Analyzing Temporal Metrics of Public Transportation for Designing Scalable Delay-Tolerant Networks

Adriano Galati, Vladimir Vukadinovic, Maria Olivares, Stefan Mangold
Disney Research
Stampfenbachstrasse 48
8092 Zurich, Switzerland

ABSTRACT

Delay-tolerant networks can complement cellular networks to address today's growth in demand for wireless data. We are interested in delay-tolerant networks for reaching out into underserved regions in growing economies, when distributing media and videos from cities to rural areas. To transport the data into these regions, public transportation vehicles equipped with wireless infostations are used instead of a traditional cellular infrastructure. We focus on media distribution in support of entrepreneurs that operate mobile cinemas in rural villages. Temporal metrics of a public transportation network are determined, together with the overall value of exploiting the mobility of participating vehicles. A dense transportation network of Zurich (Switzerland), which is known to operate reliably, is analyzed for the purpose of benchmarking. Nodes are assumed to be bus stations that are temporarily in contact with each other when a bus travels between them. Our analysis provides guidelines to design and build a hierarchical topological structure of similar networks.

Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: Wireless communication; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*; C.4 [**Performance of Systems**]: Modeling techniques

General Terms

Measurement, Experimentation

Keywords

Temporal Metrics; Measurement; Experimentation; Delay Tolerant Networks; Media Distribution

PM2HW2N'13, November 3–8, 2013, Barcelona, Spain.
Copyright 2013 ACM 978-1-4503-2371-0/13/11 ...$15.00.
http://dx.doi.org/10.1145/2512840.2512846.

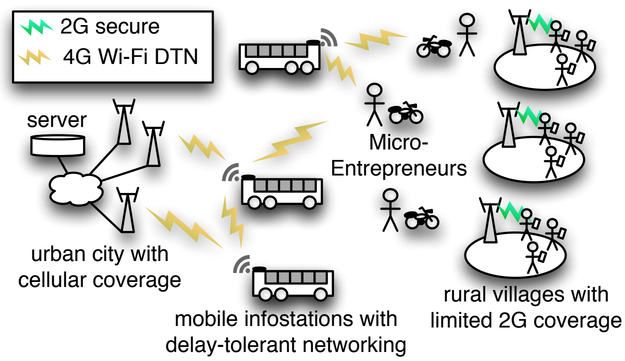

Figure 1: Media distribution with infostations: A mobile infostation is operated with the help of buses that travel to rural villages. There, the mobile cinema receive updates while in wireless range of mobile infostations.

1. INTRODUCTION

Mobile voice, data, and in particular, video services are becoming an essential part of consumers' lives. The number of mobile users continues to grow, and capacity demand due to mobile data and video distribution is increasing [1]. It is known that scaling up today's cellular networks towards the required capacity will be cost intensive [2]. While voice and low bit-rate data connectivity (GSM or IS-95 2nd Generation (2G) voice, or GPRS data (2.5G)) are available nearly everywhere across the world, mobile Internet access (3rd Generation, 3G, LTE and beyond) is often expensive or unavailable in rural regions. We want to reach out into these regions with a novel low-cost approach for media distribution. It is our overall objective to enable scalable network access and a global footprint of video services, everywhere in the world. The ultimate goal is to develop and test a framework that uses Delay-Tolerant Networks (DTNs) to unlock new mobile business opportunities, with particular regard to rural areas. DTNs can provide a viable low-cost alternative to cellular wireless communication networks in areas that are underserved [3]. There is growing interest in opportunistic networks, as a special case of DTNs (in which routes are established only if nodes are in reach of each other), because of their ability to provide flexible multi-hop connectivity by leveraging intermittent contacts among mobile radio devices. Opportunistic networking uses a paradigm where data is delivered from a source to a destination through the physical mobility of intermediate nodes. Such nodes communicate only with their direct neighbors using local radio communi-

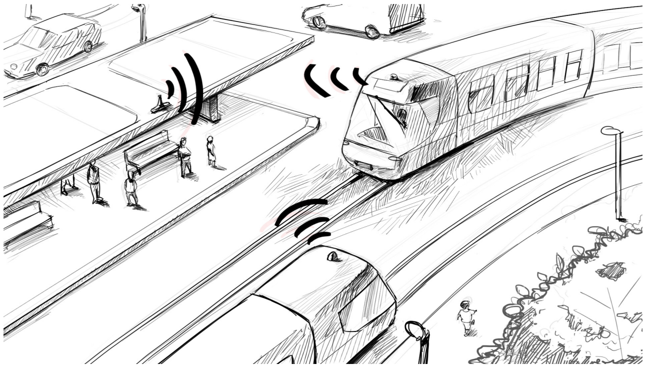

Figure 2: Media distribution through DTN in a transportation network with regular traffic. Other ways of distributing media are used in city centers (WLAN, xDSL), but the analysis is useful for comparison and establishing design guidelines (© Disney).

