

Delay Tolerant Networking to Extend Connectivity in Rural Areas Using Public Transport Systems: Design And Analysis

Salah Abdeljabar, *Graduate Student Member, IEEE*, Marco Zennaro, *Senior Member, IEEE*, and Mohamed-Slim Alouini, *Fellow, IEEE*

Abstract—In today's digital age, access to the Internet is essential, yet a significant digital divide exists, particularly in rural areas of developing nations. This paper presents a Delay Tolerant Networking (DTN) framework that utilizes informal public transportation systems, such as minibus taxis, as mobile data mules to enhance connectivity in these underserved regions. We develop a probabilistic model to capture the randomness in vehicle mobility, including travel times and contact durations at bus stops. Key performance metrics are analyzed, including average data transmission rate and Peak Age of Information (PAoI), to assess the effectiveness of the proposed system. An analytical approximation for the Mean PAoI (MPAoI) is derived and validated through simulations. Case studies from real-world datasets in Nouakchott, Accra, and Addis Ababa demonstrate the practical applicability and scalability of our framework. The findings indicate that leveraging existing transportation networks can significantly bridge the digital divide by providing reliable internet-like connectivity to remote areas.

Index Terms—Delay Tolerant Networking (DTN), Informal public transport, and Digital divide.

I. INTRODUCTION

A. MOTIVATION

In today's digital age, Internet access is widely recognized as essential for economic and social progress. The United Nations (UN) emphasizes this importance in their Common Agenda, declaring universal Internet access by 2030 as a fundamental human right [1]. Despite rapid advancements in connectivity solutions, a significant disparity in access persists between developing and developed nations. According to the latest reports from the International Telecommunication Union (ITU), approximately 2.9 billion people, which is nearly one-third of the global population, still lack Internet access, predominantly in rural areas and developing regions [2]. This gap contributes to a persistent digital divide, placing those without access to the Internet and other information and communication technologies (ICTs) at a marked socio-economic disadvantage. Although various efforts have been made to improve rural connectivity through traditional backhaul infrastructure, these solutions often face significant challenges [3]. In many cases,

Salah Abdeljabar and Mohamed-Slim Alouini are with the Computer, Electrical and Mathematical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia (e-mail: salah.abdeljabar@kaust.edu.sa, slim.alouini@kaust.edu.sa). Marco Zennaro is with the Abdus Salam International Centre for Theoretical Physics, Trieste, Italy (e-mail: mzenaro@ictp.it).

the geographic and economic conditions of rural areas make it difficult to deploy and maintain such infrastructure, further widening the access gap [3].

Delay Tolerant Networking (DTN) has emerged as a viable solution for extending connectivity to rural and underserved areas [4]. Operating on a store-carry-forward paradigm, DTN allows data to be stored locally on mobile devices, transported as devices move, and forwarded when they enter the communication range of another device. This type of DTNs, known as mule-based DTNs, enables direct device-to-device communication using local radio signals, which reduces the need for fixed network infrastructure. DTN is particularly suited for resource-constrained environments, offering a cost-effective backhaul-like connectivity approach to bridging the digital divide where traditional connectivity infrastructure may be impractical or costly [5]. This communication paradigm makes use of existing transportation systems, such as buses and trucks, to physically carry data between network points. These mobile carriers serve as the primary mechanism for data movement in the absence of continuous network coverage.

Informal public transportation systems offer a promising opportunity for supporting DTNs in developing regions. One such system is the use of minibus taxis, which play a central role in the daily transportation of people in many developing countries. For example, in South Africa, medium-sized minibus taxis dominate the public transport sector, with over 60 % of the population relying on them for mobility [6]. These taxis transport approximately 14 million passengers daily, proving their significance within the transportation network. When used as mobile data mules, these vehicles can carry information between rural areas that lack Internet access and urban centres that are connected to the Internet. Such integration of mobility and data transfer offers a practical and cost-effective way to enhance connectivity in regions where traditional infrastructure is limited or absent.

B. RELATED WORKS

DTN was originally developed by NASA for interplanetary satellite communication; however, it later showed significant potential in extending connectivity to rural regions where traditional communication infrastructure is unavailable or costly. DTN enables intermittent and opportunistic access to the Internet through a store, carry, and forward mechanism, allowing data to be transferred whenever communication opportunities

arise. Several studies have investigated the application of DTN in rural connectivity projects across different parts of the world. In [7], authors proposed the use of public transport buses to carry data and establish intermittent links between remote villages in Bangladesh, effectively overcoming infrastructural barriers. Similarly, a study in South Africa [8] examined the use of taxis and other transport modes to facilitate data transfer between rural communities. Additional studies have emphasized DTNs flexibility in addressing geographic constraints. For example, the authors of [9] explored DTN deployments in remote mountainous areas of Sweden, where helicopters were used as mobile data mules. In another case, the authors of [10] demonstrated DTNs adaptability through the use of Android-based devices to support communication in resource-constrained environments. Taken together, these studies demonstrate that DTN can adapt to a wide range of rural environments and transportation contexts, offering a flexible and cost-effective solution for addressing connectivity challenges in underserved areas.

While previous studies have demonstrated the feasibility of DTN systems for rural connectivity using mobile data mules such as buses, ferries, and helicopters [7], [8], [9], [10], [11], [12], [13], many of these works remain primarily empirical or descriptive, offering limited formal analysis of system performance. In particular, the integration of informal public transport systems, such as the ones found in many cities in developing countries, into DTN framework is still underexplored, especially with respect to quantifying key communication performance metrics. This paper addresses these gaps by introducing a probabilistic modeling framework based on renewal theory to capture the stochastic movement patterns of informal transport systems and incorporate them in the DTN design. Analytical expressions are derived for performance metrics such as the Mean Peak Age of Information (MPAoI), and cost considerations are incorporated to determine the minimum number of DTN data mules required to meet performance targets. The framework is validated using real-world mobility datasets from cities including Nouakchott, Accra, and Addis Ababa, reflecting the irregularities in informal transport behavior. This integrated approach of formal modeling, analytical evaluation, and validation with real-world mobility traces distinguishes this work from prior efforts and provides a more systematic basis for deploying mule-based DTNs in rural environments.

C. ARTICLE OUTLINE

Motivated by the pressing need to bridge the digital divide, this work investigates the performance of DTN frameworks that leverage informal public transportation systems as data mules. We propose a general methodology to evaluate DTN performance from a communication perspective, with particular focus on optimizing the number of DTN data mules required to meet communication performance targets. The primary research question explored in this article concerns the effectiveness of such systems in delivering Internet-like connectivity to underserved or remote regions. Specifically, we study scenarios where two disconnected areas, an Internet-

connected urban center and a rural location with limited or no connectivity, are linked through the movement of DTN data mules. Our aim is to assess the quality of this connection by analyzing data flow and evaluating performance metrics.

To this end, we develop a probabilistic model based on renewal theory to describe the mobility behavior of informal transport systems, including transportation routes, travel durations and waiting times. We also incorporate cost and quality-of-service (QoS) constraints into the design process. Particularly, we consider the MPAoI and average data transmission rate as key QoS metrics, drawing parallels to optimization strategies used in traditional cellular networks, but adapted to the mobile nature of DTN nodes. The framework is validated through real-world mobility datasets from cities including Nouakchott, Accra, and Addis Ababa, capturing the irregular dynamics of informal transport. Our goal is to provide a scalable framework that provides indicative metrics for the predicted communication performance that such systems can achieve, and hence offering insights into their practical feasibility and design optimization in rural contexts. The main contributions of this article are summarized as follows:

- We introduce a comprehensive framework to model the mobility patterns of informal transport systems, including travel durations and waiting times at bus stops. This framework is based on renewal theory to characterize data mule movement between rural and urban regions.
- We analyze key performance metrics such as the mean transmitted data size, average data transmission rate, and Mean Peak Age of Information (MPAoI). Although exact analytical expressions for MPAoI is complex, we provide an analytical approximation of the MPAoI based on the superposition of renewal processes at equilibrium, and validate the approximation via simulations.
- We validate the proposed framework using real-world mobility datasets from informal public transport systems in cities including Nouakchott, Accra, and Addis Ababa. These case studies demonstrate the practical applicability and scalability of our framework.

