Analyzing Metropolitan-area Networking within Public Transportation Systems for Smart City Applications

Torsten Zimmermann, Hanno Wirtz, Oscar Puñal, Klaus Wehrle
Communication and Distributed Systems, RWTH Aachen University
{zimmermann, wirtz, punal, wehrle}@comsys.rwth-aachen.de

***Abstract*—In the scope of smart cities, mobile participatory sensing and metropolitan area networking on top of public transportation systems for communication offers widespread dissemination of information in both time and spatial domains. Specifically, the transportation network naturally reflects urban human mobility patterns between places of interest and interconnects hotspots where information is created and consumed by city-wide applications. Previous work has targeted communication exclusively between users of the public transportation system. In this paper, we provide an analysis of metropolitan networking *within* the transportation system itself. Instead of relying on user-generated contact traces purely between mobile entities, i.e., busses, we build on a comprehensive data set that contains the schedules and location data of busses as well as the location of infrastructure elements, such as bus stops. Our analysis shows the general feasibility of such a network as well as the, previously not considered, impact of infrastructure elements for information dissemination. The latter motivates delay-tolerant and location-driven communication, as well as participatory sensing using the transport system as a communication infrastructure.**

I. Introduction

In urban scenarios, public transportation systems are ubiquitous, spanning the area of the city and connecting vital places of interest, commerce, and transportation. This motivates the use of mobile elements of the transportation system, such as busses, as the basis for a mobile, contact-driven communication network along the transportation routes. Previous work has investigated the feasibility and performance of such a network on campus [1], between users of the Chicago bus system [2], and even in whitespace networking [3], [4].

The applications for a metropolitan mobile network based on a public transportation system are manifold. One of the possibilities is the realization of *Delay-Tolerant Networking* (DTN) by leveraging mobile elements as data mules for time-insensitive messages. Examples for such applications are wireless advertisements, social networking, and image sharing. Moreover, users may exploit the quasi-deterministic mobility patterns to address messages to specific locations. Within the city, messages then may have semantic meaning only at their target locations, for example a lecture announcement on campus. Additionally, mobile elements can be equipped with sensors that record noise and air quality, among others. The collected information enables *Participatory Sensing* and *Data Collection* applications on a metropolitan scale [5]. We argue that communication on top of transportation networks provides a contribution to smart city scenarios. In this, the resulting communication network allows capturing phenomena on the fly and in an integrated fashion, thus ”capturing the pulse of the city”. In addition, reusing the existing infrastructure of the transportation systems mitigates the need for dedicated deployments, reducing capital and operating expenditures.

In this paper, we analyze the feasibility of establishing a communication network *within* the transportation system, i.e., comprehensively utilizing the infrastructure, mobile elements, and domain knowledge as provided by the transportation system operator. We extend existing analytical efforts that sketch communication networks *on top* of transportation systems [6] or within synthetic traces [7]. Conversely, we propose wireless communication between the vehicular mobile elements as well as with infrastructure such as bus stops. Providing (open) 802.11 wireless network access at and through both these element types opens the network up for user communication and involvement, resulting in a city-wide store-and-forward network that facilitates a variety of applications.

A. Existing Work

Utilizing contacts between mobile transportation entities for communication offers a promising venue of establishing a mobile communication network. In this section, we thus delimit our contribution from existing works.

In *DieselNet* [1], the authors leverage a bus transportation network serving the UMass Amherst campus to establish a DTN. Deriving mobility and contact traces from their findings, the authors analyze the number and duration of bus contacts. *RUTS* [6] proposes a delay tolerant forwarding mechanism that takes the characteristics of urban transport systems into account. In this, the authors focus on exploiting the transportation system for message forwarding between DTN users. The approach is evaluated using synthetically generated bus traces superimposed onto OpenStreetMap (OSM) cartographical material within an opportunistic networking simulator environment[1]. In a similar approach [2], the same authors incorporate trace data from the Chicago Transport Authority (CTA) Bus Tracker API[2]. An analysis of the trace data reveals the cyclic nature of the underlying transportation network. However, no analysis of this real-world setting for communication is provided.