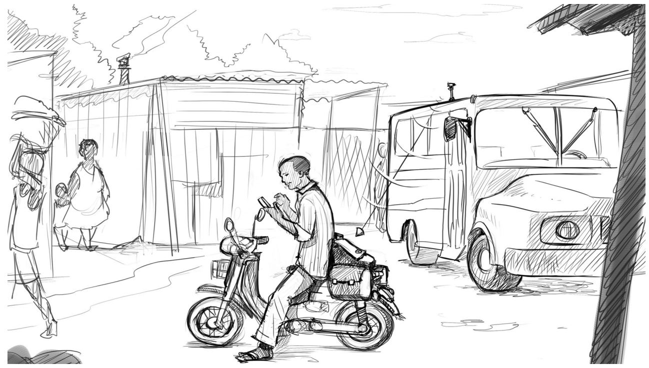

Figure 3: The DTN in rural areas employs infostations on bus stations and buses that commute from a city to villages. Mobile cinemas operated by local entrepreneurs rely on this low-cost approach for media distribution (© Disney).

cation like Wireless Local Area Networks (WLANs). When nodes move, and with enough redundant communication, data eventually reaches the destination. It is known that DTNs require only minimal (if any) investment in supporting infrastructure (e.g., infostations). However, DTNs compromise network reliability and the quality of service by introducing unpredictable delays, because data is moved from one wireless relay to another in a store-carry-forward approach. Applications must be able to tolerate the delays to avail from DTNs.

We model a bus-station network in which buses work as message ferries delivering data exclusively between two bus stops (a bus stop is also referred to as bus station, or station, in the following). Figure 2 illustrates such a scenario where buses and stations are assumed to be equipped with infostations. Such a representation of a network is different from those used in previous related research, which assume buses to be network nodes and able to exchange data with each other when in radio range. In our approach, bus stops are assumed to be nodes of a graph. Edges of the graph appear when a bus is in motion between two nodes at a given time: Two nodes are in contact when a bus travels between them. A node with higher geodesic locality and connection rate as other nodes is more important because it can forward data more rapidly to a larger number of other destinations. However, analyzing and understanding such an environment is difficult as it is inherently dynamic, with communication opportunities changing over time.

Time-varying graph metrics (as proposed in [4, 5]) have, so far, found little use in the DTN field. Such metrics allow the identification of key nodes that have higher capacity to transfer data towards other nodes through time. Temporal analysis provides understanding of network dynamics and a more accurate classification of important nodes compared to traditional metrics such as inter-contact time and contact duration [6–9] that ignore time ordering of contacts. Such important nodes, once identified, will create a backbone when designing an opportunistic communication network architecture. At this stage we are not analyzing rural area scenarios. Instead we focus on the public transportation in the city of Zurich using temporal network metrics. The approaches adopted and the experience gained through the analysis of the urban public transportation system will be used for future rural scenarios. Our analysis and classification provides us with guidelines to design, building and maintaining a hierarchical topological structure of the network as well as certain quantification degrees which identify the nature of network nodes and bus lines. For example, they may be identified and classified as primary bus lines which might act as DTN network backbones, and secondary bus lines which might represent sub-network domains.

In the following, we will first introduce the mobile cinema application in Section 2, and the scenarios we are considering. A detailed study of temporal metrics of the Zurich public transportation will be presented in the Section 3 and 4. Section 5 is a first attempt to create general design guidelines for network planning of delay tolerant networks, with our target application (mobile cinemas for rural areas) in mind. Section 6 summarizes related work, and Section 7 concludes the paper and points out future works.

2. MOBILE CINEMA SCENARIO

Our overall goal is to develop and test a new framework that uses DTNs to unlock new mobile business opportunities, especially in rural areas. Mobile cinema entertainment, possibly combined with educational content for healthcare, will be the use case. Content will be delivered with the help of mobile infostations and opportunistic networking. We are interested in equipping micro-entrepreneurs with the required tools and skills to launch their own mobile cinema businesses in their respective communities.

Instead of relying on cellular data, we use a DTN approach to provide affordable technical ways for the micro-entrepreneurs to obtain and distribute multimedia cinema contents, and to run their business at low cost. We are interested in exploring if such technology platforms introduced in rural and urban areas can serve as opportunity for people to establish their own businesses. In its simplest form, our multimedia content delivery mechanism for mobile cinema assumes two-hop communication between the multimedia server, which is located somewhere in the Internet cloud, and the mobile cinema device carried by the micro-entrepreneur. In a more flexible media distribution approach, mobile infostations help broadcasting the content. Figure 1 illustrates the DTN topology. First, the content is downloaded from the server to an infostation. An infostation is a battery-powered WLAN-enabled device placed on a bus, taxi. Second, the content is downloaded from the infostation to the WLAN-enabled mobile cinema device once

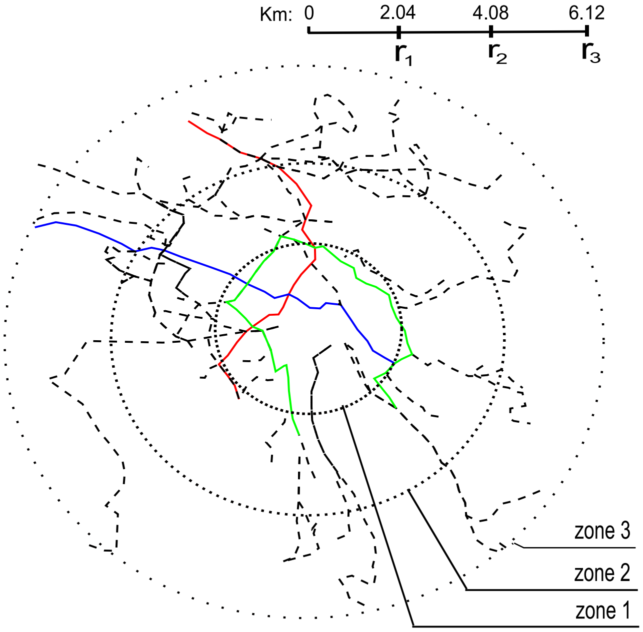

Figure 4: Public bus transportation plan for Zurich (most relevant bus lines only).

the vehicle arrives to the rural area. Hence, the mobile infostation serves as an intermediate relay (referred to as data mule) that carries the content between the server and the mobile cinema device. It is owned by one or multiple micro-entrepreneurs that reside on the same bus route. Commercial off-the-shelf components should be used to build the WLAN enabled mobile cinema device and infostation devices. The mobile cinema device includes a mobile media projector, a 2G cell phone, a WLAN mesh point with storage, a GPS receiver, and a rechargeable battery. Any fixed infostation located at a bus stop, or located inside a vehicle (public bus) will include a programmable WLAN mesh point, a storage device, and a rechargeable battery.