D. NOTATION SUMMARY

To facilitate understanding and maintain consistency throughout this manuscript, Table I provides a comprehensive summary of all key notations, symbols, and variables used in our analysis. The notations are organized into several categories: system parameters that define the network configuration, random variables that characterize the stochastic behavior of the transport system, performance metrics that quantify system effectiveness, and derived quantities used in the analytical framework.

II. SYSTEM MODEL

A. OVERVIEW

In this section, we present a model for public transportation systems that connect two distant regions, an urban city and

TABLE I: Summary of Notations

Notation	Description
<i>System Parameters</i>	
N	Number of vehicles serving the route between city and village
n	Number of vehicles equipped with DTN modules ($n \leq N$)
R_{link}	Wireless link data rate between a DTN-equipped vehicle and a fixed DTN gateway during contact time (Mbps)
\mathcal{R}	Effective DTN uplink/downlink data transmission rate (Mbps)
c_1	Minimum contact time at bus stops (minutes)
c_2	Maximum contact time at bus stops (minutes)
T_{\min}	Minimum travel time between regions A and B (minutes)
C	Cost of a single DTN node placed on vehicles (USD)
C_{gateway}	Cost of a fixed DTN gateway (USD)
<i>Random Variables</i>	
$T_{c,A}$	Contact time at region A (urban gateway), $T_{c,A} \sim \text{Unif}(c_1, c_2)$
$T_{c,B}$	Contact time at region B (rural gateway), $T_{c,B} \sim \text{Unif}(c_1, c_2)$
T_d	Random delay during vehicle's travel, $T_d \sim \text{Exp}(1/\bar{T}_d)$
\bar{T}_d	Average travel time delay, $\mathbb{E}[T_d]$
T_{AB}	Travel time from region A to region B, $T_{AB} = T_{\min} + T_d$
T_{BA}	Travel time from region B to region A, $T_{BA} = T_{\min} + T_d$
$T_{v,i}$	Inter-arrival time of the v -th vehicle for its i -th round trip
T_s	Inter-arrival time of the superimposed process (all n vehicles)
m_v	Data size transmitted or received per vehicle arrival (Mbits)
<i>Performance Metrics</i>	
AoI	Age of Information
MAoI	Mean Age of Information
PAoI	Peak Age of Information
MPAoI	Mean Peak Age of Information
$\mathbb{E}[\mathcal{R}]$	Mean data transmission rate of the DTN system (Mbps)
$\mathbb{E}[m_v]$	Mean transmitted data size per vehicle arrival (Mbits)
λ_s	Arrival rate of the superimposed process (vehicles per minute)
<i>Derived Quantities</i>	
$N_v(t)$	Renewal process for the v -th vehicle
$N_s(t)$	Superimposed renewal process for all n DTN-equipped vehicles
$\mathbb{E}[T_{v,i}]$	Mean inter-arrival time of a single vehicle (mean round-trip time)
$\mathbb{E}[T_s]$	Mean inter-arrival time of the superimposed process
$F_X(x)$	Cumulative distribution function (CDF)
$f_X(x)$	Probability density function (PDF)
t_l	Generation time of the l -th update at the source
t'_l	Reception time of the l -th update at the monitor
$u(t)$	Timestamp of the most recent update received by time t
μ	Mean round-trip time, $\mathbb{E}[T_{c,A} + T_{AB} + T_{c,B} + T_{BA}]$

a rural village. Broadly, public transport systems can be classified into two categories. The first includes formal transport systems that operate on fixed schedules and predefined routes, often regulated by governmental or municipal authorities. These are commonly found in urban areas. The second category comprises informal systems, which are typically unregulated and lack predefined schedules or fixed routes. Examples include minibuses, shared taxis, and similar modes of transport frequently used in developing countries [14]. In informal transport systems, vehicles often wait at central stops until a sufficient number of passengers with similar destinations are gathered before departing. The departure times and the routes taken are not fixed in advance, and both the waiting time at stops and the travel time between regions are inherently random. This work focuses on such systems, modeling their stochastic behavior through a probabilistic framework. The framework captures the key sources of randomness and provides analytical tools to support the design and deployment

of DTNs using these transportation systems as data mules.

B. CONSIDERED SCENARIO

We consider a DTN system that operates between two geographically separated regions: an urban area with Internet connectivity and cloud services, and a rural village with limited or no Internet infrastructure. Fig. 1 illustrates the architecture we consider in this work. In the rural region, various DTN-based services operate under the assumption of non-real-time communication, such as asynchronous email, web caching, digital learning platforms [5], and IoT-based environmental or agricultural monitoring. These services rely on intermittent but reliable data delivery, making DTN a suitable solution in such contexts. Connectivity between the two regions is facilitated by informal public transportation vehicles, such as minibuses or shared taxis, which travel between a designated bus stop in the city and another in the rural area. Among the total fleet of \mathcal{N} vehicles serving this route, only n ($n \leq \mathcal{N}$) are equipped with DTN communication modules. Each region's main bus stop is equipped with a DTN gateway that establishes wireless links (e.g., via Wi-Fi) with arriving vehicles.

For simplicity, these gateways, referred to as node A in the urban region (region A) and node B in the rural region (region B), serve as fixed communication endpoints for data transfer. As DTN-equipped vehicles move between regions A and B, they carry data in both directions: from the Internet (via node A) to rural users (via node B), and from rural users back to Internet-based services. These vehicles thus function as mobile data mules, bridging the connectivity gap between urban and rural regions. The DTN architecture described in this work is particularly relevant for regions where informal public transportation systems dominate intercity mobility, which are prevalent in many cities in developing countries [14].

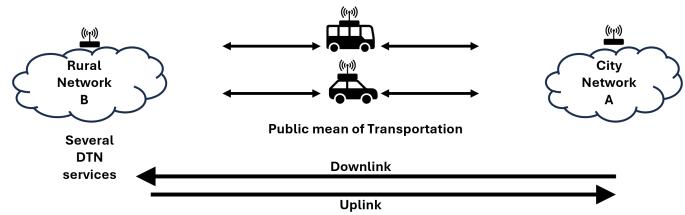


Fig. 1: The considered DTN scenario, where DTN modules carried via informal public transport means are used to connect the city and rural networks.

For simplicity and analytical tractability, the proposed model considers communication only between DTN-equipped vehicles and the fixed DTN gateways (nodes A and B), excluding vehicle-to-vehicle (V2V) communication. In the context of informal rural-urban transport, such vehicles typically travel long distances on disjoint schedules and with random corridors, making in-transit V2V encounters infrequent. Moreover, informal systems often operate sequentially, with vehicles departing only after a sufficient number of passengers are on board, further reducing the likelihood of overlap between DTN-equipped vehicles during transit. It should be noted that the discussed DTN framework targets applications that

do not require instant data delivery. Some IoT scenarios fit this category, including environmental monitoring, sensor data collection in agriculture, and wildlife tracking, where measurements are stored and forwarded when possible [15], [16], [17]. However, the framework is not intended for IoT services that need continuous or low-latency connections, such as critical emergency alerts, industrial automation, or real-time health monitoring. Such services must rely on alternative connectivity solutions that guarantee immediate and continuous data transmission.

C. DTN DATA MULE MODEL

In this part, we develop a probabilistic model that describes the DTN data mules movement traversing between the two regions. We assume that these vehicles are mainly informal public transport providing shared transportation for multiple passengers, such as medium-sized minibus taxis and vans that dominate the public transport sector in many developing countries [6]. These vehicles operate without fixed, predefined intermediate stops or schedules but with roughly fixed service corridors that start/end at a main bus stop at each region [14]. Therefore, the time it takes for each vehicle to travel between the two regions and the waiting time at the main bus stop are both random. In addition, a key metric of interest in our framework is the time interval between successive DTN-equipped vehicles' arrivals at each bus stop, as it directly impacts the data transmission rate and the effectiveness of the DTN system. The following subsections provide detailed modeling of these timing distributions and explain how they influence communication performance.