We differ from the aforementioned works in that we strive to integrate the transportation system, and the resulting communication network, into the operation of a smart city. We

[1] http://www.netlab.tkk.fi/tutkimus/dtn/theone/
[2] http://www.transitchicago.com/developers/bustracker.aspx

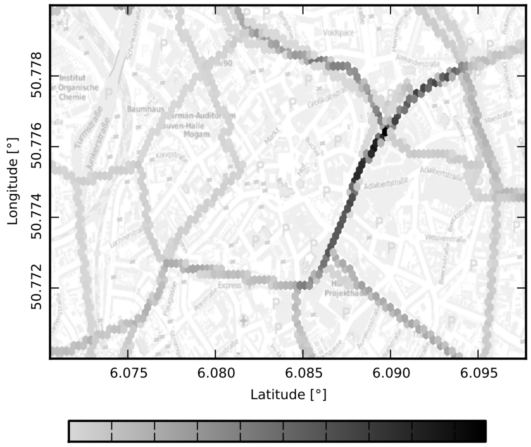


Fig. 1. Heat map of bus locations in Aachen's city center over the evaluation time frame as reported by the simulator in one second granularity.

take a holistic view of the communication network, thereby enabling the multitude of specific communication approaches proposed in previous works within the spatial coverage and existing infrastructure of a city-wide, mobile contact network substrate engrained into the city operation and landscape.

B. Paper Structure

We first describe the data set used for our analysis and identify the factors that influence both the analysis and the generality of our results (Section II). We then analyze the feasibility and performance of network-wide communication in the underlying contact network (Section III). We then analyze the available support for directed communication (Section IV). Our evaluation shows the feasibility of communication networks within public transport systems (Section V), motivating our ongoing efforts of implementing a trial version in cooperation with the public transportation provider[3] and technical operator[4].

II. Evaluation Scenario and Characteristics

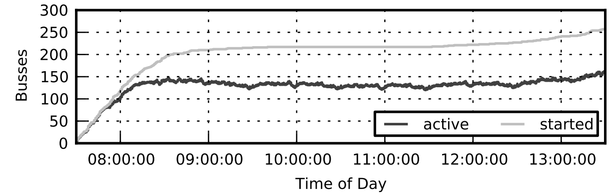
We build our evaluation on a data set describing the bus operation in the city of Aachen. The data set contains the number, assignment, and schedule of busses as well as their routes and the location of traversed bus stops. Hence, it allows us to analyze the number, duration, and location of contacts between busses and bus stops as the basis for the contact network and subsequently its communication characteristics.

While the data set lists bus schedules of the whole week, for now, we build our analysis within the time between 07:30 and 13:30 on Mondays. The reason for this restriction is the repetition of schedules on workdays and a significantly smaller set of operating busses on the weekend. Thus, to assess the possibilities of a city-wide network, we analyze a representative time frame in which "normal" bus operation can be observed.

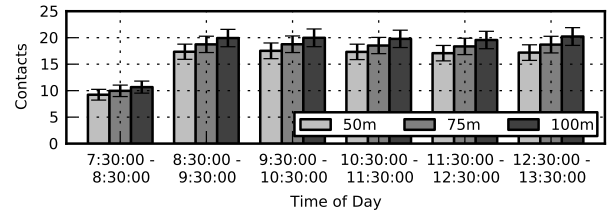
In our evaluation, we project the GPS coordinates of bus stops onto OSM cartographical material and simulate the path of each bus line along the streets given in the map and the respective schedule. From this, we obtain, with a granularity of 1 s, all active busses in the simulation, their velocity as

