METhoD: a Framework for the Emulation of a Delay Tolerant Network Scenario for Media-Content Distribution in Under-Served Regions

Adriano Galati[†], Sandra Siby[*], Theodoros Bourchas[†], Maria Olivares[†], Stefan Mangold[‡] and Thomas R. Gross[*]
[*]Department of Computer Science, ETH Zurich, Switzerland
[†]Disney Research, Zurich, Switzerland
[‡]Lovefield Wireless, Liebefeld, Switzerland

***Abstract*—Wireless communication is a cost-effective method of providing access to information for users in developing economies. We are interested in Delay Tolerant Networking (DTN) for distributing multimedia content to micro-entrepreneurs in under-served rural areas of South Africa. Buses equipped with WLAN-enabled devices (infostations), providing DTN connectivity, will be used to ferry data between urban and rural areas. Before the actual deployment, rigorous experiments must be performed to study the performance of the proposed network design. Such an activity is time consuming and requires complex preparation. Simulators, while convenient, do not provide the most realistic results. Given the complexity of DTN testing, we try to find a middle ground approach by building a mobility emulator testbed, which we use to emulate the network scenario. In this chapter, we outline the design and implementation details of the network emulator. We then describe the experiments that were performed to study the impact of different network dynamics on content delivery and their results. The results from the experiments are used to improve our current network design.**

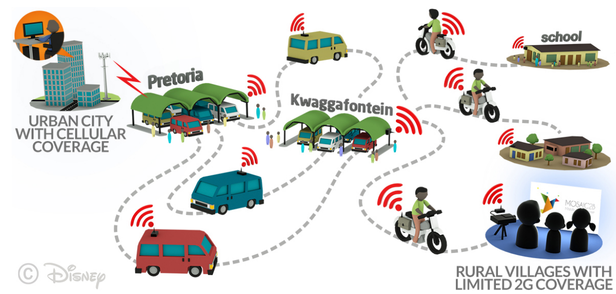

Fig. 1: The MOSAIC 2B scenario; a DTN network between the bus station in the city of Pretoria and rural bus station in Kwaggafontein.

I. Introduction

The Global ICT statistics of the year 2014 by the International Telecommunications Union indicate higher penetration rates for mobile cellular and broadband subscriptions in developing countries as compared to fixed, wired, subscriptions [1]. Access to information is essential for economic growth and development in these areas. The lower costs associated with wireless networks make them an attractive means of providing access to information for developing economies.

However, there are challenges that arise when deploying connectivity solutions in these areas; the technologies are usually built keeping the developed world in mind. They are based on the assumption that there will be reliable end-to-end paths, steady connections and power sources, which is not always the case in rural areas of developing countries [2].

Delay Tolerant Networking (DTN) is an architecture aimed at providing communication in situations where end-to-end connectivity is not possible. To ease the development process of DTN applications, the DTN Research Group (DTNRG) [3] has defined an experimental network protocol for challenged networks known as bundle. A bundle is a protocol data unit of the DTN bundle protocol. Bundles are transported in a store-carry-forward manner by the nodes in the network. The layer that implements the bundle protocol operations acts as an overlay network over existing transport layers. The bundle protocol describes all the entities and operations of the bundle layer such as the description of the bundle, routing rules, processing and security issues. The protocol specification is described in detail in RFC5050 [4]. DTNs may be a low cost alternative to traditional wireless networks. The MOSAIC 2B project [5] aims to unleash business opportunities for micro-entrepreneurs in rural areas of South Africa by providing them with entertainment and educational media content. DTN is used as a method of delivery for this content. MOSAIC 2B uses DTN to take content from urban areas and distribute the content to recipients in the rural areas. However, before the actual deployment of the project, the entire network design must be validated, and the software and hardware components must be tested to ensure that they work as expected. Simulators, while convenient, do not provide the most realistic results. Given the complexity of DTN testing, we try to find a middle ground approach by building METhoD, a Mobility Emulator Testbed for DTNs, which we use to emulate the MOSAIC 2B scenario. In this chapter, which extends our previous work [6], [7], we present the METhoD framework for the emulation of delay tolerant networks and show network performance in the MOSAIC 2B settings before the actual implementation in South Africa.

An overview of the technical concepts of the MOSAIC 2B project is given in Section II. Delay tolerant networking and the DTN-enabled infostation are introduced in Section III and Section IV, respectively. The cinema-in-a-backpack kit is presented in Section V. The METhoD testbed is described in Section VI and evaluated in Section VII. Section VIII describes

Fig. 2: Mobile infostation (left) mounted in a bus (right).

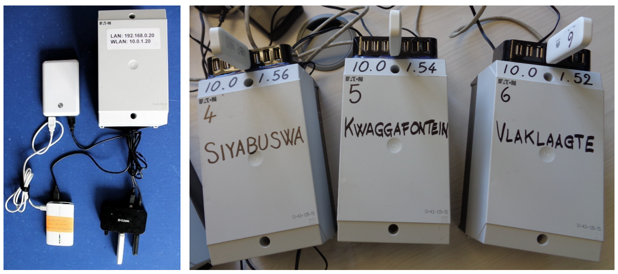

Fig. 3: Fixed infostations ready for deployment.

an emulation of the South African DTN scenario, including evaluation results: System parameters are defined based on the testbed results. Related work is compared in Section IX. Finally, Section X concludes the chapter.

II. MOSAIC 2B Overview

MOSAIC 2B is a research project aiming to provide business opportunities for micro-entrepreneurs living in rural South Africa by delivering multimedia content to them in a low cost manner. Since cellular data access is usually unavailable or expensive in rural areas, content delivery will be performed using DTN. Figure 1 gives an overview of the project. Multimedia content will be delivered with the help of DTN-enabled mobile infostations. Infostations are battery-powered Wireless Local Area Network (WLAN)-enabled devices mounted in buses and bus depots (see Figure 2 and Figure 3). We refer to the infostations placed in the bus depots as fixed infostations and the infostations placed in the buses as mobile infostations. Such infostations act as peers that broadcast the content. The multimedia contents are archived at the fixed infostation located at the main bus depot in the city of Pretoria, which is within a 3G/LTE covered area. Every day, such a fixed infostation fetches a list of contents requested by the micro-entrepreneurs from a server located in the Internet cloud (micro-entrepreneurs are provided with a catalog of the available contents that they can order by sending an SMS to the server), finds such requests in the local archive, and injects them in the DTN network. An advantage of this approach is that the Internet costs are strongly reduced as only a few kilobytes will be downloaded every day, using the cellular network.

From the fixed station in Pretoria, data is sent throughout the DTN network once infostations are in radio range. The mobile infostations serve as intermediate relays (data mules),

Fig. 4: Buses downloading content from the bus station in the city of Pretoria.

which carry the content between the server and final destination, which in our case, is a fixed infostation installed in a bus depot located in a rural area. We have identified three bus depots of the PUTCO bus transportation, Siyabuswa, Vlaklaagte, or Kwaggafontein, about 135 km north-east of Pretoria, which serve several rural communities (see Figure 5).

Buses usually travel between the two locations in predetermined paths, for three hours in the morning and three hours in the evenings. The buses travelling between rural and urban bus depots will act as carriers of data. This can be achieved by placing low cost WLAN-enabled devices in the buses and bus stops at both locations (see Figure 4). When mobile infostations come in contact with the fixed infostations, i.e., when buses arrive at the bus stop, data is forwarded to the mobile infostations via DTN (see Figure 4). The mobile infostations transport then the content between the infostation in Pretoria and another fixed infostation at the rural bus depots. Micro-entrepreneurs living in rural areas near Kwaggafontein, Vlaklaagte and Siyabuswa are equipped with the cinema-in-a-backpack kit that allows them to obtain content from the final fixed infostation and screen it. The infostations are wireless routers (see fixed and mobile infostations in Figure 3 and Figure 2, respectively) equipped with a USB hub, battery supply, external memory, GPS receiver and 3G dongle (the DTN-enabled infostation is presented in more details in Section IV).