1) **Contact Time:** A vehicle arrives at the main bus stop and waits for passengers to board or alight from the vehicle. The waiting time for the vehicles varies depending on the time of the day and the availability of customers to travel to the intended destination. During this time, a DTN-equipped vehicle establishes a wireless connection with the DTN gateway located at the bus stop and starts to send/receive data packages with data rate R_{link} (in Mbps, representing the wireless link data rate between DTN-equipped vehicles and DTN gateways). We call this period the *contact time* between the vehicles and DTN gateway, and it determines the data size sent/received for each time the vehicles stop at the bus stop. In practice, the waiting time is random in nature and could be few minutes or tens of minutes; hence, we model the contact time, denoted as T_c , for each vehicle stopping near DTN gateways at each region (nodes A and B, see Fig 1) as uniform distribution $T_{c,\{A,B\}} \sim \text{Unif}(c_1, c_2)$, where c_1 and c_2 are the minimum and maximum observed waiting time for drivers at the bus stops, respectively. This typically captures the randomness of the time needed for passengers to board and alight from the vehicle and/or the drivers to take a break after each trip. In addition, the uniform distribution also assumes no inherent bias within the range, which reflects the variability caused by operational factors, and provides a straightforward and effective representation of the randomness observed in practice. We note that more sophisticated distributions could potentially offer greater accuracy in specific contexts, particu-

larly in highly dense or disrupted environments. However, such refinements would require extensive empirical data collection on actual waiting time patterns. Importantly, our analytical framework focuses on the expected value of system delays, where the mean contact time plays a dominant role. This focus ensures that our key conclusions remain robust regardless of whether the underlying distribution is uniform or follows alternative forms with comparable mean values.

2) **Travel Time:** Once passengers board or alight from the vehicle, it starts to travel to the other region. In practice, during the journey, the drivers often adjust their route and stop at intermediate locations dynamically, for example, to find more customers, circumvent traffic, or avoid police checkpoints [14]. Therefore, we model the time to travel from region A to B (and from B to A), denoted as $T_{\{AB,BA\}}$, to be a combination of two parts; T_{\min} , which is the minimum time needed for a vehicle to travel between the city and village (following an optimum path, and without considering any delay sources), and T_d , a delay that captures the drivers' intermediate stopping, traffic encountered during its journey, and any other sources of random delay. T_{\min} is constant and calculated based on the distance and optimum route between the city and the village, while T_d is modeled as exponentially distributed random variable $T_d \sim \text{Exp}(1/\bar{T}_d)$, where \bar{T}_d is the average delay of the travel time of the vehicle. We model the delay T_d as an exponentially distributed random variable because the exponential distribution is commonly used to capture random delays, reflecting the dynamic and unpredictable nature of informal transport systems, where delays arise due to independent factors such as intermediate stops, traffic, or detours. The exponential model also provides mathematical simplicity, allowing analytical traceability while investigating communication metrics for DTNs. Hence, $T_{\{AB,BA\}} = T_{\min} + T_d$, which means that $T_{\{AB,BA\}}$ is a shifted-exponentially distributed random variable with mean delay $T_{\min} + \bar{T}_d$.

3) **Inter-arrival time of vehicles:** In our scenario, we consider \mathcal{N} vehicles serving passengers between the city and village, of which n vehicles ($n \leq \mathcal{N}$) are equipped with DTN modules to deliver data packages. These vehicles traverse between the two regions throughout the day, with round-trip times modeled as described in Sections II-C1 and II-C2. One key metric in this analysis is the inter-arrival time of vehicles at each region, which directly impacts data transmission rates and communication metrics. Here, we define the *inter-arrival* time of a single vehicle as the time between two successive arrivals of the same DTN-equipped vehicle at a region (e.g., region A). This is equivalent to the vehicle's full round-trip time, encompassing its contact time at region A, travel to region B, contact time at region B, and return to region A. To model this, we represent each vehicle movement (the v -th vehicle) as a renewal process, denoted as $N_v(t)$, with independent and identically distributed inter-arrival times. The i -th inter-arrival time of the v -th vehicle is denoted by $T_{v,i}$. For a vehicle starting at any region (region A or B), the time to complete successive round trips follows the renewal pattern $T_{v,1}, T_{v,2}, \dots$, with each renewal time accounting for both travel and waiting durations. In our considered scenario, the

inter-arrival time $T_{v,i}$ is given as the time needed for the vehicle to complete a round-trip starting from one of the regions. Formally, we define the inter-arrival time of a DTN-equipped vehicle v as the total time taken to complete one full round-trip starting at node A (urban), travelling to node B (rural), and returning to node A. This includes the contact time at node A ($T_{c,A}$), the travel time from A to B (T_{AB}), the contact time at node B ($T_{c,B}$), and the return travel time from B to A (T_{BA}). Hence, the i -th inter-arrival time for the v -th DTN-equipped vehicle is expressed as:

$$T_{v,i} = T_{c,A} + T_{AB} + T_{c,B} + T_{BA} \quad (1)$$

where each term in the right-hand side of Equation (1) corresponds to the v -th vehicle during its i -th round trip, and the subscripts v and i are omitted for notational simplicity because the random variables ($T_{c,A}, T_{AB}, T_{c,B}, T_{BA}$) are assumed independent and identically distributed (i.i.d.) across vehicles and trips, with the understanding that each is implicitly indexed by (v, i) in analysis and simulation. To evaluate the aggregate performance of the system with n DTN-equipped vehicles, we extend the above analysis to include all n vehicles. Suppose that the travel time of each vehicle traversing between the two regions is independent of the others. We model the arrival time of n vehicles at region A (or B) as a superposition of n independent renewal processes, denoted by $N_s(t)$ and given as follows:

$$N_s(t) = N_1(t) + N_2(t) + \dots + N_n(t) \quad (2)$$

$N_s(t)$ in Equation (2) describes the superimposed arrival process from n vehicles as seen at each region, i.e., the i -th inter-arrival time of the superimposed process $N_s(t)$, denoted as $T_{s,i}$, represents the time elapsed between the arrival of any consecutive vehicles at both regions. In general, the superposition of independent renewal processes is not a renewal process, and finding the renewal time (inter-arrival time of aggregated vehicles) distribution of the superimposed process is not straightforward [18]. Nevertheless, we focus on limiting distribution for intervals remote from the time origin, or equivalently, we look at the component process as an equilibrium renewal process. In other words, we assume that the vehicles have been running long before it is first observed. The assumption of observing the process at equilibrium is reasonable because vehicles are typically in continuous operation over the day, and the DTN is likely to have reached a steady state by the time it is first observed. Evidently, assuming the superimposed process is observed at equilibrium allows for direct computation of the mean inter-arrival time of the superimposed process. To clarify, suppose that the mean inter-arrival time of each of the individual processes is given by $\mathbb{E}[T_{v,i}] = \mathbb{E}[T_v] = \mu$ (the i subscript was omitted as all inter-arrival times are assumed to be identically distributed across vehicles and round trips). It was shown in [18] that, at equilibrium, the mean inter-arrival time of the superimposed process in Equation (2), is given by:

$$\mathbb{E}[T_{s,i}] = \mathbb{E}[T_s] = \frac{\mathbb{E}[T_v]}{n} = \frac{\mu}{n}, \quad (3)$$

with arrival rate, λ_s , of the superimposed process is:

$$\lambda_s = \frac{1}{\mathbb{E}[T_s]} = \frac{n}{\mu} \quad (4)$$

This analytical relationship shown in Equation (4) not only confirms known intuitions but also enables a quantitative design framework for determining the minimum number of DTN-equipped vehicles, as to be discussed in the following section, required to meet specific communication performance goals, such as average data transmission rates and age of information thresholds, which is particularly useful to evaluate the effectiveness of DTNs. In addition, the distribution of the inter-arrival time of the superimposed process in Equation (2), denoted by $g(x)$, can be expressed as [18]:

$$g(x) = -\frac{d}{dx} \left[\mathcal{F}(x) \left(\int_x^\infty \frac{\mathcal{F}(u)}{\mu} du \right)^{n-1} \right], \quad (5)$$

where $\mathcal{F}(x)$ is the survival function of the inter-arrival time of the individual renewal process. In the above model, we assume that the DTN-equipped vehicles operate independently, and their inter-arrival times are drawn from independent and identically distributed renewal processes. This assumption simplifies the analysis and allows the derivation of closed-form expressions for key performance metrics. However, this assumption may not be applicable if the vehicles pass by congested urban environments or shared road segments, as vehicle delays may exhibit correlation due to common influences such as traffic jams, accidents, weather conditions, or scheduled events. Such dependencies could lead to bursty arrivals or prolonged service gaps, impacting the effectiveness of the DTN system. It is worth noting that modeling DTN-equipped vehicles as travelling independently is a reasonable approximation, particularly when the waiting times at bus stops are sufficiently long. In such cases, the variability introduced by independent boarding and alighting processes, along with driver behaviour, can help decouple successive trips, which further justifies our independence assumption.

