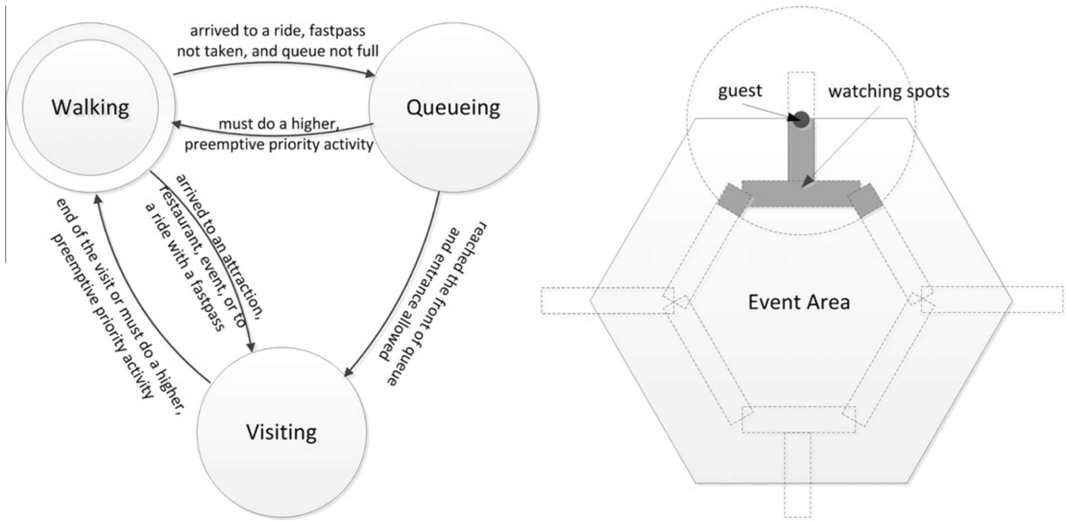


Fig. 11. Left: A park guest can be in one of the three states: walking, queuing, or visiting. The diagram shows possible transitions between the states. Right: An example of an event area with an indication of possible watching spots.

Table 6
Different types of activities are scheduled and prioritized.

Activity area	Visits driven by	Priority
Attraction, ride	Activity matrix	Low non-preemptive
Restaurant	Guest's hunger	Medium non-preemptive
Event	Event timetable	Medium non-preemptive or high preemptive
Ride (with a fastpass)	Time on the fastpass	High preemptive

Different activity areas (attractions, rides, restaurants, event areas) have different priorities. It may happen that, for example, a visitor interrupts his visit to an attraction in order to attend an event. Therefore, an event may have a preemptive priority over some other activities. A visitor may, for example, decide to visit a restaurant, but he will do so after his visit to an attraction is finished. Hence, a restaurant visit has a non-preemptive priority over attractions. Priorities of different activities are summarized in Table 6. Note that different events may have different priorities.

In the following, we provide more detailed descriptions of how visitors queue to enter rides, and how they visit restaurants and events.

6.3.1.1. Fastpass model. In some entertainment parks, a visitor who does not want to wait in a queue to enter a ride may take a so-called *fastpass*. The fastpass specifies a 30-min slot in which the owner has to return to the entrance of the ride in order to enter without queuing. In our simulator, a visitor decides to join the queue or to take a fastpass depending on the estimated waiting time T_W, which is calculated based on the number of visitors in the queue, the capacity of the ride, and the ride's duration. In entertainment parks, estimated waiting times are often displayed on screens. Since we lack empirical evidence, we assume that the probability of taking a fastpass P_{FP} increases linearly with T_W, and it is equal to one for $T_W > 90$ min. The visitor joins the queue with the probability $1\text{-}P_{FP}$. The starting time of the fastpass validity is calculated as the current time plus the estimated waiting time, rounded up to the next full or half-to-full hour. A visitor can visit other activity areas until the validity of the fastpass starts. However, he gives a high priority to the ride for which he holds a fastpass: if he is visiting another activity area, he will estimate the time to walk back to the ride and leave the current area early enough to use his fastpass. A visitor can hold only one fastpass at a time.