The cinema-in-a-backpack kit (see Figure 7 (left)) carried by the entrepreneurs consists essentially of a tablet, a projector, speakers, extra-battery, and a backpack (see Section V). The tablet allows the entrepreneur to request contents and download them from the fixed infostation at the rural bus depots.

Audio watermarking [8], [9] is used to detect copyright infringements. For this purpose, we have designed and implemented a software application able to capture location based audio watermarks embedded in the soundtrack of the movies (see Figure 7 (right)). Such watermarks may contain, for example, GPS coordinates where the projection takes place. In our case, when the watermarks are detected, the mobile application displays the GPS coordinates where the micro-entrepreneurs are supposed to screen the media content and the MOSAIC 2B logo.

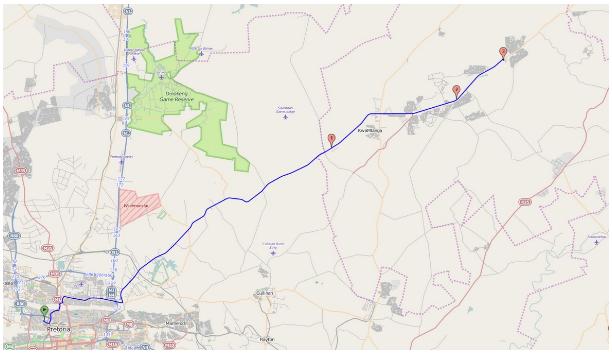

Fig. 5: Bus route between the bus depot in the city of Pretoria (lat: -25.71125, lon: 28.168642) and the bus depots in Vlaklaagte [lat: -25.3796, lon: 28.846872], Kwagafontein [lat: -25.321132, lon: 28.925758], and Siyabuswa [lat: -25.125010, lon: 29.062170] (screenshot taken from http://www.openstreetmap.org).

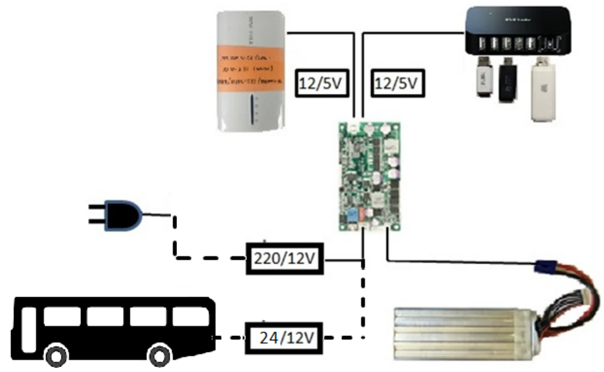

Fig. 6: Diagram of the mobile infostation.

III. Delay Tolerant Networking

Delay tolerant networking [3] is an approach to communication systems that seeks to address technical issues in heterogeneous networks, such as lack of continuous network connectivity mainly due to mobility and limited power, of wireless communication devices. The acronym “DTNs” has been often used to identify either Delay- or Disruption-, or Disconnection- Tolerant Networks, sometimes referring to one or the other without distinction. The InterPlaNetary Internet Special Interest Group was the pioneer in facing issues concerning delay experienced in transferring data between different planets of the solar system. The end-systems must have a free line of sight to be able to communicate since radio waves cannot pass through large solid objects such as planets and moons. In such an environment network protocols and algorithms, unlike the ones for terrestrial communications, must support delay. In this scenario interruptions are somewhat predictable compared to unexpected disturbances that might occur in terrestrial networks, where disconnections can be caused by several factors, such as human mobility or obstacles, or natural disasters, such as earthquakes, seaquakes, floodings, terrorist attacks, etc.

DTNs [10] were conceived for networks in which patterns of connectivity are known or predictable, such as space communication systems (LEO satellite) [11]–[13], sparse mobile ad-hoc networks [14], infostation-based systems [15] and carrier based data collection in sensor networks [16]. However, they can also handle the unpredictable connectivity among mobile devices (e.g. PDAs) [17] and try to address most of the issues raised in networks lacking continuous connectivity.

Intermittent connectivity, long or variable delay, asymmetric data rates, high error rate, high mobility, unknown mobility patterns, energy and storage exhaustion comprise just a few of the potential issues that make end-to-end communication unstable and unlikely in such networks. DTNs overcome such issues by adopting a so-called store-carry-forward paradigm. In these types of networks any synchronous communication paradigm does not perform well. Basic synchronous systems rely on a connected path between sender and receiver, and they negotiate communication parameters (such as clocks) at the data link layer before communication begins. On the other hands, asynchronous systems may simply transmit with no negotiation with the receivers. This may be required when the parties are not in the same portion of network. In fact, networks may be partitioned because nodes may not be in range with one another due to their physical distance and/or because of their mobility.

IV. DTN-Enabled Infostation

Infostations are wireless weatherproof boxes, each of which contains a Wi-Fi router equipped with external memory storage, battery supply, GPS receiver, 3G dongle and a mini-UPS (see Figure 6). We have selected the TP-Link TLMR3040 Ver. 2.0 Wi-Fi router for the infostations. Since the memory of the router is not sufficient for our purposes, we connect a USB hub to the router and use it to accommodate external memory storage. In addition, the GPS receiver and the 3G dongle are also attached to the hub.

The GPS receiver is used to track mobility of the infostations. The 3G dongle allows the fixed infostation in Pretoria to fetch the list of requests of the micro-entrepreneurs from the server on a daily basis. Besides, it allows each infostation to send useful information, system and network performance metrics and mobility traces (GPS coordinates) whenever cellular network is available. This provides some monitoring of the system and enables us to detect failures. All of the infostations have sufficient memory to store all the data sent by the source. The mobile infostations have been designed to be powered up by three different sources: the vehicle battery, the supplementary internal lithium-ion polymer battery, or the power grid. Such power sources are connected to a mini-UPS (Uninterruptible Power Supply) which provides instantaneous protection from input power interruptions by supplying energy stored in the supplementary battery. The mini-UPS can handle a maximum and minimum input operating voltage of 30V and 6V respectively, and an input current limit of 10A. The actual upper bound load current of the system is 4A. We added two 12V/5V converters to connect the router and the USB hub to the mini-UPS. An AC/DC converter (INPUT: 100-240VAC, OUTPUT: 12V 5A) is also provided to connect the mobile infostations to the power grid.

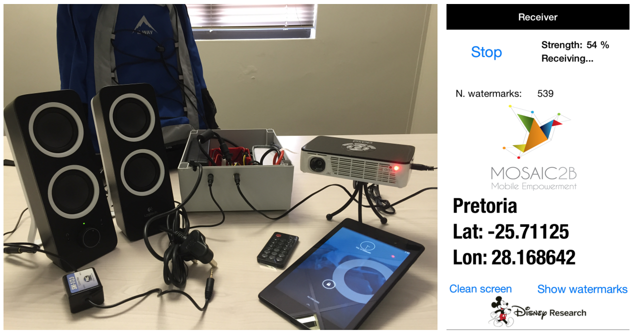

Fig. 7: The cinema-in-a-backpack kit (left) and the audio watermarking application (right).

The infostations have been configured with an OpenWrt release [18], an embedded operating system based on the Linux kernel, and the IBR-DTN, a C++ implementation of the Bundle Protocol (rfc5050) [4] designed for embedded systems [19], [20]. IBR-DTN provides different routing schemes and supports the TCP and UDP convergence layers.