III. PERFORMANCE METRICS

In this section, we provide an analysis of the performance metrics of the considered mule-based DTN. We focus on two key performance metrics critical to DTNs: (i) the down-link/uplink data transmission rates and (ii) the Peak Age of Information (PAoI). These metrics provide insight into the efficiency and timeliness of data delivery, which are crucial for assessing the feasibility and scalability of mule-based DTNs. We analyze each of these metrics to determine the minimum number of informal public transport vehicles that need to be equipped with DTN capabilities to meet specific performance requirements. We begin by discussing the wireless link data rate between DTN-equipped vehicles and DTN gateways during the contact time. Subsequently, we analyze the transmitted data package size per-each contact time and the data transmission rate at each node. Finally, we analyze the PAoI and derive an analytical approximation for the Mean Peak Age of Information (MPAoI), which represents the expected value of PAoI across multiple update cycles.

A. WIRELESS LINK DATA RATE

This subsection examines the physical layer wireless link data rate between DTN-equipped vehicles and fixed DTN gateways during their contact time. As vehicles approach the bus stop, their radio terminals begin communicating with the fixed DTN gateway. Although affected by mobility factors with expected variable link quality, modern wireless technologies offer coverage ranging from a few meters to several hundred meters [19], [20], [21]. The data rate improves as the vehicle and gateway move closer, significantly impacting the amount of data exchanged during the contact time. While characterizing the transient period during a vehicle's approach to the gateway is challenging due to rapidly changing link quality, such as the signal-to-noise ratios (SNR), most data transmission occurs when the vehicle stops near the DTN gateway. For simplicity, we assume a fixed wireless link data rate R_{link} (in Mbps) during the vehicle's contact time, representing the peak physical-layer transmission capacity when the vehicle is stationary. Additionally, we assume symmetrical uplink and downlink data rates during the contact time. This is a reasonable assumption, as modern Wi-Fi technologies share bandwidth between uplink and downlink transmissions using advanced multiple-access schemes [22].

B. DOWNLINK/UPLINK DATA TRANSMISSION RATE

In DTN systems, communication often involves long delays and is designed for scenarios where continuous connectivity cannot be guaranteed. To account for these delays, we consider the total time a vehicle spends on its journey, including its contact time with fixed gateways, as part of the data transmission process. Consequently, the data transmission rate in such networks is defined differently, reflecting the intermittent nature of data exchanges between nodes. Specifically, data is transmitted in discrete chunks whenever DTN-equipped vehicle establishes contact with fixed DTN gateways. We define the effective data transmission rate, denoted as \mathcal{R} , as the total amount of data transmitted or received (in Mbits) for uplink or downlink over an observation period. Note that the transmission is not continuous, and only a portion of the observed period involves active data exchange. This metric provides a meaningful representation of network performance, enabling a direct analysis of the effect of vehicle density on data transmission rate. To quantify the downlink and uplink transmission rates for our scenario, we leverage the established inter-arrival time of vehicles described in Section II-C3. First, we calculate the average transmitted data size for each vehicle v , denoted as m_v , during the contact time for both uplink and downlink.

Lemma 1 (Mean Transmitted Data Size). *The data size for uplink/downlink is proportional to the time needed to pick up the data package from region A (or B) and deliver it to region B (or A). Thus, the transmitted data size for the v -th vehicle during contact time is given by:*

$$m_v = R_{link} \cdot \min\{T_{c,A}, T_{c,B}\} \quad (6)$$

where the complementary cumulative distribution function (CCDF) of the data size is:

$$F_m(m) = \left(\frac{c_2 - m/R_{link}}{c_2 - c_1} \right)^2, \quad m \in [c_1 R_{link}, c_2 R_{link}] \quad (7)$$

where c_1 and c_2 are the minimum and maximum contact times at the bus stops, respectively, as defined in Section II-C1. The mean transmitted data size for uplink/downlink is:

$$\mathbb{E}[m_v] = R_{link} \cdot \left(\frac{2c_1 + c_2}{3} \right) \quad (8)$$

Proof: See Appendix -A. ■

Given the known inter-arrival time of the aggregated vehicles, the mean effective data transmission rate can be expressed as:

Lemma 2 (Mean Data Transmission Rate). *For the mean data size per vehicle arrival, $\mathbb{E}[m_v]$, and vehicles arrival rate, λ_s , the mean data transmission rate, denoted as $\mathbb{E}[\mathcal{R}]$, is given by:*

$$\mathbb{E}[\mathcal{R}] = \mathbb{E}[m_v] \lambda_s = \mathbb{E}[m_v] \cdot \frac{n}{\mu}, \quad (9)$$

This result shows that the mean data transmission rate is linearly proportional to the number of vehicles. This is consistent with Equation (4), where the rate of vehicle arrivals is also linearly proportional to the number of vehicles.

C. PEAK AGE OF INFORMATION

The concept of Age of Information (AoI) was introduced to quantify the freshness of knowledge about a process observed from a distance [23]. In other words, AoI measures the time elapsed since the last status update received from a source was generated. For clarity, assume a stochastic process $Y(t)$ is observed by a source node at the village, where samples of the process are extracted as time evolves. This process could represent the state of IoT devices, the status of a web-caching server, or any other service providing updates from the village. These samples, along with their timestamps t_i , are transmitted to a monitor node at the city network via DTN-equipped vehicles. The timestamp allows the monitor to track the freshness of the received samples. For example, if a sensor reading was created at time t_g and arrives at the monitor at time t_r , the AoI at the moment of reception is $t_r - t_g$. After the update is received, the AoI grows linearly over time until a new update arrives and resets the age. Unlike traditional delay metrics that measure how long it takes for a packet to travel from source to destination, AoI captures how timely the available information is, accounting for both delivery delay and the frequency of updates. This makes it especially relevant in DTN scenarios where updates may arrive sporadically due to intermittent connectivity. Lower AoI values indicate fresher information, which is important in applications with strict delay requirements. AoI can be mathematically defined as follows:

Definition 1 (Age of Information). *Consider a communication system comprising a source that generates data updates and a monitor that receives these updates. Let t_l denote the*

generation time of the l -th update at the source, and let t'_l denote the reception time of the l -th update at the monitor. At any time t , let $u(t)$ represent the generation timestamp of the most recent update received by the monitor. Specifically, if the monitor has received updates up to the ϑ -th update by time t , then $u(t) = t_\vartheta$, where $\vartheta = \max\{l : t'_l \leq t\}$. The Age of Information (AoI) at time t , denoted as $A(t)$, is defined as the time elapsed since the generation of the most recently received update:

$$A(t) = t - u(t). \quad (10)$$

The AoI quantifies the freshness of information available at the monitor. When a new update is received at time t'_l , the AoI instantaneously drops to $A(t'_l) = t'_l - t_l$, which represents the end-to-end delay of that update. Between consecutive update arrivals, the AoI increases linearly with time, creating a sawtooth pattern.

Another key metric of interest is the Peak Age of Information (PAoI), which describes the maximum AoI observed just before the arrival of a new update. PAoI can be considered as the upper bound of the AoI (worst-case data freshness) and is considered as the metric to measure the freshness of information in systems, as it typically provides more tractable solutions. Formally, PAoI is defined as:

Definition 2 (Peak Age of Information). *Following the definition of AoI in 1, the PAoI when receiving the l -th update is given by:*

$$PAoI[l] = t'_l - t_{l-1}, \quad (11)$$

where t'_l is the time the l -th update is received and t_{l-1} is the generation time of the previous update.

Based on Definitions 1 and 2, a sample realization of the AoI in a system with two DTN-equipped vehicles operating between two regions illustrating the PAoI peaks (A_1, A_2, \dots) from the perspective of a monitor node at the city receiving updates from the village is shown in Fig. 2. At any time t , the AoI grows linearly until a fresh update is delivered to the monitor node, where PAoI occurs.