[3]Aachener Strassenbahn und Energieversorgungs-AG – http://www.aseag.de

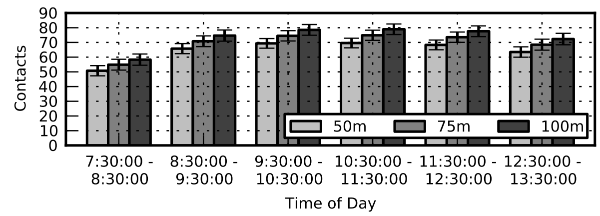
[4]IVU Traffic Technologies – http://www.ivu.de



(a) Cumulative number of started busses over the evaluation time and number of busses that are active at the respective point in time.



(b) Number of contacts between busses over time.



(c) Number of contacts between busses and bus stop infrastructure elements over time.

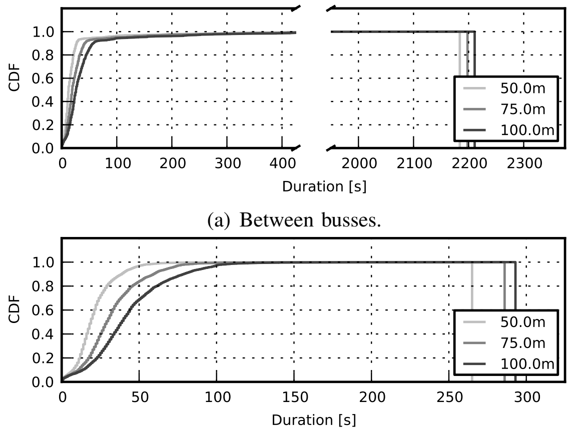
Fig. 2. Basic characteristics of the evaluated data set and resulting network.

derived by the street-wise distance between two stops, and the current position of each bus. To create a communication network, we assume all busses to be equipped with an 802.11 module. We do not make any assumption about the particular standard amendment, in order to derive a general understanding of the network. Moreover, we also evaluate the case of bus stops being equipped with 802.11 modules.

Fig. 1 shows the occurrences of bus locations reported during the simulation in the city center as a heatmap, highlighting the density of bus lines along streets or important places, with the most frequented place being the bus terminal. Please note that the heat map only shows the city center, i.e., an area of $2.12\,\mathrm{km}^2$ for visibility reasons, as occurrences decrease significantly with increasing distance to the bus terminal. In total, bus locations are spread out over an area of $906.72\,\mathrm{km}^2$.

In the regarded time frame, 259 individual busses frequent a total of 2111 bus stops. In detail, Fig. 2(a) shows the cumulative number of started busses and active busses over the complete evaluation time frame. Note that we only include busses that start within the evaluation time frame, while other busses may be active before the initial point in time. Over the main duration of the regarded time frame, about 145 busses are simultaneously active. Hence, as a first result, the bus network offers a large-scale network of mobile communication entities that may collect and aggregate data and transport messages over the, relatively confined, area of the city of Aachen.

A crucial performance characteristic of the resulting communication network is the number and duration of contacts between communicating entities. In this, we regard both contacts between busses (in a purely mobile network scenario as in [2], [6]) and contacts between busses and bus stops (i.e., in an



(a) Between busses.

(b) Between busses and bus stops.

Fig. 3. Cumulative distribution of contact durations.

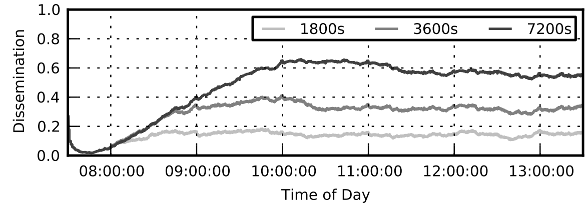


Fig. 4. Dissemination rate exclusively using busses as forwarders.

infrastructure-assisted metropolitan-area network). We justify this assumption with the fact that, in Aachen, a reasonable number of bus stops is already equipped with 802.11 modules. We hope that the data gathered in this paper ultimately serves as motivation to extend this coverage to even more bus stops, enabling a truly city-wide network.