V. Cinema-in-a-Backpack Kit

Mobile cinema entertainment, possibly combined with educational content, will be the use case. Micro-entrepreneurs will be provided with a low-complex, small cinema-in-a-backpack system, which consists of the following components: a tablet, a pico-projector, speakers, and an extra-battery (see the cinema-in-a-backpack kit in Figure 7 (left)). They can download the multimedia content from the DTN network and screen it in remote villages by means of the tablet. The projector can be connected to the tablet via HDMI/VGA. A battery is provided for use when power supply is not available. The components of the cinema-in-a-backpack are shown in Figure 8. An AC/DC converter and a supplementary lithium-ion polymer battery (22.2V 6200mAh) are connected to a mini-UPS and packed in a weatherproof box. The actual upper bound load current of the system is 4.5A. Such a box powers up both the pico-project and the speakers, whose power cables have been modified to be plugged in. Because of the common power source, a ground loop isolator has been used to filter the noise and to improve sound quality.

VI. METhoD Framework

There are several methods of testing DTN setups before deployment. Real-life testbeds provide the most accurate results. However, creating and running an actual testbed can be a time consuming and costly affair, requiring constant supervision. In addition to this, testbeds involving mobility usually require large areas for producing node disconnections. A popular alternative is simulation. Most DTN applications are tested through simulators. Simulation can be easily and conveniently performed in the laboratory. However, most simulators make use of simple models and methods and do not capture various aspects of real life testing. A viable middle ground is an emulation approach. Rather than simulating the mobile nodes, we can use the actual wireless devices and replicate their movement in a testbed. This would be an improvement over simulation since we use the same hardware

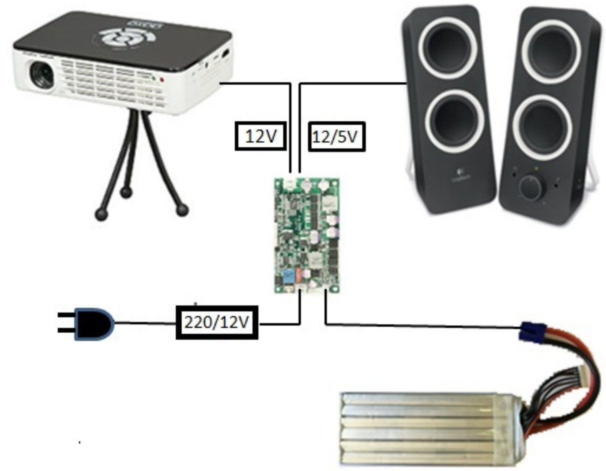

Fig. 8: Diagram of the cinema-in-a-backpack.

and software components that we intend to use in the real-world deployment. The design of the METhoD architecture is driven by the following goals: mobility emulation, DTN layer communication and node cluster isolation. The testbed detects when mobile nodes are within range of each other. When nodes are in contact, they can connect and exchange data at DTN level. Besides, it supports separate clusters of connected nodes. This allows groups of nodes communicating with each other in an intra-cluster fashion at the same time. Keeping these requirements in mind, a centralized controller is designed to coordinate the location and connectivity of all the nodes in the testbed (see Figure 9). Every node has two connection interfaces: one interface is used to communicate with the controller, while the other one is used for communication with the other nodes in the testbed. GPS traces, representing node mobility information, are fed as input into the testbed. It can take GPS traces from existing datasets as input, provided that they are processed into the proper format, that is accepted by the testbed. However, for the emulation of the MOSAIC 2B scenario, GPS traces are generated to represent vehicles moving between Pretoria and Kwaggafontein. Such traces are used by the controller to calculate the time at which two nodes come in contact, so as to simulate their mobility. Connections and disconnections can be set up by changing the firewall rules of the devices on the fly. In this experiment, the bus route between the depot in Pretoria and in Kwaggafontein has been considered as all of the bus routes between urban and rural bus depots show similar mobility pattern.

The mobile nodes are driven by the controller through the WLAN interface, while the data transfers among them take place via Ethernet cables connected to a 48-port switch. Figure 9 shows the overview of the METhoD testbed. The aforementioned configuration of interfaces is done to study the effects of node connectivity without taking into account the nature of the wireless channel. However, METhoD works even if the interfaces are switched. The controller is connected to the same ad-hoc network as the mobile nodes.

As shown in Figure 10, METhoD consists of four main components: the mobility trace generator, the mobility trace processor, the switching module and the visualizer. The trace

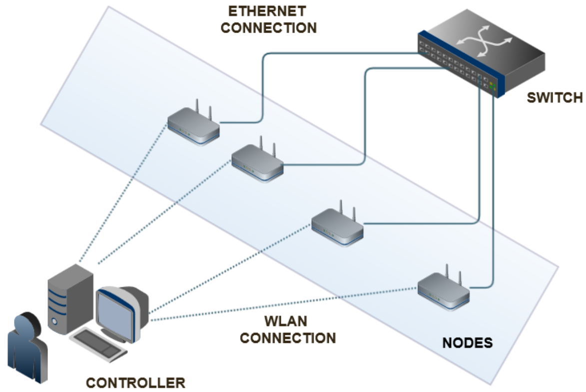

Fig. 9: The testbed overview.

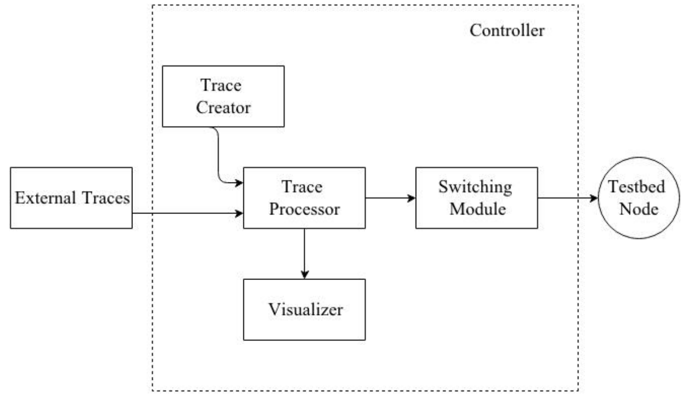

Fig. 10: The METhoD emulator framework.

generator is used to generate mobility traces. The trace processor takes the traces as input and produces connectivity traces. Such information is used by the switching module to run the actual test. The visualizer is a software tool to display the movement of the nodes. The functionality and implementation of the components are described in the following sections.

A. Trace Generator

METhoD provides a module to generate mobility traces of nodes. Mobility can be simulated using GPS traces obtained by drawing paths on an OSM map. The mobility trace creator can generate GPS traces for points placed along a path on a map. In addition, it can generate traces for vehicles of different speeds and start times, and this feature gives us the flexibility to emulate different mobility scenarios.

B. Mobility Trace Processor

This module reads individual mobility traces and output connection information of each node. It achieves this by first extracting common timestamps between two nodes. Two nodes are never in contact if they have no matching timestamps. Then, it checks during which timestamps the nodes are within a certain radio range of each other. It collates this information into time intervals of connections and disconnections. Connectivity traces are also generated in JSON format, which contains node ID, location and connection information for the visualizer.

C. Switching Module

The switching module performs several steps before the actual execution starts. First, it sets firewall rules on all the nodes. Then, it assigns the connectivity traces from the trace processor to all of them. Finally, the switching module reads the connectivity traces and spawns a new thread to handle it. The thread checks the connection status, if the status specifies that two nodes are to be connected, it creates a firewall accept rule to open a connection. If a disconnection is specified, it removes the existing accept rule for the relevant node. Once all of the nodes are ready, the emulation can start.

D. Visualizer

To better visualize the movement of nodes, a simple visualizer is provided. The JSON file containing node IDs, location and connection information, generated by the trace processor is submitted to the visualizer. The visualizer plots the nodes on an OSM map. Figure 11 shows two screenshots while reproducing the MOSAIC 2B scenario.