6.3.1.2. Restaurant visit model. The probability that the next activity area visited by a visitor is a restaurant increases linearly with the time spent in the park. Six hours after entering a park, the visitor will eat at least once. The probability of choosing a particular restaurant is inversely proportional to the walking distance to the restaurant. In addition to the distance, restaurant popularity could be accounted for. However, due to the lack of empirical data, we currently assume that all restaurants are equally popular. The popularity of restaurants could not be extracted from the GPS traces because most restaurants are collocated with the attractions (e.g. inside the same building). Therefore, it was impossible to conclude if a person is sitting in a restaurant or visiting/watching the attraction. A side-effect of this is that the visit time to attractions extracted from the GPS traces may include the time spent in the restaurant inside that attraction.

6.3.1.3. Event visit model. An event area is defined in the OSM map as a polygon that encompasses some of the park's walking areas. The walking areas inside the event area can be used by visitors to select a spot to watch the event (e.g. firework, street performance). When a visitor arrives to the event area, he will select a random spot in the walking areas that are within a certain radius, as shown in Fig. 11 (right). This avoids overcrowding of the walking areas when only few walkways enter the event area. An event has a starting time, duration, spot selection radius, popularity, and priority as parameters. The popularity is the probability that a random visitor in the park will attend the event. An event may have a medium non-preemptive priority (e.g. a street parade) or high preemptive priority (e.g. a park evacuation).

6.3.2. Micro mobility

Micro-mobility determines how visitors select routes between activity areas, how they avoid colliding with each other in walking areas, and how they move inside activity areas and queues. Micro mobility is responsible for the formation of pedestrian crowds (e.g. platooning on congested sidewalks). In opportunistic networks, micro mobility affects the time that two radio devices stay within each other's range and the number of devices they encounter per unit time.

6.3.2.1. Routing. In theme parks, most people choose the shortest path to the next activity area based on, for example, a map of the park that is handed out to them at the ticket counter or based on posted signs. In many cases, the next attraction will be within their sight. ParkSim uses the Dijkstra algorithm to calculate shortest paths between activity areas. The algorithm computes an ordered list of walkways that determines waypoints for walking. A visitor walks straight towards the next waypoint if not disturbed by other visitors; otherwise the direction of his movement is determined by a collision avoidance algorithm. To avoid cases where all visitors walk in the middle of the walkway, a random offset is added to the waypoint so that the visitor actually moves towards a randomly selected point on the ending edge of the current segment. Target walking speeds of visitors are drawn from a specified speed distribution. The actual walking speeds depend on the crowdedness of a park.

6.3.2.2. Collision avoidance. Simulated visitors try to avoid colliding with each other while walking. The visitors are represented by circles of a specified radius. Each visitor has a rectangular field of view, as shown in Fig. 12 (left). The figure illustrates a scenario where two visitors would collide after three simulation timeslots/steps if their walking directions remain unchanged. The collision avoidance algorithm extrapolates the trajectories of all persons with-

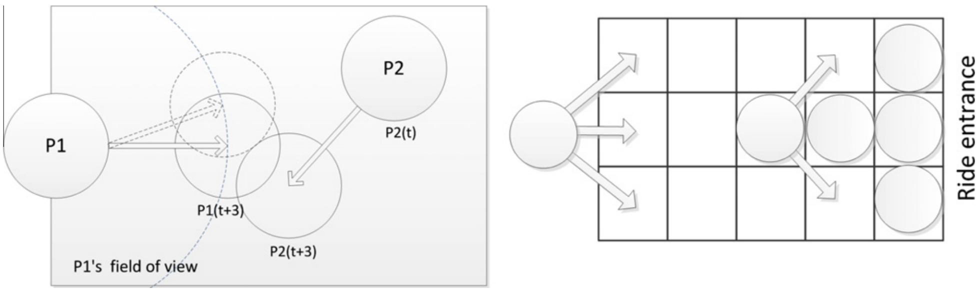


Fig. 12. Left: An example of a potential collision that is avoided by changing the direction of movement. Right: An example of a queue line structure.