Figure 3 illustrates a scenario with a vehicular node serving as a local community member who requires some digital content for his micro-entrepreneur business. In this figure, a bus that follows regular mobility patterns between city centers carries an infostation that transmits data to a local device in the vicinity. Instead of relying on a network infrastructure, the local device receives data in a DTN manner; the communication relies on the bus traveling from one location to another. When mobile devices meet, they use the opportunity to exchange messages and data that is not necessarily intended to themselves. Instead, devices might receive data that they should carry along while traveling, until they meet another devices which creates another opportunity to communicate. In our example, the infostation mounted on the bus can be referred to as data mule. We aim to use the approaches adopted and experience gained through the urban scenario in the city of Zurich as targeted in this paper as pilot study for our future rural scenarios.

3. TIME-VARYING BUS NETWORK

Temporal graphs have recently been proposed [4,5,10–13] to model real time-varying networks, with the intuition that the behavior of dynamic networks can be more accurately captured by a sequence of snapshots of the network topology as it changes over time instead of using a single static graph with a fixed topology. Temporal analysis provides a better understanding of dynamic systems and a more accurate identification of important nodes compared to traditional static methods [4,5,11]. A directed temporal graph is defined as an ordered sequence of directed graphs.

3.1 Zurich Case Study Analysis

The environment we use for our initial analysis is not a rural area, instead we focus on the city of Zurich (Switzerland) and its public transportation system. Zurich is a city in Switzerland and its public transport system is known for its reputation for punctuality, synchronised timetables, efficiency, frequency and high quality of service and innovation. Within Zurich and throughout its canton, ZVV (Zurcher Verkehrsverbund), the Zurich transport network, has traffic density ratings among the highest worldwide [14]. Three means of mass-transit exist: the S-Bahn (local trains), trams, and buses. In this work we are interested in investigating only the bus transportation within Zurich which spans 44 lines over about 150 square kilometers and serving 336 bus stations. We retrieved bus timetables and other useful information of Zurich public transportation from the ZVV database. We extracted useful information among which arrival and departure times of each bus at each station which we are able to locate through its UTM (Universal Transverse Mercator) coordinates. As it can be observed in Figure 4, the overall city structure is not Manhattan-like. At a glance, three kind of bus lines can be spotted: two span across the city, one runs around the center, and the rest are mostly shorter, connected to the first two and covering small areas. This topological structure represents an interesting case study as the first three bus lines might represent network backbones interconnecting smaller bus network domains. We consider Zurich bus transportation as a sample large-scale bus network for our study with the intention to propose similar approach to public transportation in developing regions of growing economies. Because of its efficiency, the Zurich bus network also provides optimal estimates which might be reference values when analysing similar networks in different environments.

We model the Zurich bus network as a time-varying graph $\mathcal{Z}_t^w(t_{min}, t_{max})$, starting at t_{min}, ending at t_{max}, and defined as a sequence of graphs $\mathcal{Z}_{t_{min}}, \mathcal{Z}_{t_{min+w}}, \ldots, \mathcal{Z}_{t_{max}}$. The parameter w is the size of each time window expressed in minutes. The number of graphs in the sequence is denoted by

$$W = ((t_{max} - t_{min})/w) = |\mathcal{Z}_t^w(t_{min}, t_{max})|.$$

Here, nodes of a graph are bus stations (bus stops). All the temporal graphs have the same number of nodes, while each one may have a different number of edges. We represented a time-varying graph $\mathcal{Z}_t^w$ by means of a $N \times N$ time-dependent adjacency matrix $A(t)$, $t = 1, 2, \ldots, T$, where N is the number of bus stations and $a_{i,j}(t)$ identifies a possible connection between the stations i and j in the adjacency matrix of the t-th graph. We assume the station i to be in contact with the station j if a bus is traveling from j to i within the same time-slot. In this sense, a sequence of graphs is convenient to describe systems where each connection starts at a specific time and has a temporal duration.

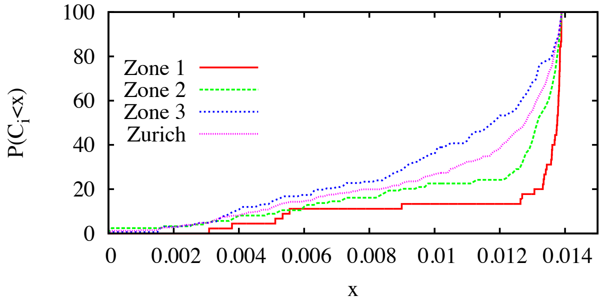

Figure 5: Cumulative distributions of the temporal correlation coefficients (C_i) of the bus stations during peak time in zone 1, zone 2, zone 3 and the whole Zurich.

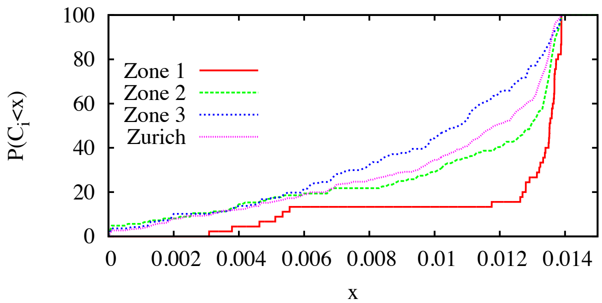

Figure 6: Cumulative distributions of the temporal correlation coefficients (C_i) of the bus stations during off-peak time in zone 1, zone 2, zone 3 and the whole Zurich.