VII. Validation

Before using the testbed to perform experiments for MOSAIC 2B, we have validated METhoD with the UMass dataset [21]. This dataset has been selected because, as for MOSAIC 2B, it also involves mobility of buses. The dataset consists of connectivity traces of buses travelling in the UMass Amherst campus over a period of three years (2005-2007). The data was collected during the spring and fall months. Each file in the dataset represents connectivity of the buses during a single day. Mobility traces of fifteen buses travelling in the campus have been selected. The dataset has connectivity information over a period of 17.5 hours. The contact times of all of the nodes have been extracted by monitoring the pinging of each device with all the others. Although the experiment runs for the entire 17.5 hours and everything worked as expected, we noticed that sometimes nodes did not ping each other. In addition, some devices reboot occasionally. We notice that such nodes are the ones that have the highest number of connections, that is, they cannot handle many processes running on them. We confirm this by rerunning the entire experiment without launching the DTN services to reduce the number of processes, and then the experiment successfully completes multiple times without any issues.

These results indicate that the number of processes running on a device needs to be taken into account during the deployment in South Africa. Besides, the METhoD emulator has been cross-evaluated by using four different metrics: inter-contact time, contact duration, number of nodes with n-sightings, and number of pairs with n-contacts. Figures 12 and 13 show the contact duration and the inter-contact time [22]–[25] based on ping activity for fifteen nodes. The contact duration between two nodes is the time during which they are in contact with each other. The inter-contact time of two nodes is the time between two consecutive contact durations of the same nodes.

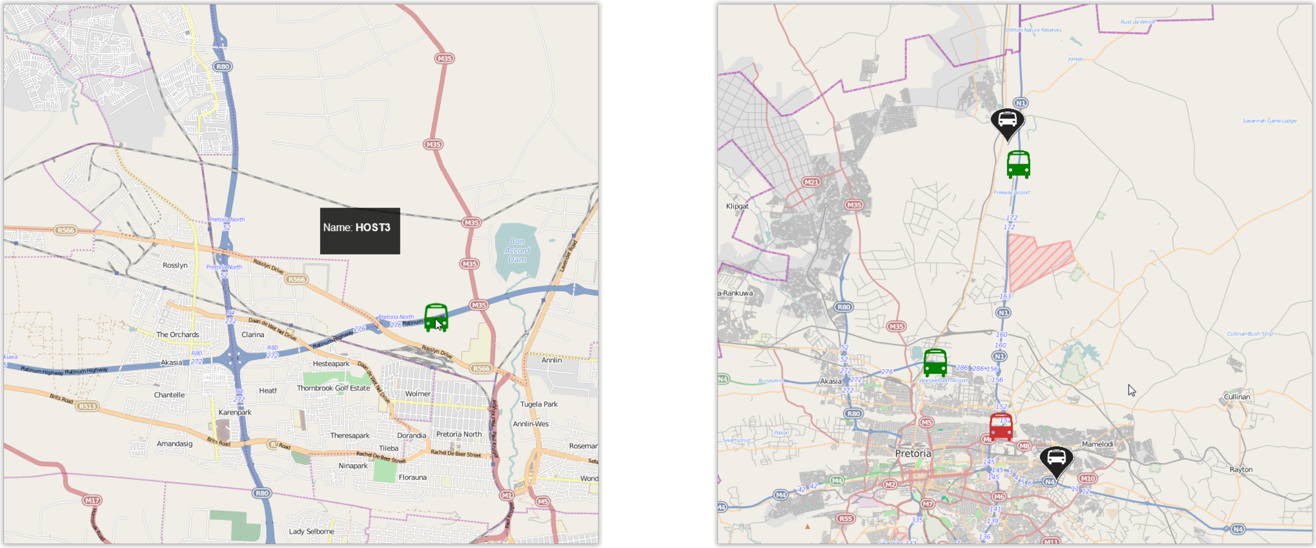

Fig. 11: Snapshots of the MOSAIC 2B scenario in the visualizer of the METhoD emulator.

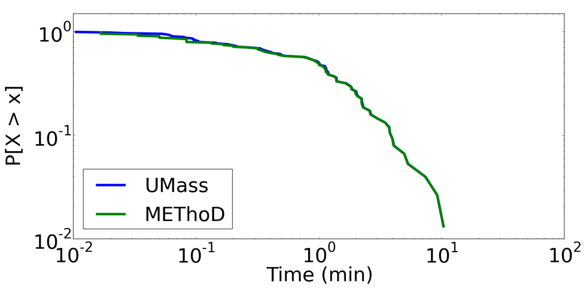

Fig. 12: CCDF of contact durations obtained from the UMass dataset and the METhoD emulator.

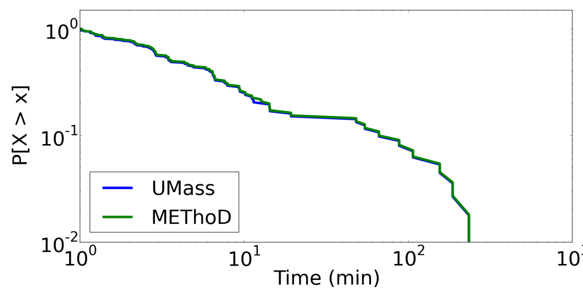

Fig. 13: CCDF of inter-contact durations obtained from the UMass dataset and the METhoD emulator.

Figure 12 shows the CCDF of the contact time distribution generated by METhoD against the actual UMass distribution. Similarly, Figure 13 shows the CCDF of the inter-contact times of the METhoD against the UMass. In both plots, the distributions overlap with the real-world traces of the UMass dataset.

The plot in Figure 14 presents the number of nodes with respect to how many times they were sighted, while the plot in Figure 15 presents the number of pairs with respect to how many times they got in contact with each other, distinguishing between the UMass real-world traces and the METhoD generated traces. Both of the two plots show similar distributions. In Figure 14, all of the devices but one express the same number of sightings, and in Figure 15 all of the pairs but one present the same number of contacts.

Such results show that the METhoD distributions are very good approximations of the UMass distributions, indicating that the connection setup occurs almost immediately, with negligible latency. The latency overhead, namely the excess time taken by METhoD to change the firewall rules with respect to the real traces, is 0.5%.

VIII. MOSAIC 2B Emulation

Experiments to study the performance of the MOSAIC 2B network were reported previously [6], [7]. However, those experiments had shortcomings; outdoor experiments were time consuming and required constant supervision while the indoor experiments involved stationary nodes and were not true representations of the MOSAIC 2B scenario. To perform an in-depth analysis of the system we built METhoD, a testbed that allows us to replicate the target scenario. We use the METhoD testbed to reproduce an environment similar to the one in MOSAIC 2B and run experiments (see Figure 16). In this way, we analyse the network performance in such settings and eventually identify and fix issues before the real deployment.

A. Experimental Setting

Since datasets of mobility traces of buses between Pretoria and Kwaggafontein are not available, we create our own traces using the testbed trace creator module provided by the METhoD framework. As an initial MOSAIC 2B scenario, we assume that buses wait for 15 minutes at the Pretoria stop, travel towards the Kwaggafontein stop and wait there

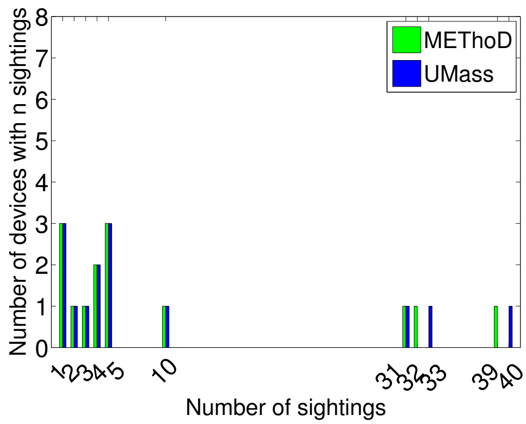

Fig. 14: UMass dataset and the METhoD emulator.