Fig. 2(b) and 2(c) show the average and corresponding 95 % confidence intervals of the number of contacts between busses and between busses and bus stops, respectively. Please note that contacts only depend on whether the respective entities are in communication range, i.e., a bus does not have to actively stop at a bus stop for a contact to arise. Because the 802.11 communication range fluctuates in real-world scenarios, we further vary the assumed 802.11 range between 50 m and 100 m to investigate the impact of this major factor on the contact network. As Fig. 2(b) and 2(c) show, busses experience a large number of contacts throughout the evaluated time of operation. The larger number of contacts with bus stops stems from bus lines operating the outskirts of Aachen, which contribute positively to the bus stop count but, in comparison to inner city lines, negatively to the bus contact count. From this evaluation, we deduce a negligible impact of the assumed communication range. In contrast, the communication range shows a discernible impact when we regard the duration of contacts between busses and between busses and bus stops. Fig. 3(a) and 3(b) show the cumulative distribution of contact durations measured over the complete simulation. Note that we do not artificially model stop times in the simulations to avoid random assumptions. While the figures only show "drive-through" contact durations, any reasonable model of stop times could just be added to this durations. Contacts between busses (cf. Fig. 3(a)) on average last between 26 s (50 m range) and 44 s (100 m range) and show a significant impact of the assumed 802.11 communication range. Predominantly, both entities are mobile, i.e., distinct busses may follow the same route for a period of time, causing the long tail of the distribution that exhibits longer durations.

Intuitively, contact durations between stationary bus stops and mobile busses show a great impact of the assumed range (cf. Fig. 3(b)), given that the respective range exclusively determines whether a contact occurs. Because of this, contact durations are tightly limited. On average, contacts between busses and bus stops last between 21 s (50 m range) and 43 s (100 m range). As an overall result, the number of contacts and their respective durations indicate the feasibility of contact-driven wireless communication. For a back-of-the-envelope calculation, we pessimistically assume that wireless links require 3 s for discovery and establishment, even without the need for DHCP, and that links achieve 30 % of their designated throughput, i.e., 32.5 MBit/s for 802.11n (single spatial stream, 20 MHz channel, 800 ns guard interval). Busses are then able to transmit, within each contact, between 31.14 MB (50 m range) and 55.52 MB (100 m range) to other busses and between 24.37 MB (50 m range) and 54.16 MB (100 m range) to bus stops, on average. We argue that for the regarded applications this offers a viable communication substrate.

III. Evaluation: Network-wide Communication

The main evaluation focus of this paper is the feasibility of using the existing transportation network (and infrastructure) as a viable metropolitan-area communication network that enables smart city applications. In this section, we thus analyze the performance of establishing a store-and-forward network solely between busses (Section III-A). Subsequently, we build on the observation that multiple bus stops are already equipped with 802.11 capabilities and study the benefits of using bus stops as storage entities for bus-transferred data and as forwarding shortcuts in an Internet-attached backbone (Section III-B).

To model a communication scenario, we assume that passengers enter and leave busses at each bus stop. Each new passenger creates a broadcast message that is inserted into the communication network and has to be disseminated to all busses, i.e., transitively all passengers in the busses. Examples of broadcast communication are notifications about traffic incidents as well as offloading or prefetching approaches that aim to disseminate Internet content via the contact network [8]. For this evaluation we assume an 802.11 range of 50 m.

At each point in time, a number of user-generated messages thus exist in the network that, in principle, can reach all currently active busses. We thus measure the performance of the communication network as the network-wide content dissemination rate, i.e., the sum of messages that are present at active busses (the current state) over the product of active messages and active busses (the possible maximum). To represent different application scenarios, messages have a lifetime, after which they are discarded from the network. Examples for short-lived messages would be dissemination of a news item, while longer lifetimes could serve event announcements or traffic incidents.