In our analysis we also distinguish between peak and off-peak times. Peak time is a part of the day during which traffic congestion on roads and crowding on public transportation is at its highest. The frequency of public transportation is usually higher in the peak hours, and longer buses, or larger vehicles are often employed. Normally, it happens twice a day, once in the morning from 7:00 am to 10:00 am and once in the evening from 4:00 pm to 7:00 pm (Monday to Friday), the times when the most people commute. Instead, off-peak refers to all the other hours when fewer people travel. Here, we assume off-peak time to be from 10:00 am to 4:00 pm. Furthermore, we analyse the bus transport network within three different zones covering the entire city to measure network efficiency in different areas, from the city center to the suburbs. This might help choose proper locations where to connect to communication services as they might consider costs and time for traveling to the destination. Figure 4 shows the three zones: one covering the center of the city with radius r_1 and counting 45 bus stations (zone 1), and two concentric rings around it with radius $r_2 - r_1$ (zone 2), and $r_3 - r_2$ (zone 3) including 124 and 167 bus stations respectively.

3.2 Temporal Correlation Coefficient

We calculate the temporal-correlation coefficient [4] C_i of each bus station of the Zurich bus network $\mathcal{Z}_t^w$ and the three zones during peak C_{PT} and off-peak C_{OPT} time (see Figures 5 and 6). This provides us with an indication of the topological overlap degree of $\mathcal{Z}_t^w$ through time in two distinct parts of the day. We process the complete temporal graph during peak and off-peak hours with directed links, using windows of five minutes. We measure the average temporal correlation coefficient C of the neighbour set of a station i between two successive graphs of $\mathcal{Z}_t^w$ by deriving the temporal correlation coefficient of each bus station C_i:

$$C = \frac{\sum_i C_i}{N} \tag{1}$$

$$C_i = \frac{1}{T-1} \sum_{t=1}^{T-1} \frac{\sum_j a_{ij}(t)\, a_{ij}(t+1)}{\sqrt{\left[\sum_j a_{ij}(t)\right]\left[\sum_j a_{ij}(t+1)\right]}} \tag{2}$$

In particular, $C = 1$ if all graphs in the sequence are equal. Figures 5 and 6 plot the cumulative distributions of the temporal correlation coefficients C_i of the bus stations during peak and off-peak time in the zones 1, 2, 3 and in the whole city of Zurich. In both of the plots C_i decreases from zone 1, the city center, to zone 3, the suburbs of the city, while zone 2 tends to approximate the Zurich distribution. This behavior is confirmed by the average topological overlap values presented in Table 1. C_{PT} expresses slightly higher values than C_{OPT}. All of the three zones and the entire Zurich bus system express very similar values through peak and off-peak hours. This shows that the Zurich bus transportation and each zone express similar average topological overlap with a good approximation.

3.3 Temporal Path Length

The average temporal path length [4, 11] of a network quantifies how fast information spreads to all nodes by means of transitive connections between them. We briefly recall the definition of temporal path and shortest temporal distance over $\mathcal{Z}_t^w(t_{min}, t_{max})$ as a sequence of k hops via a distinct path. A temporal path $p_{ij}^h(t_{min}, t_{max})$ between two nodes i and j is defined as the set of paths starting from i and finishing at j that pass through the nodes $n_1^{t_1} \dots n_k^{t_k}$, where $t_{k-1} < t_k$ and $t_{min} < t_k < t_{max}$ is the time window that node n is visited. The shortest temporal distance between two nodes i and j of a generic graph $\mathcal{Z}_t^w$ is defined as the shortest temporal path length $d_{ij}^h(t_{min}, t_{max})$. Therefore, starting from time t_{min}, this is the number of time windows (or temporal hops) for information to spread from node i to node j. h indicates the maximum number of hops allowed within each window w. We have processed the complete temporal graph $\mathcal{Z}_t^w$ during peak and off-peak hours with directed links, using windows of five minutes and h=1. If a bus is traveling from station i to station j in a temporal window, a link between the two nodes will be added to the graph representing the temporal snapshot for that time. Generally, buses will take more than three minutes to travel between two bus stations. If a temporal path from i to j does not exist we set their distance $d_{ij} = T$, the longest temporal path. That is, we assume that information expires after a certain time period t_{max}. In order to compute the temporal distances d_{ij} for all node pairs $i, j = 1, 2, \dots, N$ of the Zurich bus network $\mathcal{Z}_t^w$, we have implemented a generalization of the breadth first search algorithm. We measure the average temporal connectivity properties of $\mathcal{Z}_t^w$ by deriving its temporal path length L:

$$L = \frac{1}{N(N-1)} \sum_{ij} d_{ij} \tag{3}$$

Low values of L indicate that the bus stations in $\mathcal{Z}_t^w$ can communicate more efficiently. Table 2 shows the average

Table 1: Average Temporal Correlation Coefficient (C)

	Zone 1	Zone 2	Zone 3	Zurich
C_{PT}	0.0125	0.0115	0.0103	0.0109
C_{OPT}	0.0122	0.0105	0.0093	0.0101

Table 2: Average Temporal Path Length (L)

	Zone 1	Zone 2	Zone 3	Zurich
L_{PT}	9.456	36.146	34.244	41.472
L_{OPT}	11.146	25.25	28.325	31.897