in the field of view of a visitor in order to predict impeding collisions. If one or more collisions are predicted, the algorithm changes the direction and/or the speed of the visitor to avoid the collision(s). The algorithm will first try to change the directions without changing the speed. It this is not possible, for example because the visitor would have to step outside of the walking area, the speed is changed. We omit the details of the algorithm for brevity. The algorithm is executed at each time step for each visitor independently. It is a best-effort algorithm, which does not ensure a completely collision free movement, but it captures the effect of crowded walking areas on the speeds of the visitors.

6.3.2.3. Queuing behavior. A queue line is modeled as a grid of cells with specified width and length. Each cell provides space for one visitor, as shown in Fig. 12 (right). When entering the queue, a visitor selects a free cell that is closest to the entrance; ties are broken randomly. The visitor moves forward whenever some of the visitors in front of him are allowed to enter the ride. Since the number of cells is limited, it may happen that an arriving visitor finds all cells occupied. In such cases, he selects another activity area to visit, according to the activity matrix. Note that a visitor who decides to take a fastpass does not enter the queue.

6.3.2.4. Intra-area mobility. When they are inside activity areas, the mobility of visitors is different from the mobility in walking areas. Currently, there are two intra-area mobility models in the ParkSim: Random Waypoint and Random Sitpoint models. In the popular Random Waypoint model [23], a visitor moves from one random waypoint to the other along a zigzag line that connects them. This model is used for some attractions and rides. In the Random Sitpoint model, a visitor moves towards a random point inside an activity area and remains there until the end of his visit. This model is used for restaurants, event areas, and some attractions and rides. In the future, we will modify the indoor mobility models for some of the activity areas to better capture the true behavior of people at those locations.

6.4. Model calibration and validation

We use the Epcot park as an example to describe how we calibrated and validated the model. The layout of the park is shown in Fig. 1. The ParkSim model of the park contains 22 attraction areas, six rides, 12 restaurants, and an event area centered around the lake. During the calibration, various parameters of the model, described in Sections 6.2 and 6.3, are set based on data collected in the park. The data comes from two sources: long-term attendance statistics collected by the park management and GPS traces of park visitors described in Section 3. The hourly distribution of arrivals to the park, schedules and durations of park events (e.g. fireworks on the lake), and ride durations are obtained from the park management. The activity matrix and the distributions of visit durations for some of the activity areas are obtained from the GPS traces. Most of the data is confidential and, therefore, not disclosed in this paper. Some of the model parameters are set arbitrary due to the lack of relevant data (e.g. visit durations for the restaurants). A sample mobility trace obtained from the ParkSim is shown in Fig. 13 as an illustration. The model validation focuses on the opportunistic broadcasting scenario described in Section 4 and on the encounter statistics analyzed in Section 5.

We used synthetic mobility traces produced by the ParkSim to simulate the message dissemination in the Epcot. We measured the time needed to deliver the message to 75%, 90%, and 98% of devices at different times of

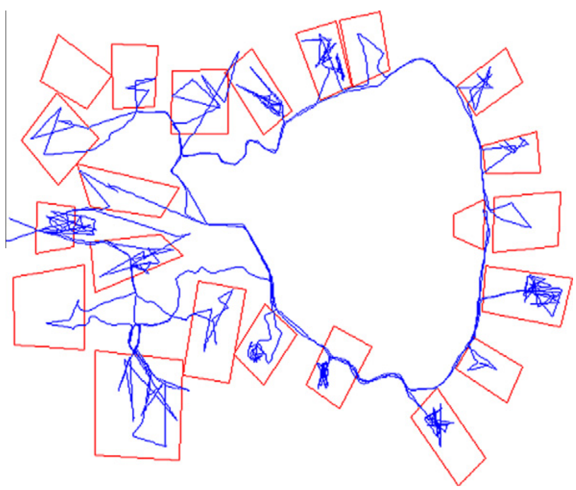


Fig. 13. A sample mobility trace from ParkSim. Not all activity areas are shown in the figure.