A. Bus-driven Data Dissemination Performance

As a first step, we follow the epidemic routing evaluation of existing works [1], [6] within our scenario. In this setting, only busses communicate to disseminate messages among them.

Fig. 4 shows the dissemination rate over the full time frame with regard to message lifetimes of 1800 s, 3600 s, 7200 s. Intuitively, increasing message lifetimes benefits dissemination rates as more busses contacts occur within the lifetime. Note that, in comparison to the allowed delivery delays in [1], we set rather small message lifetimes. This is because, in the confined area of a city, we do not regard arbitrary delivery times or lifetimes to be meaningful.

Overall, bus-driven dissemination within the relatively short time frame of our evaluation scenario only yields satisfying results for long message lifetimes. However, the results factor in all busses spread out over the whole evaluation area. For messages with a short lifetime, a limited geographical dissemination, resulting in a limited dissemination rate in Fig. 4, can be acceptable. An example for such a scenario are the aforementioned traffic incident reports, which only carry meaning with the (limited) affected geographical area.

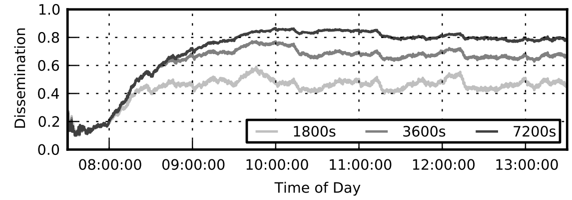
B. Infrastructure-assisted

Fig. 2(c) highlighted the fact that bus contacts with bus stops exceed contacts with other busses. This fact, along with the observation that bus stops are already equipped with 802.11 capabilities and 802.11 modules are available in small form factors and at manageable costs, motivates our aim of incorporating bus stops as communication entities in the metropolitan-area network. As such, we now evaluate the impact of leveraging bus stops in the communication network as storage entities (Section III-B1) and as members of an Internet-connected backbone that serves as a distribution shortcut between bus stops (Section III-B2).

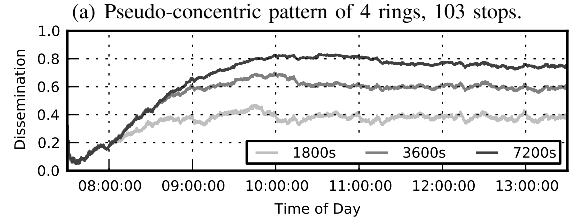
1) Autonomous: As a first step, we evaluated whether storing messages at 802.11-equipped bus stops benefits the overall dissemination, when busses synchronize messages with each encountered bus stop. Unfortunately, our evaluation showed no additional dissemination progress in this scenario compared to exclusive bus-to-bus communication. This is based on the fact that we measure the spreading of information both geographically and towards participating entities, a performance factor that stationary bus stops can not contribute much to. Still, 802.11-augmented bus stops would enable informational, location-based, or entertainment applications as well as storage capacities inside the network, among others.

2) Backbone-assisted: Extending the point of the previous section, geographical distribution and thereby dissemination to busses farther away can be achieved by establishing an Internet-connected backbone of bus stops that provides an immediate forwarding shortcut within the network. Information then becomes available at all included bus stops and can be exchanged with nearby busses quicker than relying solely on forwarding within contacts of mobile entities. The number of included bus stops presents a tradeoff between the achieved coverage and the associated installation and operation costs.

In this section, we thus evaluate the possibilities of assigning bus stops to be included in the backbone network. We regard one calculable geographical pattern, namely a ring pattern, and a manual selection based on the frequentation of bus stops. Fig. 5 shows the dissemination performance on top of each pattern for increasing message lifetimes.



(a) Pseudo-concentric pattern of 4 rings, 103 stops.



(b) Manual selection by bus stop frequentation, 43 stops.