Table 3: Average Temporal Global Efficiency (E)

	Zone 1	Zone 2	Zone 3	Zurich
E_{PT}	0.172	0.087	0.067	0.053
E_{OPT}	0.15	0.08	0.07	0.055

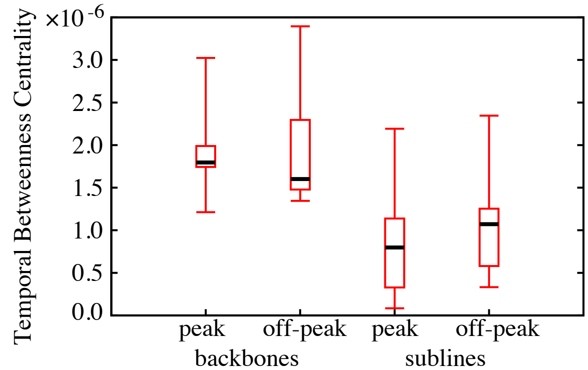

Figure 7: Temporal betweenness centrality of bus stations connecting backbones and sub-lines during peak and off-peak time.

temporal path lengths during peak-time (L_{PT}) and off-peak-time (L_{OPT}) hours for the zone 1, zone 2, zone 3 and the whole city of Zurich. Average path lengths in the city center are lower with respect to the other zones and time, in particular during peak hours. This means that bus stations in the center of Zurich can communicate more efficiently with respect to the others. Instead, the temporal path lengths in the zones 2, 3 and in the whole Zurich express smaller values during off-peak hours. This suggests that a delay tolerant service could exploit the bus network in the city center during the peak hours because their data can travel faster, or, eventually, from any zone during off-peak hours. In Table 2 the performance of the bus network decreases from zone 1 to Zurich. This implies that different zones have different speed to spread information on average.

3.4 Temporal Global Efficiency

The temporal global efficiency E [4, 11] captures the dynamics of the whole network, in particular how easy information flows from source to destination across the whole time space. The temporal network efficiency is derived from the temporal path length and avoids the potential divergence due to pairs of nodes that are not temporally connected. If a temporal path from i to j does not exist we set $1/d_{ij} = 0$ in Equation 4. Temporally disconnected node pairs have zero efficiency.

$$E = \frac{1}{N\,(N-1)} \sum_{ij} \frac{1}{d_{ij}} \tag{4}$$

The temporal global efficiency values extracted from $\mathcal{Z}_t^w$ are shown in Table 3. High values of E indicate that the nodes of the graphs can communicate more efficiently. The zone 1 expresses higher efficiency and in particular during peak-time. Notice that the zone 1, 2, 3 and the whole Zurich show similar values during peak and off-peak hours while this is not the case for the temporal path lengths in Table 2. This is because pairs of nodes that are never in contact will not impact on the calculation as for the temporal path length.

4. TEMPORAL CENTRALITY METRICS

In this section we compute two temporal metrics, temporal betweenness centrality and temporal closeness centrality [5], which are metrics associated to the identification of central nodes in dynamic networks. More specifically, temporal betweenness distinguishes individuals who act as key mediators between the most communication paths over time, while temporal closeness quantifies how fast a node can disseminate information.

4.1 Temporal Betweenness Centrality

The temporal betweenness centrality [5] of a node i is defined as the ratio between all the temporal shortest paths that pass through i and all the shortest paths between all of the pairs within a certain time. It takes into account not only the number of shortest paths which pass through a node, but also the length of time for which a node along the shortest path retains a message before forwarding it to the next node. The temporal betweenness centrality C_i^B of a node i over the entire temporal graph $\mathcal{Z}_t^w$ is:

$$C_i^B = \frac{1}{W} \sum_{t=1}^{W} C_i^B\left((t \times w) + t_{min}\right) \tag{5}$$

and $C_i^B\,(t)$ of node i at time t is defined as:

$$C_i^B\,(t) = \frac{1}{(N-1)\,(N-2)} \sum_{\substack{j \in V \\ j \neq i}} \sum_{\substack{k \in V \\ k \neq i \\ k \neq j}} \frac{U\,(i,t,j,k)}{|S_{jk}^h|} \tag{6}$$

The function $U\,(i,t,j,k)$ returns the number of shortest temporal paths from j to k in which node i has received a message at time t. S_{jk}^h is the set of shortest temporal paths between j and k with h hops per time window. Here, we compute the betweenness values assuming $h = 2$ with a time window $w = 300\,s$ (five minutes). If $S_{jk}^h = \emptyset$ (empty set), station i is totally isolated and we set $C_i^B\,(t) = 0$.

The plot in Figure 7 shows the temporal betweenness centrality of bus stations belonging to backbones and sub-lines. We assume the bus lines 31, 32 and 33 being network backbones through the whole Zurich public transport system and all of the others belonging to sub-busnetworks (see Figure 8). Bus stations belonging to the backbone lines express higher median values with respect to sub-lines and in particular during peak time. This shows that such backbones can act as key mediators between sub-networks over time. They might serve as gateways to route data between the three main network backbones and the sub-networks, namely the bus lines 31, 32, 33 and all the others, or eventually to keep data traffic restricted within the sub-network domains (see Figure 8).