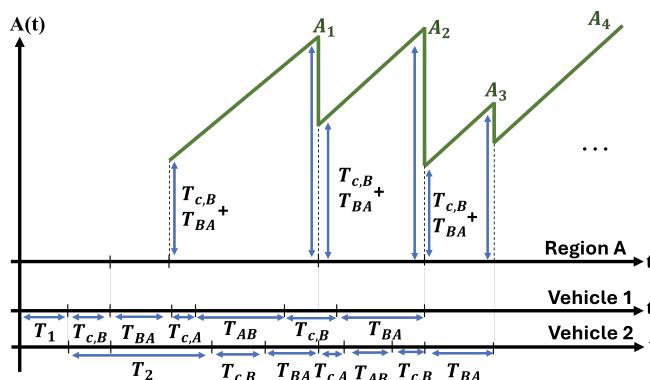


Fig. 2: A sample realization of AoI and PAoI for a system with two DTN-equipped vehicles. The peaks A_1, A_2, \dots correspond to the PAoI values.

To analyze the PAoI in our considered scenario, the following assumptions are made for the system under consideration.

First, each DTN terminal (at the city, village, and mules) has sufficiently large buffers, ensuring no packet loss due to buffer overflow. This assumption is motivated by the availability of high-capacity and low-cost storage units. Second, the contact time during a vehicle's stop at a node is long enough to allow the transmission of a complete update packet. This ensures that the updates are appropriately sized to fit within the contact time. For a concrete example, we refer to Fig. 2 and consider observing the system starting at $t = T_1$, with vehicle 1 at region B and vehicle 2 arriving later at $t = T_1 + T_2$. Here, T_1 and T_2 are arbitrary initial times and do not affect the system's average performance. Vehicle 1 returns to region A after collecting data from region B at $t = T_1 + T_{c,B} + T_{BA}$. At this point, the AoI reaches $T_{c,B} + T_{BA}$. The AoI continues to grow until either vehicle 1 completes a round trip or vehicle 2 delivers a fresher update. In Fig. 2, vehicle 2 arrives first, and the first PAoI is $A_1 = T_2 + T_{c,B} + T_{BA}$. The AoI then drops to $T_{c,B} + T_{BA}$ and resumes its linear growth until the next update arrives.

In general, obtaining a closed-form expression for the PAoI is challenging due to the inherent randomness of vehicle trip durations and the possibility that a vehicle may carry outdated updates if another vehicle delivers fresher data earlier. To make the analysis tractable, we assume that vehicles complete their trips sequentially, which aligns with typical operational patterns in informal public transport and is supported by the relatively long waiting times observed at main bus stops. Under this assumption, we leverage the results derived for the inter-arrival time in a system with n DTN-equipped vehicles and observe that the PAoI depends primarily on three factors: the contact time at the rural bus stop ($T_{c,B}$), the travel time required for the data package to move between the rural village and the city (T_{BA}), and the difference in arrival times between two consecutive DTN-equipped vehicles, which follows the inter-arrival time distribution of the superimposed renewal process (T_s). Accordingly, the PAoI can be approximated as $PAoI \approx T_{c,B} + T_{BA} + T_s$. Additionally, the Mean PAoI (MPAoI), denoted as $\mathbb{E}[PAoI]$, is expressed as $\mathbb{E}[T_{c,B} + T_{BA} + T_s]$. We verify how close this approximation is to the actual quantity in the simulation section. Moreover, as the number of vehicles increases, the MPAoI is expected to decrease and approach $T_{c,B} + T_{BA}$. Specifically, when the mean inter-arrival time of vehicles is reduced, the MPAoI follows. Thus, the MPAoI can be approximated as:

$$MPAoI \approx \mathbb{E}[T_{c,B} + T_{BA}] + \frac{\mu}{n}, \quad (12)$$

This relationship highlights the effect of deploying more DTN mule nodes at vehicles to minimize the system's AoI (and PAoI).

D. DATA TRAFFIC REQUIREMENTS

In mule-based DTN systems, the amount of data exchanged between nodes primarily depends on the contact time during vehicle stops. Estimating the data traffic generated by users at the village is critical to optimizing system components, including storage capacities and wireless access points. DTN

systems prioritize reliability over latency, favouring non-real-time applications such as email [24], web-caching [25], asynchronous messaging [26], and asynchronous digital learning platforms [5]. Additionally, DTNs can support IoT applications and collect sensor data from remote areas [15], [16], [17], [27], where cost-effectiveness and scalability play a vital role for resource-constrained IoT ecosystems. As discussed, DTN systems are most effective when the data traffic does not depend on an immediate response. Suitable IoT examples include batch collection of sensor readings for agriculture, wildlife tracking, and long-term environmental monitoring. For these use cases, periodic uploads and opportunistic data transfer are sufficient. On the other hand, applications that require real-time communication or immediate action, such as critical alerts or industrial controls, fall outside the range of services supported by the mule-based DTN designs.

The total data traffic generated at the village depends on the user density, denoted as u , and the traffic demand per user, denoted as d in Mbits/day. Typical estimates suggest that email and asynchronous chat applications generate 10–50 Mbits/day per user, while digital learning platforms may produce hundreds of Mbits/day [28]. Multimedia content requests (e.g., video streaming) generate significantly higher traffic, between 300–700 Mbits/hour, with general web browsing consuming around 50 Mbits/hour [29], [30]. Based on these estimates, the daily data demand per user could reach approximately 1 Gbits/day, though this depends on specific DTN applications and user activity levels. To meet the data traffic demands of a village, the total daily traffic, denoted as \mathcal{D}_{day} , can be calculated as:

$$\mathcal{D}_{\text{day}} = u \cdot d \quad (\text{Mbits/day}). \quad (13)$$

From this, the required number of vehicles to be equipped with DTN can be estimated, considering their capacity and frequency of travel. This calculation informs decisions regarding the deployment and configuration of DTN infrastructure, ensuring sufficient resources for data exchange.

E. COST CONSIDERATIONS

The proposed DTN architecture consists of two types of nodes: fixed DTN gateways located at urban and rural village bus stops and mobile DTN nodes (data mules) mounted on public transport vehicles. A typical DTN node comprises a processing unit (e.g., a small server or computer), a wireless adapter (e.g., Wi-Fi), a storage unit, a power supply with battery backup, and an enclosure for all components. The cost of each node ranges from a few hundred to over a thousand dollars depending on the hardware configuration [9]. In practice, the cost of a DTN node depends heavily on its hardware configuration. Low-cost designs based on devices such as Raspberry Pi boards equipped with Wi-Fi dongles and microSD or SSD storage can bring the total cost to under a few hundred USD. For mobile DTN nodes mounted on vehicles, power can be drawn directly from the vehicle's electrical system, eliminating the need for standalone battery systems or solar panels and thereby reducing both initial costs and maintenance requirements. This vehicle-powered approach is particularly practical since the

nodes only need to operate during vehicle contact times. On the other hand, fixed DTN gateways at rural locations without reliable grid power may require solar panels with battery backup to ensure continuous operation. Additionally, integrated systems with rugged enclosures, onboard batteries or solar panels, and pre-installed software stacks may raise the cost of the DTN terminals. Our framework is flexible and supports a wide range of cost parameters to accommodate this variability. While connecting the urban DTN gateway to the Internet incurs additional expenses, most costs are one-time investments amortized over the network's operational lifespan. In this DTN system, the primary variable cost is associated with deploying the mobile DTN nodes. To minimize these costs, we aim to determine the minimum number of mobile DTN nodes required to meet specific communication performance metrics. Assuming the cost of a single DTN node placed on vehicles is denoted as C , and for the fixed DTN gateways denoted as C_{gateway} , the total deployment cost can be formulated as an optimization problem while meeting QoS requirements, including MPAoI and mean data transmission rate. The problem is formulated as follows:

$$\min_n \quad n \cdot C + 2 \cdot C_{\text{gateway}}, \quad (14)$$

$$\text{s.t.} \quad \mathbb{E}[T_{c,B} + T_{BA}] + \frac{\mu}{n} \leq \mathbb{E}[\text{PAoI}]_{\text{thr}}, \quad (15)$$