Fig. 5. Dissemination performance per message lifetime with regard to the number and pattern of assigned bus stops.

In the ring pattern, bus stops are selected based on their distance to the geographical city center, creating four pseudo-concentric rings of increasing distance. On each ring, we require bus stops to be a minimum of 250 m apart from each other to ensure a good distribution. In total, we select 103 of 2111 bus stops (i.e., about 5 %), increasing the dissemination rate by at most 40 %, 35 %, and 20 % for lifetimes of 1800 s, 3600 s, and 7200 s, respectively (cf. Fig. 5(a)), compared to bus-only dissemination (cf. Fig. 4). Intuitively, short-lived messages especially benefit from this shortcut forwarding, enabling farther distribution in the available time frame.

Contrary to the ring pattern, the frequentation of bus stops indicates which of the latter more efficiently serve as hot spots in the communication network. We therefore selected 43 bus stops based on their position and frequentation. We include the position because, in the city center, multiple highly frequented bus stops are very close to each other and thus an assignment solely based on frequentation would result in redundant assignments. As Fig. 5(b) shows, leveraging such domain knowledge in the selection provides a performance that is within 10 % of the geographical ring pattern, while requiring 60 % less infrastructure support.

From this evaluation, we conclude the feasibility of establishing a metropolitan-area communication network that leverages the existing mobile and stationary infrastructure of public bus transportation system. Moreover, the system operator is able to boost the communication performance of the network by installing an Internet backbone according to the possible financial dimensions. Last, restricting the regarded area of communication to the inner city, i.e., within a distance of 1.5 km to the city center, may be an attractive option of establishing a more geographically focused city network. Fig. 6 shows the dissemination rate for an Internet backbone of 13 selected bus stops within such a concentrated deployment. Avoiding the low frequentation and lower population density of outer city areas allows for a relatively high-performance network with moderate infrastructure requirements.

IV. EVALUATION: LOCATION-CENTRIC COMMUNICATION

Undirected communication, as evaluated in the previous section, serves applications that benefit from a maximal set

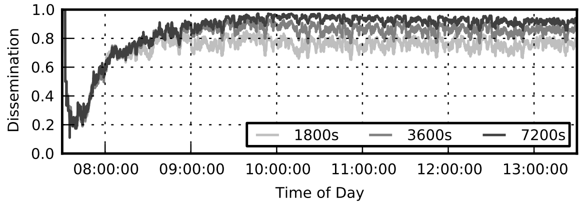


Fig. 6. Dissemination confined to the city center with 13 selected bus stops.

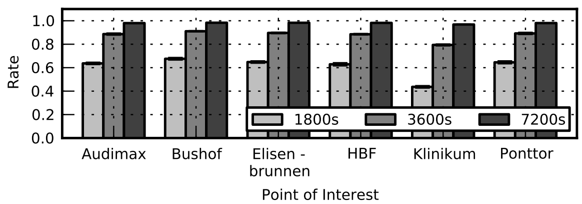


Fig. 7. Delivery rate of directed geographical communication per POI over message lifetime.

of recipients. In contrast to this, we envision smart city applications that allow information to originate at any point in the network, i.e., the city, but require information to be delivered to designated locations, according to the application semantics. Examples are location-based gaming and services as well as participatory and city-wide mobile sensing. In the former, user messages may only have meaning at their target locations, while in the latter collected data needs to be delivered to a designated processing entity.

Therefore, we evaluate the performance of the communication network in delivering messages to a designated location. Messages are created at random points in space with sufficient time left in the simulation to reach their destination before their lifetime expires. We choose six points of interest in Aachen as destinations (cf. Fig. 7). Of these, the Elisenbrunnen and Bushof are located in the city center and have touristic relevance while the Audimax and Ponttor are located within 1 km of the center and are frequented by students. The train station (HBF) is located about 1.3 km away from the center and may serve as an ingress/egress point for store-and-forward communication. Last, the University hospital (Klinikum) is more than 3 km away from the center. Please note that in this evaluation, forwarding occurs exclusively between busses.