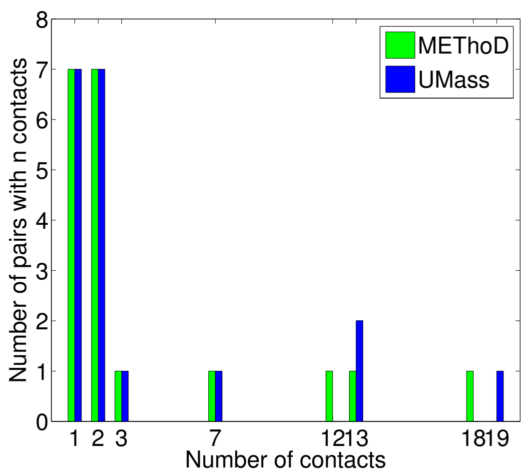

Fig. 15: UMass dataset and the METhoD emulator.

for 15 minutes. Buses have a constant speed of 60 Km/h. The route between Pretoria and Kwaggafontein is the one followed by the buses during the actual South African experiments. Mobility traces for five buses carrying the mobile infostations between two fixed infostations at the end points are generated. The time between two consecutive buses leaving an end point was one hour. The connection range between two infostations is derived using the Haversine formula [26]–[28]), which results to be about thirty meters. Such a distance guarantee a stable connection in real-world scenarios, where obstacles and the nature of the surrounding environment affect the communication link. The infostations are configured with OpenWrt, bleeding edge version (Barrier Breaker), a Linux-based operating system for embedded devices. They are also configured with IBR-DTN [19], [20], a C++ implementation of the bundle protocol targeted at embedded systems running OpenWrt. Since we want to observe data transfer at the DTN layer, we set up METhoD to turn on the IBR-DTN daemon on all the nodes. Initially, all of the nodes are set up with Epidemic routing [29]. To make the scenario more realistic,

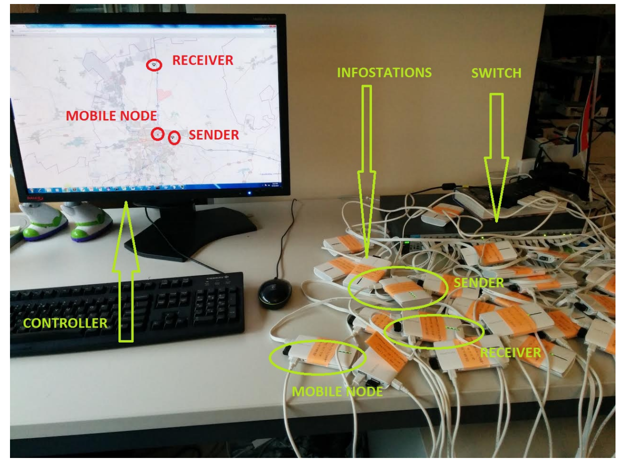

Fig. 16: Testing real devices in the testbed while tracking them on the visualizer map.

the link speed of the Ethernet interface has been changed from 100Mb/s to 10Mb/s using `ethtool`, a standard Linux utility for controlling network drivers and hardware.

B. Emulation with a single movie

Initially, the source in Pretoria sends 1GB data as a single bundle to Kwaggafontein bus station. The MOSAIC 2B scenario considers transfer of large amount of data, that is, movies and multimedia content. The content is therefore forwarded to the mobile infostations mounted on the buses and carried to Kwaggafontein where it is forwarded to the final recipient. However, the contact time between the sender and the mobile nodes is not enough to transfer 1GB bundles. Hence, reactive fragmentation occurs and the mobile nodes get part of the data. Notice that, only 1GB data file is considered for transmission, and all of the mobile nodes receive about 800MB of it.

Figure 17 shows the transmission activity at the receiver during connection with each mobile node. Based on our previous results [6], which show proactive fragmentation expressing higher delivery times, the data file is provided as a whole bundle. The first mobile node transfers the entire fragment to the receiver. Meanwhile, the second mobile node obtains a fragment and reaches the receiver. At this point, the data bundle stored at the receiver is forwarded back to the mobile node. This happens because both of the nodes store fragments of the original bundles identified by different signatures. The receiver assumes that the mobile node does not have its fragment, received by the first carrier, and forward it to the second one. Such a two-way data exchange behaviour is seen in all the following meetings between the receiver and the mobile nodes. Notice that, in the last two trips both nodes try to send data to each other and ultimately, only a few hundred kilobytes are sent while in contact, which cannot be clearly seen in the plot. From this initial experiment, we observe that epidemic routing cannot be set on the receiver. One option would be to set up such a fixed node with default routing that will allow transfer of data bundles only to the intended recipient.

Figure 18 shows the bundle activity at the receiver, which is

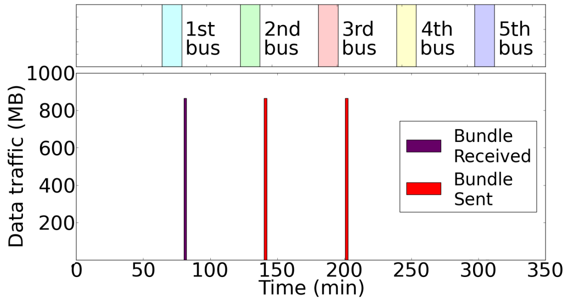

Fig. 17: Data traffic at the receiver. 1GB movie transfer with no proactive fragmentation and receiver with Epidemic routing.

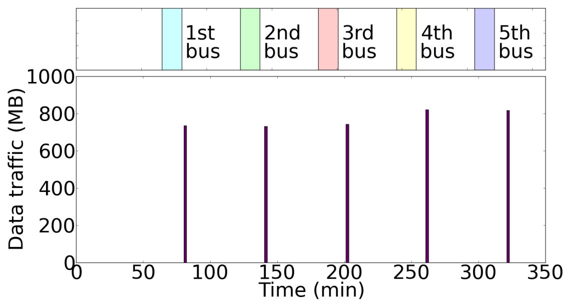

Fig. 18: Data traffic at the receiver. 1GB movie transfer with no proactive fragmentation and receiver with Default routing.

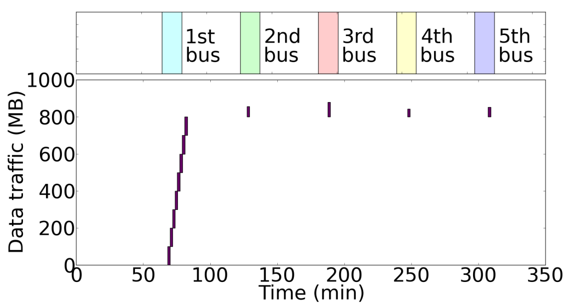

Fig. 19: Data traffic at the receiver. 1GB movie transfer with 100MB bundle size.

set up with default routing, during connection with each mobile node. It shows that 15 minutes is not enough time for the mobile nodes to obtain the entire 1GB bundle from the sender. Hence, reactive fragmentation occurs, and the fixed infostation in Kwaggafontein gets several fragments of the movie. Each mobile node transfers roughly the same initial part of the entire bundle. If the contact time is not sufficient to transfer the entire bundle, the DTN node may fragment a bundle cooperatively when a bundle is only partially transferred. A reactive fragmentation process occurs after an attempted transmission has taken place. A DTN node may also divide application data into multiple smaller blocks and transmit each block as an independent bundle. This approach is used primarily when contact volumes are known in advance. In the MOSAIC 2B scenario this is exactly the case, where buses follow a predefined timetable.

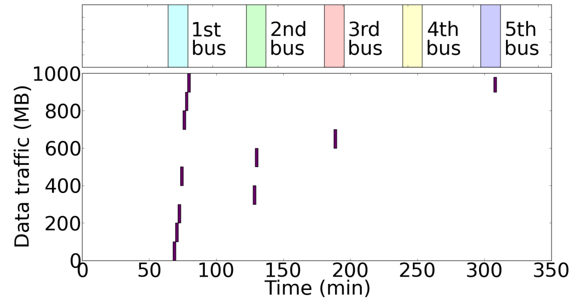

Fig. 20: Data traffic at receiver. 1GB movie transfer with 100MB bundle size and random distribution of bundles by the sender.