$$\mathbb{E}[m_v] \cdot \frac{n}{\mu} \geq \mathbb{E}[\mathcal{R}]_{\text{thr}}, \quad (16)$$

where $\mathbb{E}[\text{PAoI}]_{\text{thr}}$ is the threshold for the MPAoI, and $\mathbb{E}[\mathcal{R}]_{\text{thr}}$ is the required average data transmission rate. The constraints in Equations (15) and (16) are to satisfy the QoS requirements presented by the MPAoI and mean data transmission rate, respectively. The optimal solution for the optimization problem with objective function in Equation (14) and constraints in Equations (15) and (16) is straightforward. Let $\alpha = \frac{\mu}{\mathbb{E}[\text{PAoI}]_{\text{thr}} - \mathbb{E}[T_{c,B} + T_{BA}]}$ and $\beta = \frac{\mathbb{E}[\mathcal{R}]_{\text{thr}} \cdot \mu}{\mathbb{E}[m_v]}$. The minimum number of mobile DTN nodes required to satisfy both constraints at the same time minimizes the total cost is:

$$n_{\text{opt}} = \lceil \max\{\alpha, \beta\} \rceil. \quad (17)$$

This formulation provides a cost-effective approach to deploying a DTN system while ensuring that communication performance metrics are met.

IV. NUMERICAL RESULTS

In this section, we validate the investigated framework through numerical simulations. We first present results under various system parameters and demonstrate how closely the proposed MPAoI approximation aligns with simulation outcomes. Additionally, we utilize datasets from informal public transport systems in Nouakchott, Accra, and Addis Ababa, to evaluate system constraints in realistic scenarios.

To validate the MPAoI approximation derived in Equation (12), we developed a simulation framework based on the probabilistic model described in Section II-C. The main simulation parameters are summarized in Table II, where we used the same waiting time parameters (c_1 and c_2) for regions

TABLE II: Simulation Parameters

Parameter	Value
c_1	3 minutes
c_2	5 minutes
T_{\min}	100 minutes
\bar{T}_d	20 minutes
R_{link}	10 Mbps

A and B, and the same minimum delay parameter (T_{\min}) for both T_{AB} and T_{BA} . The framework simulates discrete-time events over a maximum simulation duration of 10^5 time units, with each unit representing one minute. This is to ensure steady-state metrics of the considered scenario. The mobility of DTN-equipped vehicles in our simulations is modeled using the probabilistic framework described in Section II-C, where round-trip durations are composed of random contact times (modeled as uniform distributions) and travel times (modeled as shifted exponential distributions). For validation, we incorporate real-world trace data from [14] to extract travel time statistics and compare performance across different urban-rural scenarios. We calculate and plot the Mean AoI (MAoI), the MPAoI, and the approximated MPAoI (labeled 'MPAoI Approx.' in the figure), based on Equation (12). The results are shown in Fig. 3. For a network with DTN-equipped vehicles ranging from 1 to 20, the maximum approximation error for the proposed MPAoI calculation is observed to be 4.66 minutes, demonstrating high accuracy for practical applications. The maximum approximation error represents the largest observed difference between the analytical approximation in Equation (12) and the simulation-based MPAoI across all vehicle configurations tested (1 to 20 vehicles). Additionally, in a network with a single DTN-equipped data mule, the MPAoI is approximately 375 minutes. However, as the number of data mules increases to 20, the MPAoI decreases to around 150 minutes. This substantial reduction shows the effectiveness of deploying additional DTN-equipped vehicles in improving data freshness and overall network efficiency. It can be also noted from Fig. 3 that, as the number of data mules increases, the MPAoI approaches its theoretical minimum (red dashed line in Fig. 3), which corresponds to the average time required for a vehicle to complete a one-way trip from the village to the city, T_{oneway} , i.e., the age of the data package just as it arrives at the monitor at the city while generated T_{oneway} minutes before at the village.

Fig. 4 illustrates the impact of round-trip times on the mean data transmission rate and MPAoI. In this analysis, the mean minimum contact time at each bus stop is set to 10 minutes, and the wireless link data rate is assumed to be 10 Mbps. While modern Wi-Fi standards such as IEEE 802.11n/ac/ax can support data rates exceeding 600 Mbps [31], we conservatively assume a fixed average rate of 10 Mbps during the vehicle's contact time with DTN gateways. This reflects a more realistic data rate achievable in rural deployments where interference, energy limitations, and hardware constraints may affect link quality. The assumption accounts for average performance during the stationary phase of the vehicle's stop, when most data transmission occurs,

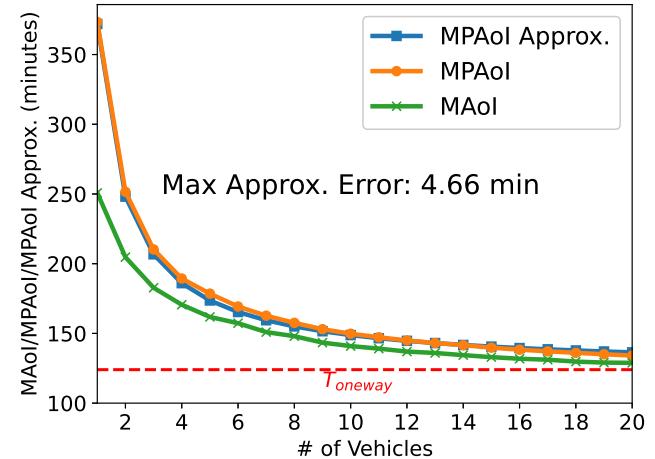


Fig. 3: MAoI, MPAoI, and MPAoI approximation as described in Equation (12). The simulation was averaged over 10^5 time units. The maximum approximation error represents the largest observed difference between the analytical approximation in Equation (12) and the simulation-based MPAoI across all vehicle configurations tested (1 to 20 vehicles). The red dashed line represents the average time required for a vehicle to complete a one-way trip from the village to the city (T_{oneway}).

and serves as a practical baseline for performance evaluation. Importantly, the analytical framework and simulation results remain valid for other data rate values, allowing adaptation to more bandwidth-constrained or unreliable rural environments. Based on Equation (8), the mean data package size per vehicle arrival is around 6 GB. From the figure, we observe that the transmission rate decreases exponentially as the mean round-trip time increases, consistent with the earlier analysis. Simultaneously, the MPAoI grows linearly with the round-trip time. Furthermore, as the number of DTN-equipped data mules increases, both the transmission rate and the MPAoI show significant improvement. For instance, in a network with a single DTN-equipped data mule and a mean round-trip time of 120 minutes, the transmission rate is approximately 2 Mbps, while the MPAoI is approximately 200 minutes. However, with 20 DTN-equipped data mules in the network, the transmission rate increases to around 20 Mbps, and the MPAoI drops to approximately 75 minutes. These results highlight that adding more data mules to the network improves the transmission rate linearly while reducing the MPAoI.

In Figs. 5 and 6, we utilize a recent dataset collected for informal public transport systems across 22 countries globally [14]. The dataset includes mobility traces for several routes, obtained by equipping vehicles with GPS transceivers to record the time required to complete their routes. While the dataset does not specifically provide information on vehicles starting and ending at bus stops or waiting times at the beginning and end of routes, it serves as a representative model of the randomness and delays commonly encountered in informal public transport systems when traveling between different locations. For this study, we selected the three longest