Table 7
Time needed to distribute the message to 75%, 90%, and 98% of the devices.

T	GPS traces			ParkSim traces		
	$\Delta T_{75\%}$ (s)	$\Delta T_{90\%}$ (s)	$\Delta T_{98\%}$ (s)	$\Delta T_{75\%}$ (s)	$\Delta T_{90\%}$ (s)	$\Delta T_{98\%}$ (s)
T'_{200}	239	651	1855	172	565	1590
T'_{400}	545	990	1984	410	787	1765
T_{800}	621	1033	1556	589	929	1415
T''_{400}	1144	1484	2790	944	1220	2394
T''_{200}	3885	4040	4320	2818	3458	3776.

the day. The results are summarized in Table 7 together with the results obtained earlier with the GPS traces. The table shows that the dissemination delay is somewhat shorter with the synthetic ParkSim traces. The following simplifications in the ParkSim model have caused the discrepancy:

(1) Some of the lesser attractions, visitor service areas, shops, and walkways are not included in the model. Hence, the model restricts the movement of visitors to fewer locations in the park, which results in higher density and more frequent encounters.
(2) The model assumes Random Waypoint or Random Sitpoint mobility inside activity areas. In reality, visitors move differently in each of the activity areas: In some they roam freely, while in others they move in groups following pre-determined routes. In some they move on foot, while in others they ride vehicles (roller coasters, mini cars, flight simulators). The simplified intra-area mobility model may overestimate or underestimate the frequency of encounters, depending on the particular activity area.
(3) The activity matrix, which contains transition probabilities between attractions, is assumed to be constant over time. It is constructed based on average day-long statistics extracted from the GPS traces. In reality, the transition probabilities may depend on the time of the day and the number of visitors in the park.

The simplifications (1) and (2) are needed in order to abstract away fine details of the park layout. Our intention is to avoid micro-modeling of the park at the level of individual activity areas. We consider the discrepancies in Table 7 acceptable, especially considering that, due to the time granularity of GPS sampling and unavailability of GPS signal at indoor location, the performance obtained with the GPS traces is only indicative, and not the "ground truth" with respect to which the correctness of the model can be accurately measured. The simplification (3) requires further attention. Since the activity matrix contains *average* transition probabilities observed in the GPS traces, it is biased towards the peak hour when the largest number of transitions is observed. Therefore, it is not surprising that the discrepancy in Table 7 is smaller for T_{800} than for T'_{200} and T''_{200}. We will continue to investigate the presence and significance of hourly changes in transition probabilities to determine if such changes should be included in the model.

We also extracted the average number of neighbors, number of contacts, and percentage of new contacts from the ParkSim traces. The results are shown in Table 8 together with the results extracted from the GPS traces. The number of neighbors in the ParkSim traces is slightly lower than in the GPS traces. This may seem contradictory to (1). Note, however, that the number of neighbors in *indoor* activity areas is overestimated in the GPS traces since all visitors of an indoor area appear to be close to its entrance. It can also be observed from Table 8 that the number of contacts per minute is significantly larger in the ParkSim traces, while the percentage of new contacts is smaller. This is due to the Random Waypoint mobility, which is assumed for some of the activity areas. Such intra-area mobility results in frequent encounters, many of which are repeated. The crucial factor for the speed of message dissemination is the number of new contacts, which is the product of the number of contacts and the percentage of new contacts. The number of new contacts per minute is approximately equal for the Epcot and ParkSim traces. While intra-area mobility may have a significant impact on contact durations and number of repeated contacts, new contacts occur mostly due to macro-mobility (i.e. movements of visitors from one activity area to the other). Therefore, for the message dissemination scenario, it is sufficient that the macro-mobility of visitors is captured in the model. For some other application scenarios, it might be necessary to expand the ParkSim with more detailed intra-area mobility models.