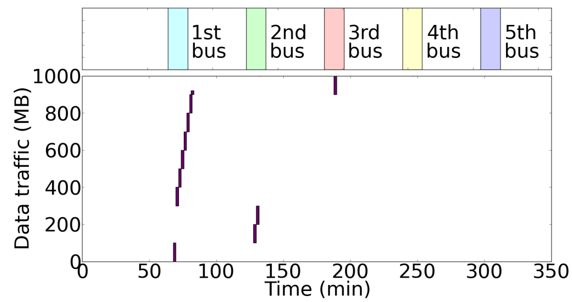

Fig. 21: Data traffic at receiver. 1GB movie transfer with 100MB bundle size and FIFO distribution of bundles by the sender.

In this first experiment, having no proactive fragmentation raises a new issue. Since the contact time is not sufficient to transfer the entire bundle, the receiver gets multiple overlapping large fragments. Eventually, it will receive the remaining parts when the same buses will be back in Kwaggafontein.

Despite longer delays, proactive fragmentation can help decreasing overlapping of data. Since mobile nodes carry fragments of the initial 1GB data bundle, the receiver cannot distinguish between them so as to identify overlapping parts. DTN nodes must receive the entire bundles to be able to identify them. By enabling proactive fragmentation and splitting up the data file in smaller bundles, such that the forwarding time of a bundle is shorter than the contact duration between sender and receiver, will reduce retransmission of data. In this case, disconnections will not occur while transferring the first bundle. Namely, at least the first bundle will not be fragmented. Thus, future nodes carrying the same bundles stored by the receiver will not be forwarded again. For that, we enable proactive fragmentation and split the 1GB movie to be sent to Kwaggafontein in smaller bundles of 100MB each. Figure 19 shows the bundle activity at receiver when proactive fragmentation is enabled. Unlike the distribution in Figure 18, from the second contact onwards bundles are not sent from the beginning. The proactive fragmentation enables the mobile node to send bundles not yet at the receiver. All of the nodes after the first one transmit only a small amount of data. This is because all of the nodes are at the sender for the same amount of time and receive the bundles in the same

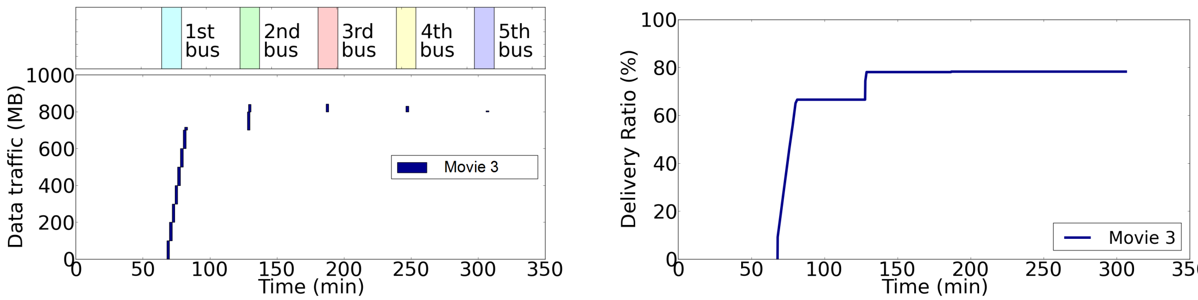

Fig. 22: (Left) Data traffic at receiver. Three 1GB movies, 100MB proactive fragmentation, default routing at receiver. No bundle scheduling. Bundles of only one movie are sent to the receiver by the mobile node. (Right) Delivery ratio.

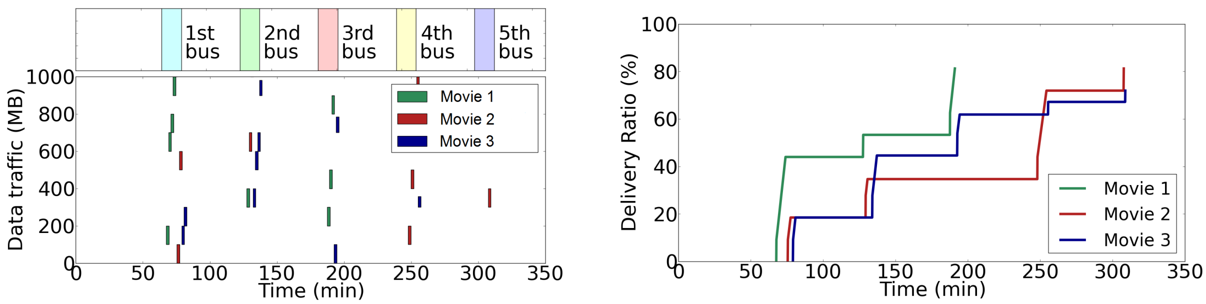

Fig. 23: (Left) Data traffic at receiver. Three 1GB movies, 100MB proactive fragmentation, default routing at receiver. Random scheduling. Bundles of different movies are now sent to the receiver by the mobile node. (Right) Delivery ratio.

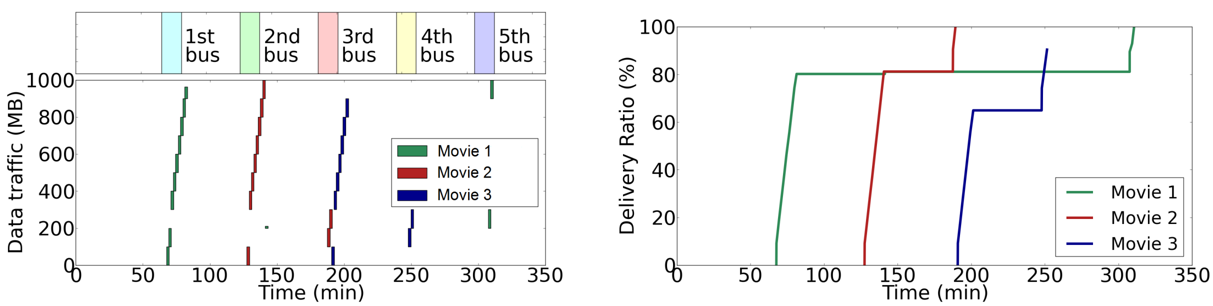

Fig. 24: (Left) Data traffic at receiver. Three 1GB movies, 100MB proactive fragmentation, default routing at receiver. FIFO scheduling. Bundles of different movies are now sent to the receiver by the mobile node. (Right) Delivery ratio.

order. All the mobile nodes take approximately eight bundles. This means that when the nodes after the first one come to the receiver, they transmit only a fragment of a bundle. This is because each fragment of bundle is created from a reactive fragmentation, and it is not recognized as one of those bundles generated by the sender. Therefore, each mobile node produces a fragmented bundle if the contact time is not sufficient to transfer all the content. Thus, the efficiency we gain from using proactive fragmentation does not really help since new bundles are never forwarded.

To ensure that the sender at the Pretoria bus station sends different sets of bundles to different mobile nodes, we make use of a scheduler to decide which bundles are to be sent at every contact. Such a scheduler provides two simple scheduling methods, random and FIFO. Figure 20 shows the effect of a random distribution of bundles by the sender, at the receiver side. It has a better performance with respect to the results in Figure 18 and Figure 19; all of the bundles get delivered over the five contacts. However, since the process of picking bundles is random, there is a chance that the same bundles could be picked during a contact, leading to a redundancy. In the FIFO distribution the bundles are sorted and sequentially forwarded to the carriers. The process is repeated when it reaches the end of the list of bundles. Figure 21 shows the effect of a FIFO distribution of bundles by the sender, at the receiver side. All the bundles are delivered within three contacts. In this case, we can better control which bundles are sent at which contact. Out of these first experiment, to get some quantitative results on which method performs better, we experiment with multiple movies in the following section.