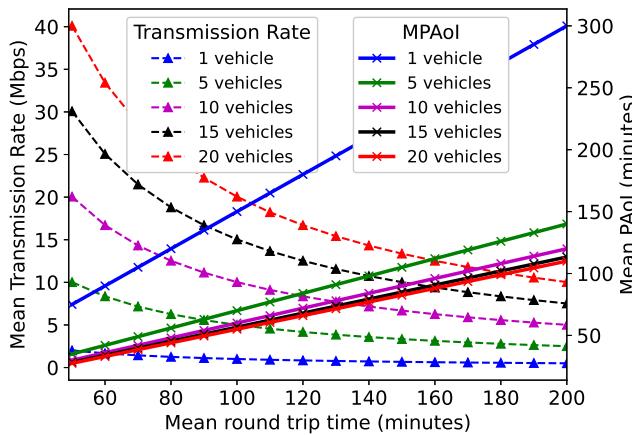


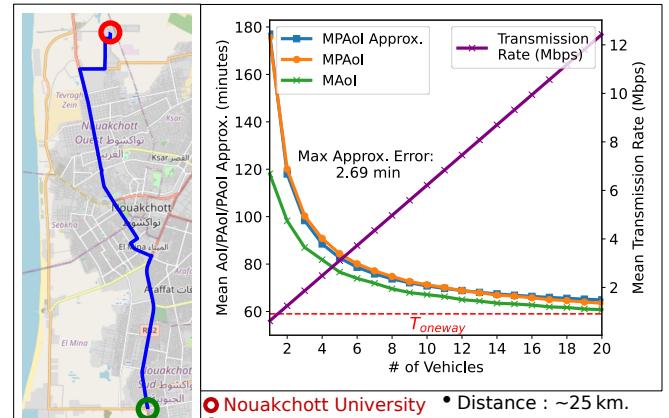
Fig. 4: MPAoI and mean data transmission rate for different mean round-trip times and DTN-equipped vehicles in the network.

routes in Nouakchott (capital of Mauritania), Accra (capital of Ghana), and Addis Ababa (capital of Ethiopia). Although some of the areas represented in our dataset have access to other connectivity solutions, we utilized this dataset as a means to demonstrate the effectiveness of our mule-based DTN framework for potentially bringing connectivity to rural areas. We used the same contact time limits as listed in Table II, while the travel time statistics were derived from the dataset and the wireless link data rate is set to 20 Mbps.

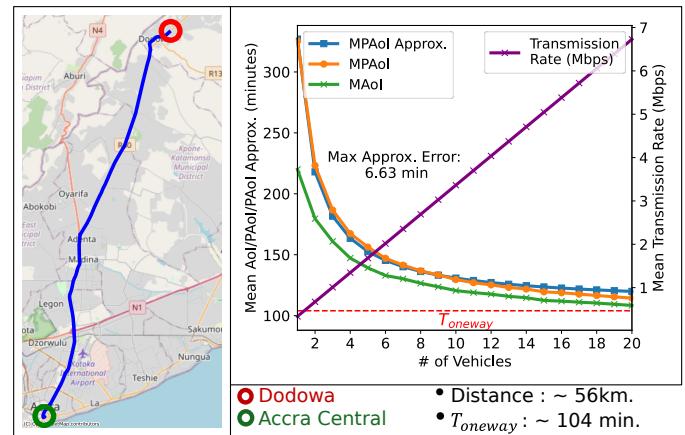
Fig. 5(a) shows the results of our proposed framework applied to the route connecting Nouakchott University and *Bamako Crossroads*. This route spans approximately 25 km and, under optimal conditions, takes around 30 minutes to complete a one-way trip. However, based on the dataset, informal transport vehicles take an average of 59 minutes to complete the one-way journey, accounting for random delays caused by various factors as analyzed earlier. The figure also illustrates the expected data transmission rate and MPAoI based on our framework. For a network with a single DTN data mule, the data transmission rate is around 1 Mbps, while it increases to approximately 12.5 Mbps with 20 data mules. Similarly, the MPAoI for one data mule is approximately 180 minutes, which decreases to around 65 minutes when 20 data mules are introduced into the network. These results highlight the interplay between the number of DTN data mules in the network and the resulting performance.

A similar analysis is conducted for the route between *Dodowa* and *Accra Central* in Accra, as shown in Fig. 5(b). This route is approximately 56 km long and requires around 55 minutes for a one-way trip under optimal conditions. However, based on the dataset, vehicles take an average of 104 minutes due to random delays. Because this route is longer, the network performance is slightly lower compared to the shorter route in *Nouakchott* when using the same number of vehicles. For instance, with one DTN data mule, the transmission rate is less than 1 Mbps, and the MPAoI is approximately 350 minutes. With 20 DTN data mules, the data transmission rate improves to around 7 Mbps, while the MPAoI decreases to

approximately 120 minutes.



(a)



(b)

Fig. 5: Network performance metrics (Transmission Rate and MPAoI) for the routes (a) between *Nouakchott University* and *Bamako Crossroads*, and (b) *Dodowa* and *Accra Central*, considering varying numbers of DTN data mules.

Finally, we consider a scenario involving a central hub connecting multiple routes in Addis Ababa. This represents a central bus stop, *Shero Meda*, serving as a hub for routes of varying lengths and travel durations. The hub connects routes to *Tulu Dimtu* (Route 1: 123 minutes), *Mexico* (Route 2: 73 minutes), *Bole Millennium* (Route 3: 61 minutes), *Abo Junction* (Route 4: 55 minutes), and *Autobis Tera* (Route 5: 26 minutes), with the average one-way trip durations provided in parentheses. This can be modeled as a star-shaped DTN where the gateway at *Shero Meda* serves multiple rural areas simultaneously. The performance metrics for this setup are shown in Fig. 6. For the route with the longest travel time (*Tulu Dimtu*), the data transmission rate is less than 1 Mbps with a single DTN data mule, increasing to approximately 8 Mbps with 20 data mules. Similarly, the MPAoI decreases from 380 minutes with one data mule to around 150 minutes with 20 data mules. For shorter routes, consistent improvements in both transmission rate and MPAoI are observed as the number of data mules increases.

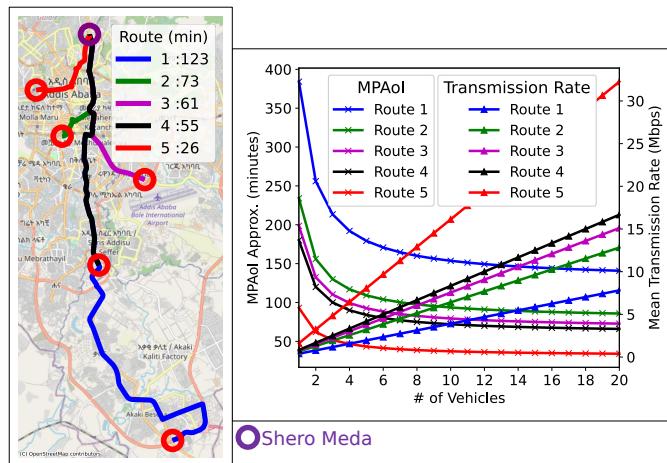


Fig. 6: Performance evaluation of a star-shaped DTN with a central hub at *Shero Meda*, connecting multiple routes in Addis Ababa with varying travel durations. The figure shows the impact of the number of DTN data mules on the data transmission rate and MPAoI for different routes, highlighting the interplay between network performance and route length.

V. CONCLUSION

This paper presented a DTN-based framework leveraging informal public transport systems to address the persistent digital divide in rural areas. This framework offers a cost-effective and scalable solution for enabling intermittent data exchange between urban regions with Internet access and rural regions without connectivity. A probabilistic model was developed to characterize the movement of data mules, incorporating realistic travel and contact time distributions. Key performance metrics, including data transmission rates and Mean Peak Age of Information (MPAoI), were analyzed, and a mathematical approximation for the MPAoI was derived and validated through numerical simulations. The framework was further tested using real-world datasets from Nouakchott, Accra, and Addis Ababa, demonstrating its applicability in diverse settings. The results quantify how the number of DTN-equipped vehicles and the route length influence achievable data transmission rates and data freshness, showing that additional data mules can improve performance even in resource-constrained settings.

While the investigated framework is well-suited for regions with limited infrastructure, its performance may decrease when vehicles travel very long distances between rural and urban areas, resulting in higher round-trip times and increased MPAoI. To address this limitation, future work could explore hybrid networking strategies that combine opportunistic DTN connections with infrastructure-based links, such as satellite backhaul, to maintain acceptable performance in remote or geographically isolated regions.

VI. ACKNOWLEDGEMENT

The authors acknowledge support from the Arab Fund Pro-

gramme at ICTP (ARF01 - AFESD Grant No. 14/2023) in the framework of the project “Advancing the Capabilities of Arab Researchers and Students” and from the ICTP/IAEA Sandwich Training Educational Programme.