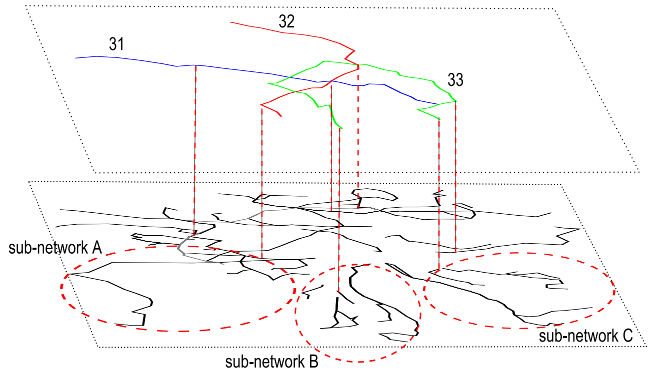

Figure 8: Clusters within public bus transportation in Zurich.

4.2 Temporal Closeness Centrality

In this section we analyse the temporal closeness centrality [5] of each bus station. Two bus stations are close to each other if their geodesic distance is small in their temporal graph $\mathcal{Z}_t^w$. It is a measure of how early data traveling from a source station can be delivered to all of the other stations. Given the shortest temporal distance d_{ij}, the temporal closeness centrality is expressed as:

$$C_i^h = \frac{1}{W\,(N-1)} \sum_{t \neq 1 \in V} d_{i,j}^h \tag{7}$$

In contrast to the previous indicators, a DTN operates more scalable and more resource efficient if the temporal closeness centrality is small. Bus stations that have, on average, shortest temporal distances to the other stations are considered more central. The two plots in Figure 9 and 10 show the cumulative distributions of the temporal closeness centrality for the peak and off-peak time respectively, with a time window $w = 300\,s$ (five minutes) and $h = 1$. In both plots the cumulative distributions present increasing temporal closeness centrality in the sequence from zone 1 to Zurich. This implies that bus stations in zone 1 can disseminate information faster with respect to stations in other parts of the city. Bus stations in zone 1 are key nodes and play an important role in forwarding data through the network.

5. DISCUSSION AND GUIDELINES

So far, the bus transportation of the city of Zurich clustered in three different concentric zones were analyzed. The primary bus lines, which may act as DTN network backbones, and potential secondary bus lines, which may be clustered within sub-busnetworks, were observed in detail. We also took into account two times of the day which might capture different behaviour of bus networks, peak time when the most people commute which happens twice a day, from 7:00am to 10:00am and from 4:00pm to 7:00pm, and off-peak time from 10:00am to 4:00pm when fewer people travel. Even though the extracted distributions show better values during peak time, they express little time dependence during the daytime. From our results, different zones have different speed to spread information; zone 1 can disseminate data more efficiently than the others, and so on from the centre to the outskirts of the city with decreasing performance. Although the distinction of connectivity in such locations of the city is somewhat obvious, we present evidence and some

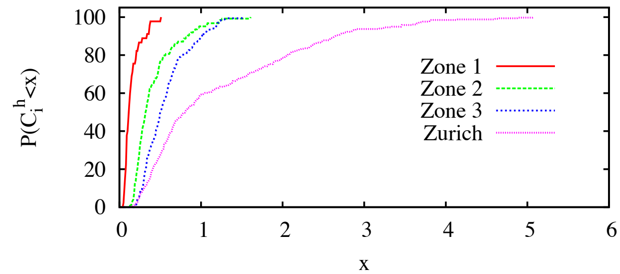

Figure 9: Cumulative distributions of temporal closeness centrality (C_i^h) of the bus stations during peak time in zone 1, zone 2, zone 3 and the whole Zurich.

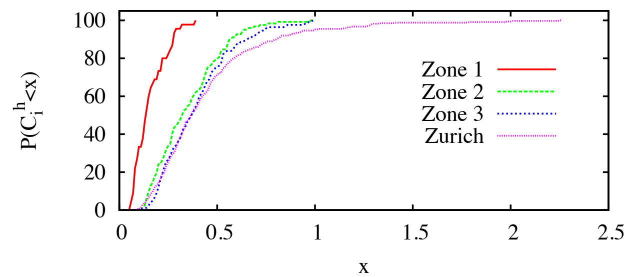

Figure 10: Cumulative distributions of temporal closeness centrality (C_i^h) of the bus stations during off-peak time in zone 1, zone 2, zone 3 and the whole Zurich.

quantification of the way in which they differ. For each zone we calculate the temporal correlation coefficient which indicates the topological overlap degree of the bus network through time in two distinct parts of the day. We calculate the temporal path length and global efficiency to estimate how easy information flows from source to destination across the whole time space. Low values of L (high values of E) indicate that the bus stations can communicate more efficiently. We also calculate temporal betweenness of each bus station to identify the ones that act as key mediators between the most communication paths over time. This shows that bus lines 31, 32 and 33 act as network backbones and key mediators between sub-busnetworks. Their bus stations might be employed as gateways to route data between the sub-busnetworks and to control forwarding of data in the entire network. Finally, we measure how early data from a source station can be delivered to all of the other stations by calculating the temporal closeness centrality. Namely, bus stations that have shortest temporal distance to the other stations have higher centrality.

From such results, micro-entrepreneurs can make their decisions on where to go (to which bus station) to download/upload a movie based on the zone, sub-busnetwork domain, network backbone, his available time and cost of the service, just to mention a few. For example, our results show that the Zurich bus network would disseminate data more efficiently from bus stations in zone 1 during off-peak time than from zone 2 during peak time. Communication services might require specific networking protocols depending on the zone in which they are used. The observed distributions suggest that forwarding to network backbones, during peak time and from the city centre, might increase significantly the likelihood of timely contact.