C. Emulation with multiple movies

Since the MOSAIC 2B scenario involves sending multiple movies to recipients, we investigate the behavior of the system when more than one movie is sent. For that, three 1GB movies are to be sent to Kwaggafontein. The devices are set up with proactive fragmentation of 100MB size. After the bundles were created and stored on the sender, the test was kicked off. We first run the experiment without any scheduling at the sender side. The bundle activity at the receiver is shown in Figure 22 (left). After the first contact with the mobile infostations, only a fragment of the same bundle of the same movie is sent to the receiver at each contact. This means that, irrespective of how many mobile nodes are there, they all receive the same bundles. Hence, unless the mobile node returns to the sender a second time, new bundles will not be received at the receiver.

The same experiment is also repeated with random and FIFO bundle distribution. As shown from the results of the random distribution of bundles in Figure 23 (left) with respect to Figure 22 (left), a simple random distribution strategy can increase the efficiency. The FIFO distribution is the most efficient, as shown in Figure 24 (left). It leads to the delivery of the largest amount of data. This is because there is greater control over which bundles are chosen by the sender to transmit to the mobile node at each contact.

The delivery ratios of the three experiments are also presented next to each plot in Figures 22, 23, and 24, respectively. None of the three movies are delivered in one journey when there is no scheduling in place because all nodes get the same bundles. The random distribution has a much better performance and manages to deliver approximately 80% of all the movies. Finally, the FIFO distribution has the best performance; it almost succeeds to deliver all of the three movies. More precisely, it delivered two movies, and 90% of the third one.

IX. Related Work

Previous work on the MOSAIC 2B project mainly consists of initial experiments to determine the performance of the system [6], [7]. They describe the effect of varying bundle sizes on delivery time under different network topologies. These works help assess the system, but the experiments were time consuming. The testbed we propose aims to facilitate the experimentation process in the laboratory and builds upon the results produced from [6] and [7].

Even though majority of the tests in DTN related work have been performed on real testbeds or using simulators, efforts have been taken to create realistic emulators for mobile scenarios. RAMON [30] is an emulator that tries to imitate realistic wireless characteristics in high speed systems. A moving node is emulated using an access point, a computer, an attenuator and an antenna. Speed is emulated by changing the signal strength according to a path loss equation. RAMON also uses a propagation model tailored towards indoor and micro-cellular environments. While RAMON is useful in the emulation of the physical layer, it does not scale well. New models have to be created for every scenario and attenuators have to be placed strategically for the most accurate readings.

An on/off based emulation approach is followed in [31]. The authors claim that a tight interaction between the application and the MAC/PHY layers is missing in existing emulation methods. Their idea is to have a grid of nodes running emulation clients. An emulation server would trigger operations. Mobility is emulated by migrating a running application from one node to the other based on a mobility pattern. In other words, on obtaining a trigger from the emulation server, the emulation client on the node would take a snapshot of the state of the running application on the node, stop it and migrate it to the next node on the path, where the application would be restarted. This method satisfies the interaction criterion but migration of the application introduces latency and the emulation method does not work well in high speed scenarios, where fast migrations would be required. Besides, the size of the snapshots impacts on the performance of the system as it creates further delay. Finally, channel quality fluctuations cannot be emulated by this method.

A large scale spatial switching emulation method is introduced in [32] to tackle large scale topologies. They make use of the m-ORBIT grid testbed and a split-stack architecture to emulate mobility. The application and network layers of a node run on a single machine, the virtual node. The physical and link layers of the grid nodes are used to emulate mobility. Communication between the virtual and grid nodes occurred via tunneling over gigabit Ethernet. The virtual nodes keep on changing tunnels to different grid nodes based on the mobility patterns. Communication between two grid nodes representing two different mobile nodes is done over the wireless interface. Additive white Gaussian noise is introduced into the testbed to disconnect nodes despite their close proximity. This approach is able to emulate mobility on a larger scale and has a much simpler noise model. However, it could not handle high speed and the tunneling introduces some amount of latency. [30] and [31] employs complex designs to emulate mobility in the physical layer. [32], while using a simpler noise model, makes use of tunneling. These decisions resulted in scalability issues. Our design circumvents these issues by forgoing the physical layer emulation and utilizing a less complex emulation method.

With regards to content dissemination, works like [33], [34] and [35] propose solutions to adapt the BitTorrent protocol to ad-hoc networks. [36] develops a content dissemination protocol for vehicular ad-hoc networks called SPAWN. This is based on a strategy known as rarest-closest first, where a content piece's distribution is determined by a neighbouring node's proximity and possession of that piece. [37] has a similar approach for opportunistic networks, where pieces were disseminated based on their prevalence in the network. The METhoD framework can be used to experiment with strategies based on this prior work and comparisons can be made with our current distribution methods.

DakNet [38] was one of the first projects to propose the use of existing transport infrastructure to bring connectivity to developing regions in India and Cambodia at a lower cost. It was a wireless ad-hoc network where buses equipped with mobile access points (MAP) would transmit data to Wi-Fi-enabled kiosks in villages when they were in contact. A MAP consisted of a PC, an antenna, an amplifier and a power supply. The main uses of DakNet were to transport land records between urban and rural areas and to provide asynchronous connectivity to the Internet. The MOSAIC 2B project is in a similar vein as DakNet. However, we make use of delay tolerant networking for transmission of larger amounts of data. Not only this, we use different equipment (section 2.4) that we hope brings down the costs further.

KioskNet [39] attempted to improve the service provided by rural kiosks in a low cost manner. These kiosks are a means to provide Internet connectivity at a low cost but technical problems lead to unreliable service. In KioskNet, buses acted as ferries that carried data to and from kiosks to a gateway that had reliable Internet connection. Both kiosks and ferries were equipped with low cost single board computers. Delay Tolerant Networking was used - the DTN2 implementation was modified for the scenario and an overlay called OCMP (Opportunistic Communication Management Protocol) that runs on top of DTN was plugged in. KioskNet also addressed the issue of cost by allowing recycled PCs the ability to run their controller software. KioskNet was deployed in India and Ghana. The work on KioskNet later led to the development of VLink - a delay tolerant method to transmit data via USB keys. Our approach is quite similar to KioskNet since we use buses as ferries and a DTN method of communication. However, we differ in the hardware and DTN software used and our aim of delivering large amounts of media content to the end users.

X. Conclusion

Our DTN network for media-content distribution using bus transportation in rural South Africa unleashes business opportunities for micro-entrepreneurs in such regions. Before the actual deployment the entire system must be evaluated during the design process. We have developed METhoD, a mobility emulator testbed for DTNs. METhoD is able to emulate the mobility of the devices and their connections on the DTN layer in real time, using port firewalls to mimic the device proximity and hence wireless connectivity taken from synthetic location traces. We analyse network performance and identify possible issues. The experiments we conducted provide interesting insights into the behavior of the devices and the whole network. We observe that setting up proactive fragmentation at the sender side leads to more efficient use of the contact duration between two nodes. Besides, since bundles are always sent in the same order from the sender, an increase in the number of mobile nodes has no positive effect on the amount of data delivered. A bundle distribution strategy is necessary for higher delivery ratios. A FIFO distribution of bundles is a better option than a random strategy. The experiments also show that the devices have a limitation when it comes to large number of processes and we must ensure that unnecessary processes are not running in the background.

Acknowledgement

We thank Johannes Morgenroth and the team of Lars Wolf at Techn. Univ. Braunschweig, Germany, for provisioning software and support of their IBR-DTN. We also thank the Disney Lab Artists, Maurizio Nitti and Alessia Marra, for providing artworks. The MOSAIC-2B project is partially funded by the EU FP7 framework program for research, technological development and demonstration under grant agreement no.611796.