As Fig. 7 shows, all inner city destinations allow delivery rates between 60 % and 100 %, depending on their lifetime. This is because the vast majority of busses commutes between the city and the outer regions of Aachen, with outbound busses meeting inbound busses regularly. For the low message lifetime, delivery rates for the University hospital drop significantly in comparison, due to the more isolated location, while higher lifetimes allow comparable delivery rates.

From this evaluation, we deduce that directed communication within the communication network can support the requirements of smart city applications. Further strategies of directed communications include semantic route and destination labels for busses and bus stops as well as users publishing interest labels when boarding a bus, with busses exposing these labels in contacts, supporting communication between busses based on passenger interests.

V. Conclusion

In this paper, we analyzed the possibility of realizing a city-wide communication network on top of existing public transport network infrastructure. Different than related works, we conduct the analysis from a provider perspective, compliant to the notion of establishing a cooperative network basis for smart city applications. Our evaluation of the transport network in Aachen, representative for middle-sized european cities, shows that the resulting network can support city-wide communication as well as directed communication. Two deciding factors influence the performance and characteristics of the resulting network, namely the time requirements of applications and the degree of augmenting the network with an Internet-bound backbone network at bus stops. Specifically, solely leveraging busses and bus stops for communication limits the temporal spreading of messages, while an appropriately placed backbone network offers instant and geographically widespread dissemination.

A. Future Work

As a next step, we will cooperate with the provider of the local transportation network as well as the technical operator in equipping busses with 802.11 modules for a real-world evaluation of the simulated scenario. Concurrently, we aim to realize prototype applications in the aforementioned usage scenarios of mobile sensing, especially in leveraging the spatial and temporal coverage for environmental monitoring [5].

Last, trading data rates for transmission range, the 802.11ah [9] standard offers an upcoming technology that might extend the communication range of entities in the network. Surpassing 802.11 communication ranges, sub-GHz transmissions offer better penetration, to be used for sensing, discovery, connection negotiation, and low-volume emergency channels. Prior to entities entering 802.11g/n ranges, 802.11ah channels might prepare usage of the high-volume 802.11 link.

Acknowledgment

This work has been funded by the DFG Cluster of Excellence on Ultra High-Speed Mobile Information and Communication (UMIC) and the DFG SFB 1053 "MAKI".

References

[1] X. Zhang *et al.*, "Study of a bus-based disruption-tolerant network: Mobility modeling and impact on routing," in *MobiCom*, 2007.

[2] M. Doering *et al.*, "A new mobility trace for realistic large-scale simulation of bus-based dtns," in *CHANTS*, 2010.

[3] R. Chandra *et al.*, "A campus-wide testbed over the tv white spaces," *SIGMOBILE MCR*, vol. 15, no. 3, 2011.

[4] Microsoft Research, "Networking Over White Spaces (KNOWS) Deployments," http://research.microsoft.com/en-us/projects/KNOWS/deployment.aspx.

[5] M. Mun *et al.*, "Peir, the personal environmental impact report, as a platform for participatory sensing systems research," in *MobiSys*, 2009.

[6] M. Doering, T. Pögel, and L. Wolf, "Dtn routing in urban public transport systems," in *CHANTS*, 2010.

[7] D. Naboulsi and M. Fiore, "On the instantaneous topology of a large-scale urban vehicular network: The cologne case," in *MobiHoc*, 2013.

[8] Y. Li *et al.*, "Multiple mobile data offloading through delay tolerant networks," in *CHANTS*, 2011.

[9] S. Aust, R. Prasad, and I. G. M. M. Niemegeers, "IEEE 802.11ah: Advantages in standards and further challenges for sub 1 GHz Wi-Fi," in *ICC*, 2012.