6. RELATED WORK

Considerable work has been devoted by the scientific community to devise reliable routing strategies in DTNs [15–19]. As special cases, opportunistic networks exploiting contact opportunities between nodes have been proposed [6–9, 20–22]. Comprehensive overviews of technical problems in supporting opportunistic DTN communication can be found in [15, 23, 24]. In this context, wireless networks exploiting public transport systems have been attracting attention in recent years. Initial work focusing on rural environments in developing regions where buses connect a number of villages spread over a large area [25–27]. Their common goal is to provide network access for delay-tolerant applications such as e-mail and non-real time web browsing. DakNet [25] uses computers with a disk and WLAN radio attached to buses on a bus route between villages. E-mails and other data are downloaded to the village and uploaded for transport to the Internet or to other villages along the bus route. On the same bus network, a system of throwboxes [28, 29] was deployed to enhance the capacity of the DTN. KioskNe [30] is also a network of rural Internet kiosks that provide data services in remote regions. Vehicles with on-board computers ferry the data between the kiosks and gateways connected to the Internet. In these cases the set of neighbors for every node is usually small and does not change frequently over time; usually encounters are highly predictable. TACO-DTN [31] is a content-based dissemination system composed of fixed and mobile infostations that allow mobile users to subscribe to certain contents for a period of time. Campus bus networks designed to serve students and faculties who commute between colleges or from/to nearby towns are proposed in [32–34]. In these settings opportunistic networks are usually characterized by a relatively small number of nodes when compared to a fully fledged urban environment. Scaling up in terms of number of nodes, urban environments generally offer a considerable number of bus lines, densely deployed, to enable people to commute inside a city. Bus networks in urban environments are usually characterized by many contact opportunities and frequent contacts [35–38]. In [35] the authors propose a commercial application based on a multi-tier wireless ad-hoc network called Ad Hoc City. It provides Internet access by means of WLAN Access points responsible for a geographical area. Using the same real data set, [36] propose a cluster-based multi-copy routing algorithm for intra-city message delivery. Here nodes are clustered based on their encounter frequency. To reduce the overhead effect of multiple copies, [37] propose an optimal stopping rule when forwarding. In [38] the public transport system of Shanghai is used to test the performance of a single-copy probabilistic forwarding mechanism. A recent work about performance analysis for deployment at urban scale is presented in [39]. In this work, the authors analysed inter-contact times of the Zurich and Amsterdam transport systems discovering that they follow an exponential distribution. Based on their findings, they are able to predict the performance of the epidemic routing protocol using a Markov chain model.

7. CONCLUSION

Delay tolerant and opportunistic networks could be useful in a number of environments if efficient forwarding algorithms are designed thoroughly. Temporal complex network metrics were used to characterize the Zurich public transportation network, and to identify stations and bus lines that can play a key role in an efficient urban DTN network. The approach taken is an example to help addressing the problem of scalability in a real city environment. We have presented evaluation results from temporal metrics of the public transportation. We have also presented evidence of distinct connectivity of bus stations in different zones of the city, and during peak and off-peak hours. Such measurements will be considered in the future to design an efficient urban DTN network. Besides, such a scenario provides us with a reference system for the analysis of other scenarios such as rural areas. The identification of such transport system's components can help in forwarding data: The various temporal metrics of the Zurich public transportation might be used to design improved forwarding algorithms. By analyzing the time-varying topology of a real urban setting we have identified network nodes that play key roles in building an efficient scalable system. For this purpose, the position of a node with respect to others can be classified and exploited. A node with higher geodesic locality and encounter rate to other nodes could forward data quickly to high numbers of nodes. We propose to exploit the three main components to route data: backbone bus lines, sub-networks and gateway-stations (see Figure 8). In future work, we intend to perform a similar analysis for public transports in other cities and surrounding rural areas, with focus on growing economies. Our final goal is to be able to assess the potential of a low-cost media distribution and content delivery system that will not rely on cellular networks and their infrastructure, nor on physical media such as flash memory or DVDs, but on delay-tolerant communication. As a first step towards such an approach, our analysis provides initial guidelines to design an efficient scalable network based on public transportation.

8. REFERENCES

[1] Cisco. Global Mobile Data Traffic Forecast Update 2012-2017, Visual Networking Index, 2013.

[2] A. Qureshi, J. Carlisle, and J. Guttag. Tavarua: Video Streaming with WWAN Striping. *in ACM Multimedia 2006*, October 2006.

[3] Vladimir Vukadinovic, Fabian Dreier, and Stefan Mangold. Impact of human mobility on wireless ad hoc networking in entertainment parks. *Ad Hoc Networks*, (0), 2012.

[4] J. Tang, S. Scellato, M. Musolesi, C. Mascolo, and V. Latora. Small-world Behavior in Time-varying Graphs. *Physical Review E*, 81(5), May 2010.

[5] C. Mascolo V. Latora J. Tang, M. Musolesi. Analysing Information Flows and Key Mediators through Temporal Centrality Metrics. *SNS '10*, April 2010.

[6] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott. Impact of human mobility on opportunistic forwarding algorithms. *IEEE Trans. Mob. Comput.*, 6, 2007.

[7] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott. Pocket switched networks: Real-world mobility and its consequences for opportunistic forwarding. *Technical Report UCAM-CL-TR-617, University of Cambridge*, February 2005.

[8] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, and C. Diot. Pocket switched networks and human mobility in conference environments. *ACM SIGCOMM*, 2005.

[9] J. Leguay, A. Lindgren, J. Scott, T. Friedman, and J. Crowcroft. Opportunistic content distribution in an urban setting. *SIGCOMM workshop on challenged networks, CHANTS '06*, 2006.

[10] A. Casteigts, P. Flocchini, W. Quattrociocchi, and N. Santoro. Time-Varying Graphs and Dynamic Networks. *CoRR*, abs/1012.0009, 2010.