References

[1] The World in 2014, ICT Facts and Figures. *ITU Technical report, Geneva, Switzerland*, 2014.

[2] Enabling Network Connectivity in Developing Nations. *NYUAD Research Report*, 2012.

[3] Delay-Tolerant Networking Research Group (DTNRG). https://irtf.org/dtnrg. Retrieved 12/2015.

[4] K. Scott and S. Burleigh. Bundle Protocol Specification. RFC 5050 (Experimental), November 2007.

[5] MOSAIC 2B: Mobile Empowerment. http://mobile-empowerment.org. Retrieved 3/2015.

[6] A. Galati, T. Bourchas, S. Siby and S. Mangold. System Architecture for Delay Tolerant Media Distribution for Rural South Africa. *in the proceedings of the 9th ACM International Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization (WiNTECH 2014)*, pages 65–72, 2014.

[7] A. Galati, T. Bourchas, S. Siby, S. Frey, M. Olivares and S. Mangold. Mobile-Enabled Delay Tolerant Networking in Rural Developing Regions. *in the proceedings of the IEEE Global Humanitarian Technology Conference (GHTC 2014)*, pages 699–705, 2014.

[8] R. Frigg, G. Corbellini, S. Mangold, and T.R. Gross. Acoustic data transmission to collaborating smartphones: An experimental study. *11th Annual IEEE Conference on Wireless On-demand Network Systems and Services (WONS)*, pages 17–24, 2014.

[9] R. Frigg, S. Mangold, and T.R. Gross. Multi-channel acoustic data transmission to ad-hoc mobile phone arrays. *ACM SIGGRAPH 2013 Mobile*, page 1, 2013.

[10] K. Fall. A delay-tolerant network architecture for challenged internets. *Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, pages 27–34, 2003.

[11] Interplanetary Internet Special Interest Group (IPNSIG). http://www.ipnsig.org. Retrieved 12/2015.

[12] S. Burleigh, V. Cerf, B. Durst, A. Hooke, R. Rumeau, K. Scott, E. Travisand, and H.Weiss. The Interplanetary Internet: The Next Frontier in Mobility. *IPN Special Interest Group*, 2001.

[13] R. C. Durst P. D. Feighery and K. L. Scott. Why not use the Standard Internet Suite for the Interplanetary Internet? *IPN Special Interest Group*, 2003.

[14] Qun Li and Daniela Rus. Sending messages to mobile users in disconnected ad-hoc wireless networks. In *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, MobiCom '00, pages 44–55, New York, NY, USA, 2000. ACM.

[15] R. H. Frenkiel, B. Badrinath, J. Borras, and R. Yates. The infostations challenge: balancing cost and ubiquity in delivering wireless data. *IEEE Personal Communications*, 7(2):66–71, 2000.

[16] R. C. Shah, S. Roy, S. Jain, and W. Brunette. Data Mules: Modeling a Three-Tier Architecture for Sparse Sensor Networks. *In Proceedings of the First IEEE Sensor Network Protocols and Applications*, pages 33–41, 2003.

[17] X. Chen and A. L. Murphy. Enabling disconnected transitive communication in mobile ad-hoc networks. *In Workshop on Principles of Mobile Computing*, pages 30–41, 2001.

[18] OpenWrt, Wireless Freedom. https://www.openwrt.org. Retrieved 12/2015.

[19] W. Pottner S. Schildt, J. Morgenroth and L. Wolf. IBR-DTN: A lightweight, modular and highly portable Bundle Protocol implementation. *Electronic Communications of the EASST*, pages 1–11, 2011.

[20] A modular and lightweight implementation of the bundle protocol. http://trac.ibr.cs.tu-bs.de. Retrieved 2/2015.

[21] J.Burgess, B.Gallagher, D.Jensen, and B.N.Levine. Maxprop: Routing for vehicle-based disruption-tolerant networks. In *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, pages 1–11, April 2006.

[22] 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.*, 6:606–620, 2007.

[23] 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.

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

[25] J. Leguay, A. Lindgren, J. Scott, T. Friedman, and J. Crowcroft. Opportunistic content distribution in an urban setting. *Proceedings of the 2006 SIGCOMM workshop on Challenged networks, CHANTS '06*, pages 205–212, 2006.

[26] L Vanajakshi, SC Subramanian, and R Sivanandan. Travel time prediction under heterogeneous traffic conditions using global positioning system data from buses. *IET intelligent transport systems*, 3(1):1–9, 2009.

[27] Julien Montavont and Thomas Noel. Ieee 802.11 handovers assisted by gps information. In *Wireless and Mobile Computing, Networking and Communications, 2006.(WiMob'2006). IEEE International Conference on*, pages 166–172. IEEE, 2006.

[28] Tommy Nguyen and Boleslaw K Szymanski. Using location-based social networks to validate human mobility and relationships models. In *Advances in Social Networks Analysis and Mining (ASONAM), 2012 IEEE/ACM International Conference on*, pages 1215–1221. IEEE, 2012.

[29] A.Vahdat, D.Becker, et al. Epidemic routing for partially connected ad hoc networks. Technical report, Technical Report CS-200006, Duke University, 2000.

[30] E.Hernandez and A.S.Helal. RAMON: Rapid-Mobility Network Emulator. In *Proceedings of the 27th Annual IEEE Conference on Local Computer Networks*, LCN '02, pages 809–820, Washington, DC, USA, 2002. IEEE Computer Society.

[31] H.Yoon, J.Kim, M.Ott, and T.Rakotoarivelo. Mobility emulator for dtn and manet applications. In *Proceedings of the 4th ACM International Workshop on Experimental Evaluation and Characterization*, WINTECH '09, pages 51–58, New York, NY, USA, 2009. ACM.

[32] K.Ramachandran, S.Kaul, S.Mathur, M.Gruteser, and I.Seskar. Towards large-scale mobile network emulation through spatial switching on a wireless grid. In *Proceedings of the 2005 ACM SIGCOMM Workshop on Experimental Approaches to Wireless Network Design and Analysis*, E-WIND '05, pages 46–51, New York, NY, USA, 2005. ACM.

[33] Sundaram Rajagopalan and Chien-Chung Shen. A cross-layer decentralized bittorrent for mobile ad hoc networks. In *Mobile and Ubiquitous Systems: Networking & Services, 2006 Third Annual International Conference on*, pages 1–10. IEEE, 2006.

[34] Neelakantam Gaddam and Anupama Potluri. Study of bittorrent for file sharing in ad hoc networks. In *Wireless Communication and Sensor Networks (WCSN), 2009 Fifth IEEE Conference on*, pages 1–6. IEEE, 2009.

[35] Mohamed Karim Sbai, Emna Salhi, and Chadi Barakat. P2p content sharing in spontaneous multi-hop wireless networks. In *Communication Systems and Networks (COMSNETS), 2010 Second International Conference on*, pages 1–10. IEEE, 2010.

[36] Alok Nandan, Shirshanka Das, Giovanni Pau, Mario Gerla, and MY Sanadidi. Co-operative downloading in vehicular ad-hoc wireless networks. In *Wireless On-demand Network Systems and Services, 2005. WONS 2005. Second Annual Conference on*, pages 32–41. IEEE, 2005.

[37] N.Belblidia, M.Dias de Amorim, L.H.M.K.Costa, J.Leguay, and V.Conan. Pacs: Chopping and shuffling large contents for faster opportunistic dissemination. In *Wireless On-Demand Network Systems and Services (WONS), 2011 Eighth International Conference on*, pages 9–16, Jan 2011.

[38] A. Pentland, R. Fletcher, and A. Hasson. Daknet: Rethinking connectivity in developing nations. *IEEE Computer Society*, 37:78–83, 2004.

[39] 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*, 37:95–100, 2007.