Table 8
Number of neighbors and contacts assuming transmission range of 10 meters.

T	GPS traces			ParkSim traces		
	# neighbors	# contacts [min^{-1}]	% new contacts	# neighbors	# contacts [min^{-1}]	% new contacts
T'_{200}	1.82	0.47	85.4	1.74	0.80	54.1
T'_{400}	2.40	0.61	86.8	2.27	0.93	53.9
T_{800}	2.49	0.77	88.0	2.31	1.19	55.2
T''_{400}	0.69	0.29	90.7	0.58	0.48	55.9
T''_{200}	0.31	0.16	89.7	0.26	0.30	52.8

7. Conclusions

Intermittent connectivity could be useful for many entertainment park applications if efficient routing/forwarding algorithms are designed. To be practical, the algorithms must target specific applications and mobility scenarios. We studied the mobility of park visitors based on a fairly large dataset of GPS traces. Using a broadcast application as an example, we showed the impact of hourly changes in the number of devices and their spatial distribution on the speed of content dissemination. We analyzed several contact-related statistics to interpret the observed performance. Since our entertainment park scenarios assume sparse deployment of infrastructure nodes, the density of mobile devices is crucial to reduce the delay of content delivery. Mobility models and simulations are needed to evaluate scenarios where the number of devices exceeds the number of available GPS traces. Our results suggest that generic mobility models are not sufficient: Targeted mobility models are needed in order to capture the non-stationarity in the number and spatial distribution of nodes. Therefore, we developed the ParkSim tool to simulate the mobility of entertainment park visitors. It implements an empirical mobility model based on data collected in entertainment parks. Synthetic mobility traces produced by ParkSim are validated against the GPS traces of park visitors. The results indicate that ParkSim can be a useful tool to assist performance evaluations of wireless ad hoc networks. The tool can easily be adapted for scenarios where pedestrians exhibit similar mobility patters, such as trade shows, zoos, open-air museums, and multi-stage festivals.

Acknowledgment

Our thanks to Amber Brown from Disney Research who collected the GPS traces used in this study.