[11] John Tang, Mirco Musolesi, Cecilia Mascolo, and Vito Latora. Temporal distance metrics for social network analysis. In *2nd ACM workshop on Online social networks*, WOSN '09, New York, NY, USA, 2009. ACM.

[12] J. Tang, C. Mascolo, M. Musolesi, and V. Latora. Exploiting temporal complex network metrics in mobile malware containment. In *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2011 IEEE International Symposium on a*, 2011.

[13] J. Tang, M. Musolesi, C. Mascolo, and V. Latora. Characterising Temporal Distance and Reachability in Mobile and Online Social Networks. *WOSN '09*, 2009.

[14] Zurich Transport Network ZVV. [online]: http://www.zvv.ch/en, June 2013.

[15] S. Jain, K. Fall, and R. Patra. Routing in a delay tolerant network. *ACM SIGCOMM*, 2004.

[16] Q. Yuan, I. Cardei, and J. Wu. Predict and relay: an efficient routing in disruption-tolerant networks. *ACM MobiHoc*, 2009.

[17] S. Nelson, M. Bakht, and R. Kravets. Encounter-based routing in DTNs. *IEEE INFOCOM 2009*, April 2009.

[18] A. Orda and R. Rom. Shortest-path and minimum-delay algorithms in networks with time-dependent edge-length. *Journal of ACM*, 37, July 1990.

[19] S. Merugu, M. Ammar, and E. Zegura. Routing in space and time in networks with predictable mobility. *Technical report, School of Computer Science, Georgia Tech, Atlanta, USA*, 2004.

[20] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, and C. Diot. Pocket switched networks and human mobility in conference environments. *2005 ACM SIGCOMM workshop on Delay-tolerant networking (ACM)*, 2005.

[21] H. Cai and D. Y. Eun. Aging rules: what does the past tell about the future in mobile ad-hoc networks? *ACM MobiHoc*, 2009.

[22] P. Hui, J. Crowcroft, and E. Yoneki. Bubble rap: social-based forward-ing in delay tolerant networks. *ACM*, 2008.

[23] L. Pelusi, A. Passarella, and M. Conti. Opportunistic Networking: Data Forwarding in Disconnected Mobile Ad Hoc Networks. *IEEE Communications Magazine*, 44(11), 2006.

[24] E. M. Daly and M. Haahr. The Challenges of Disconnected Delay-Tolerant MANETs. *Elsevier Ad Hoc Networks Journal*, 8(2), 2010.

[25] A. Pentland, R. Fletcher, and A. Hasson. DakNet: Rethinking Connectivity in Developing Nations. *IEEE Computer*, 37(1), January 2004.

[26] C. T. De Oliveira, R. B. Braga, D. M. Taveira, N. C. Fern, and O. C. M. B. Duarte. A predicted-contact routing scheme for Brazilian rural networks. *in Electrical Engineering Program, COPPE/UFRJ*, 2008.

[27] M. Demmer and K. Fall. DTLSR: delay tolerant routing for developing regions. *ACM NSDR 2007 workshop on Networked systems for developing regions*, 2007.

[28] Wenrui Zhao, Yang Chen, M. Ammar, Mark Corner, B. Levine, and E. Zegura. Capacity enhancement using throwboxes in DTNs. In *Mobile Adhoc and Sensor Systems (MASS), 2006 IEEE International Conference on*, 2006.

[29] N. Banerjee, M. D. Corner, and B. N. Levine. Design and Field Experimentation of an Energy-Efficient Architecture for DTN Throwboxes. *ACM Transactions on Networking*, 18(2), April 2010.

[30] S. Guo, M.H. Falaki, E.A. Oliver, S. Ur Rahman, A. Seth, M.A. Zaharia, and S. Keshav. Very Low-Cost Internet Access Using KioskNet. *ACM Computer Communication Review*, 37(5), 2007.

[31] G. Sollazzo, M. Musolesi, and C. Mascolo. TACO-DTN: A Time-aware Content-based Dissemination System for Delay-Tolerant Networks. *ACM Int. Workshop Mobile Opportunistic Networking (MobiOpp)*, 2007.

[32] J. Burgess, B. Gallagher, D. Jensen, and B. N. Levine. Maxprop: Routing for vehicle-based disruption-tolerant networks. *IEEE INFOCOM 2006*, April 2006.

[33] X. Zhang, J. Kurose, B. N. Levine, D. Towsley, and H. Zhang. Study of a bus-based disruption-tolerant network: mobility modeling and impacton routing. *ACM MobiCom*, 2007.

[34] A. Balasubramanian, B. Levine, and A. Venkataramani. DTN routing as a resource allocation problem. *ACM SIGCOMM*, 2007.

[35] J. Jetcheva, Y.-C. Hu, S. PalChaudhuri, A. Saha, and D. Johnson. Design and evaluation of a metropolitan area multitier wireless ad hoc network architecture. *in Mobile Computing Systems and Applications*, October 2003.

[36] S. Ahmed and S. Kanhere. Cluster-based forwarding in delay tolerant public transport networks. *32nd IEEE Conference in Local Computer Networks, 2007*, October 2007.

[37] C. Liu and J. Wu. An optimal probabilistic forwarding protocolin delay tolerant networks. *ACM MobiHoc*, 2009.

[38] M. Sede, L. Xu, L. Da, W. Min-You, L. Minglu, and S. Wei. Routing in large-scale buses ad hoc networks. *Wireless Communications and Networking Conference*, 2008.

[39] L. D. Hartog, T. Spyropoulos, and F. Legendre. Using public transportation as a DTN backbone: Mobility properties and performance analysis. *Proceedings of AOC 2010*, 2010.