References

[1] The Delay Tolerant Networking Research Group (DTNRG), <http://www.dtnrg.org> (accessed: 06.05.11).
[2] K. Fall, A delay-tolerant network architecture for challenged internets, in: Proc. ACM Sigcomm, Karlsruhe, Germany, 2003.
[3] S. Jain, K. Fall, R. Patra, Routing in a delay tolerant network, in: Proc. ACM Sigcomm, Portland, USA, 2004.
[4] M. Montola, J. Stenros, A. Waern, Pervasive Games: Theory and Design, Morgan Kaufman Game Design Books, 2009.
[5] Kim Possible, <http://disneyworld.disney.go.com/parks/epcot/attractions/kim-possible> (06.05.11).
[6] P. Hui, R. Mortier, M. Piorkowski, T. Henderson, J. Crowcroft, Planet-scale Human Mobility Measurement, in: Proc. ACM Mobisys HotPlanet, San Francisco, USA, 2010.
[7] E. Daly and M. Haahr, Social network analysis for routing in disconnected delay-tolerant MANETs, in: Proc. ACM MobiHoc, Montreal, Canada, 2007.
[8] P. Hui, J. Crowcroft, E. Yoneki, BUBBLE Rap: Social-based forwarding in delay tolerant networks, in: Proc. ACM MobiHoc, Hong Kong, China, 2008.
[9] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, J. Scott, Impact of human mobility on the design of opportunistic forwarding algorithms, in: Proc. IEEE Infocom, Barcelona, Spain, 2006.
[10] S. Kim, C.-H. Lee, D.Y. Eun, Superdiffusive behavior of mobile nodes and its impact on routing protocol performance, IEEE Trans. Mobile Computing 9 (2) (2010) 288–304.
[11] I. Rhee, M. Shin, S. Hong, K. Lee, S. Chong, On the Levy-walk nature of human mobility, in: Proc. IEEE Infocom, Phoenix, USA, 2008.
[12] T.S. Azevedo, R.L. Bezerra, C.A.V. Campos, L.F.M. de Moraes, An analysis of human mobility using real traces, in: Proc. IEEE WCNC, Budapest, Hungary, 2009.
[13] T. Karagiannis, J.L. Bourkec, M. Vojnovic, Power law and exponential decay of inter contact times between mobile devices, in: Proc. ACM Mobicom, Montreal, Canada, 2007.
[14] W. Hsu, A. Helmy, On nodal encounter patterns in wireless LAN traces, IEEE Trans. Mobile Comput. 9 (11) (2010) 1563–1577.
[15] V. Conan, J. Leguay, T. Friedman, Characterizing pairwise intercontact patterns in delay tolerant networks, in: Proc. ACM Autonomics, Rome, Italy, 2007.
[16] F. Legendre, V. Lenders, M. May, Gunnar Karlsson, Narrowcasting: an empirical performance evaluation study, in: Proc. ACM Mobicom Workshop on Challenged Networks (CHANTS), San Francisco, USA, 2008.
[17] Ó. Helgason, S.T. Kouyoumdjieva, G. Karlsson, Does Mobility Matter?, in: Proc. IEEE/IFIP WONS, Kranjska Gora, Slovenia, 2010.
[18] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, C. Diot, Pocket switched networks and human mobility in conference environments, in: Proc. ACM Sigcomm Workshop on Delay Tolerant Networking, Philadelphia, USA, 2005.
[19] S. Heimlicher, K. Salamatian, Globs in the primordial soup: the emergence of connected crowds in mobile wireless networks, in: Proc. ACM MobiHoc, Chicago, USA, 2010.
[20] K. Collins, G. Muntean, S. Mangold, Supporting mobile devices with wireless LAN/MAN in large controlled environments, IEEE Commun. Mag. 48 (12) (2010) 36–43.
[22] S.P. Hoogendoorn, P.H.L. Bovy, W. Daamen, Microscopic pedestrian wayfinding and dynamics modeling, in: Pedestrian and Evacuation Dynamics, Springer Verlag, 2001, pp. 123–154.
[23] D.B. Johnson, D.A. Maltz, Dynamic source routing in ad hoc wireless networks, in: Mobile Computing, Springer, 1996, pp. 153–181.
[24] J. Tian, J. Hahner, C. Becker, I. Stepanov, K. Rothermel, Graph-based mobility model for mobile ad hoc network simulation, in: Proc. 35th Annual Simulation Symposium, San Diego, USA, 2002.
[25] F. Bai, A. Helmy, A survey of mobility models in ireless ad hoc networks, a chapter in Wireless Ad Hoc and Sensor Networks, Springer, Verlag, 2007.
[26] F. Bai, N. Sadagopan, A. Helmy, IMPORTANT: a framework to systematically analyze the Impact of Mobility on Performance of RouTing protocols for Ad-hoc NeTworks, in: Proc. IEEE Infocom, San Francisco, USA, 2003.
[27] W.-J. Hsu, K. Merchant, H.-W. Shu, C.-H. Hsu, A. Helmy, Weighted waypoint mobility model and its impact on ad hoc networks, ACM Mobile Comput. Commun. Rev. 9 (1) (2005) 59–63.
[28] J. Kim, V. Sridhara, S. Bohacek, Realistic mobility simulation of urban mesh networks, Ad Hoc Networks 7 (2) (2009) 411–430.
[29] Q. Zheng, X. Hong, and J. Liu, An agenda based mobility model, in: Proc. 39th Annual Simulation Symposium, Huntsville, USA, 2006.
[30] F. Ekman, A. Keranen, J. Karvo, J. Ott, Working day movement model, in: Proc. ACM Workshop on Mobility Models for Networking Research, Hong Kong, China, 2008.
[31] I. Stepanov, P. Marron, K. Rothermel, Mobility modeling of outdoor scenarios for MANETs, in: Proc. 38th Annual Simulation Symposium, San Diego, USA, 2005.
[32] H. Timmermans, Models of activity scheduling behavior, Stadt Region Land 71 (2001) 33–47.
[33] B. Konig-Ries, M. Klein, T. Breyer, Activity-based user modeling in wireless networks, Mobile Networks and Applications, vol. 11, Springer Verlag, 2006, pp. 267–277.
[34] D. Helbing, P. Molnar, Social force model for pedestrian dynamics, Phys. Rev. E 51 (5) (1995) 4282–4286.
[35] VISSIM, <http://www.vissim.de> (26.06.11).
[36] Legion, <http://www.legion.com> (26.06.11).
[37] SimWalk, <http://www.simwalk.com> (26.06.11).
[38] V.J. Blue, J.L. Adler, Cellular automata micro-simulation of bi-directional pedestrian flows, Transport. Res. Part B 35 (3) (2000) 135–141.
[39] PedGo, <http://www.traffgo-ht.com/en/pedestrians/products/pedgo/index.html> (26.06.11).
[40] OpenStreetMap, <http://www.openstreetmaps.org> (26.06.11).
[41] V. Vukadinovic, S. Mangold, Opportunistic wireless communication in theme parks: a study of visitors mobility, in: Proc. ACM MobiCom Workshop on Challenged Networks (CHANTS), Las Vegas, USA, 2011.
[42] V. Vukadinovic, F. Dreier, S. Mangold, A simple framework to simulate the mobility and activity of theme park visitors, in: Proc. Winter Simulation Conference (WSC), Phoenix, USA, 2011.

[43] V. Vukadinovic, S. Mangold, Performance of collaboraive GPS localization in Pedestrian Ad Hoc networks, in: Proc. ACM Int. Workshop on Mobile Opportunistic Networking (MobiOpp), Zurich, Switzerland, 2012.

Vladimir Vukadinovic received his Ph.D. degree in electrical engineering from the Royal Institute of Technology, Stockholm, Sweden in 2010 and his M.Sc. and Dipl.-Ing. degrees in telecommunications and electrical engineering from the University of Belgrade, Serbia, in 2004 and 2000, respectively. He works as a postdoctoral researcher at Disney Research Zurich, Switzerland. His research interests are focused on wireless multimedia networking, delay and disruption tolerant communication, and modeling and simulation of human mobility. His e-mail address is vvuk@disneyresearch.com.

Fabian Dreier received his M.Sc. degree in computer science from the Swiss Federal Institute of Technology, Zurich, Switzerland in 2011. During his masters studies he worked as an intern at Disney Research Zurich for one year. The main focus of his work are performance issues of IEEE 802.11 networks and protocols. His e-mail address is fdreier@disneyresearch.com.

Stefan Mangold received his Ph.D and M.Sc. degrees in electrical engineering from the RWTH Aachen University, Aachen, Germany, in 2003 and 1997, respectively. He is a senior research scientist at Disney Research Zurich and lecturer at ETH Zurich, Department of Computer Science. Before joining Disney Research, he worked at Swisscom in research, product and business development, and at Philips Research USA. His research covers many aspects of wireless communication networks and mobile computing, such as wireless protocols and system aspects for IEEE 802.11 Wireless LAN, cognitive radio, and wide area networks. Other research interests include mobile computing, interactive toys, cellular networks, and ontology engineering. He is coauthor of two books, *IEEE 802 Wireless Systems: Protocols, Multi-Hop Mesh/Relaying, Performance, and Spectrum Coexistence* (Wiley, 2007) and *Cognitive Radio and Dynamic Spectrum Access* (Wiley, 2009). His e-mail address is stefan@disneyresearch.com.

Please cite this article in press as: V. Vukadinovic et al., Impact of human mobility on wireless ad hoc networking in entertainment parks, Ad Hoc Netw. (2012), http://dx.doi.org/10.1016/j.adhoc.2012.05.011