APPENDIX

A. PROOF OF LEMMA 1

Let $X = \min(T_{c,A}, T_{c,B})$, the CDF and PDF of the minimum of two i.i.d uniformly distributed random variables is:

$$F_X(x) = P(X \leq x) = 1 - \left(\frac{c_2 - x}{c_2 - c_1} \right)^2, \quad x \in [c_1, c_2].$$

$$f_X(x) = \frac{d}{dx} F_X(x) = \frac{2(c_2 - x)}{(c_2 - c_1)^2}, \quad x \in [c_1, c_2].$$

The mean of X can be evaluated as:

$$E[X] = \int_{c_1}^{c_2} x f_X(x) dx = \frac{2c_1 + c_2}{3}.$$

Hence, this completes the proof.

REFERENCES

- [1] “Report of the secretary-general on our common agenda,” [Online]. Available: <https://www.un.org/en/content/common-agenda-report/>, accessed: Nov. 26, 2024.
- [2] “Global connectivity report,” [Online]. Available: <https://www.itu.int/hub/publication/d-ind-global-01-2022/>, accessed: Nov. 26, 2024.
- [3] E. Yaacoub and M.-S. Alouini, “A key 6G challenge and opportunity—connecting the base of the pyramid: A survey on rural connectivity,” *Proceedings of the IEEE*, vol. 108, no. 4, pp. 533–582, Apr. 2020.
- [4] S. Perumal, V. Raman, G. N. Samy, B. Shanmugam, K. Kisenasamy, and S. Ponnan, “Enhanced disruption tolerant network (DTN) framework for improving network efficiency in rural areas,” *International Journal of System Assurance Engineering and Management*, vol. 13, no. Suppl 1, pp. 710–717, Jan. 2022.
- [5] S. Abdeljabar and M.-S. Alouini, “Bridging the educational divide: A delay-tolerant networking approach for equitable digital learning in rural areas,” Nov. 2025. [Online]. Available: <https://arxiv.org/abs/2511.20334>
- [6] I. Ndbatya and M. J. Booyse, “Modelling of inter-stop minibus taxi movements: Using machine learning and network theory,” in *Proceedings of the First International Conference on the use of Mobile Information and Communication Technology (ICT) in Africa (UMICTA 2014)*. Stellenbosch, South Africa: Stellenbosch University, Department of Electrical & Electronic Engineering, Dec. 2014, pp. 9–10.
- [7] A. Rahman, T. Nakanishi, and A. Fukuda, “Delay tolerant network for developing countries,” in *2013 International Conference on Informatics, Electronics and Vision (ICIEV)*, May 2013, pp. 1–6.
- [8] A. Galati, T. Bourchais, S. Siby, S. Frey, M. Olivares, and S. Mangold, “Mobile-enabled delay tolerant networking in rural developing regions,” in *IEEE Global Humanitarian Technology Conference (GHTC 2014)*, Oct. 2014, pp. 699–705.
- [9] S. Grasic and A. Lindgren, “Revisiting a remote village scenario and its DTN routing objective,” *Computer Communications*, vol. 48, pp. 133–140, Jul. 2014.
- [10] H. Ntareme, M. Zennaro, and B. Pehrson, “Delay tolerant network on smartphones: Applications for communication challenged areas,” in *Proceedings of the 3rd Extreme Conference on Communication: The Amazon Expedition*, Sep. 2011, pp. 1–6.
- [11] A. Pentland, R. Fletcher, and A. Hasson, “DakNet: Rethinking connectivity in developing nations,” *Computer*, vol. 37, no. 1, pp. 78–83, Jan. 2004.
- [12] 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 SIGCOMM Computer Communication Review*, vol. 37, no. 5, pp. 95–100, Oct. 2007.

- [13] M. M. Luo and S. Chea, "Internet Village Motoman Project in rural Cambodia: Bridging the digital divide," *Information Technology & People*, vol. 31, no. 1, pp. 2–20, Feb. 2018.
- [14] K. M. Mittal, M. Timme, and M. Schröder, "Efficient self-organization of informal public transport networks," *Nature Communications*, vol. 15, no. 1, p. 4910, Jun. 2024.
- [15] F. Z. Benhamida, D. Casado-Mansilla, C. Bennani, and D. López-de Ipiña, "Toward a delay tolerant Internet of Things," in *Proceedings of the 9th International Conference on the Internet of Things*, Oct. 2019, pp. 1–4.
- [16] M. Li, F. R. Yu, P. Si, W. Wu, and Y. Zhang, "Resource optimization for delay-tolerant data in blockchain-enabled IoT with edge computing: A deep reinforcement learning approach," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9399–9412, Jul. 2020.
- [17] S. Bounsiar, F. Z. Benhamida, A. Henni, D. L. d. Ipiña, and D. C. Mansilla, "How to enable delay tolerant network solutions for Internet of Things: From taxonomy to open challenges," *Proceedings*, vol. 31, no. 1, Nov. 2019.
- [18] D. R. Cox, *Renewal theory*. London: Methuen, 1962, catalogue No. (Methuen) 2/5238/11.
- [19] P. Rani and R. Sharma, "An experimental study of IEEE 802.11n devices for vehicular networks with various propagation loss models," in *International Conference on Signal Processing and Integrated Networks*. Springer Nature Singapore, Jun. 2023, pp. 125–135.
- [20] Z. El Khaled, H. McHeick, and F. Petrillo, "WiFi coverage range characterization for smart space applications," in *2019 IEEE/ACM 1st International Workshop on Software Engineering Research & Practices for the Internet of Things (SERP4IoT)*, May 2019, pp. 61–68.
- [21] M. Rjab, A. Omri, S. Bouallegue, H. Chamkhia, and R. Bouallegue, "Modeling and performance analysis of mmWave and WiFi-based vehicle communications," *Electronics*, vol. 13, no. 7, p. 1344, Apr. 2024.
- [22] E. Khorov, A. Kiryanov, A. Lyakhov, and G. Bianchi, "A tutorial on IEEE 802.11ax high efficiency WLANs," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 197–216, Sep. 2019.
- [23] G. Kallitsis, V. Karyotis, and S. Papavassiliou, "On the potential of enhancing delay-tolerant routing protocols via age of information," *Future Internet*, vol. 14, no. 8, p. 242, Aug. 2022.
- [24] A. Lindgren and A. Doria, "Experiences from deploying a real-life DTN system," in *2007 4th IEEE Consumer Communications and Networking Conference*, Jan. 2007, pp. 217–221.
- [25] J. Näslund, "Developing delay-tolerant networking applications for Arctic communities," Ph.D. dissertation, Luleå Tekniska Univ., Luleå, Sweden, 2013.
- [26] S. Grasic, "Not-so-instant-messaging service for the delay tolerant networks," in *International Multiconference Information Society*. Institut Jožef Stefan, Oct. 2011, pp. 203–207.
- [27] C.-A. Sarros, V. Demiroglou, and V. Tsoussidis, "Intermittently-connected IoT devices: Experiments with an NDN-DTN architecture," in *2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC)*, Jan. 2021, pp. 1–9.
- [28] "Pentel data bandwidth calculator," [Online]. Available: <https://www.penteledata.net/calculator>, accessed: Nov. 26, 2024.
- [29] "Data usage calculator," [Online]. Available: <https://www.omiccalculator.com/everyday-life/data-usage>, accessed: Nov. 26, 2024.
- [30] "Internet data calculator," [Online]. Available: <https://www.att.com/support/pages/data-calculator/>, accessed: Nov. 26, 2024.
- [31] E. J. Oughton, G. Geraci, M. Polese, M. Ghosh, W. Webb, and D. Bubley, "The future of wireless broadband in the peak smartphone era: 6G, Wi-Fi 7, and Wi-Fi 8," *IEEE Wireless Communications*, vol. 32, no. 5, pp. 245–252, Apr. 2025.



Marco Zennaro (Senior Member, IEEE) received the Ph.D. degree from the KTH-Royal Institute of Technology, Stockholm, Sweden, and the M.Sc. degree in electronic engineering from the University of Trieste, Trieste, Italy. He is a Researcher with the Abdus Salam International Centre for Theoretical Physics, Trieste, where he coordinates the Wireless Group of the Science, Technology and Innovation Unit. He is a Visiting Professor with the KIC-Kobe Institute of Computing, Kobe, Japan. His research interests include ICT4D, the use of ICT for development, and in particular he investigates the use of IoT in developing countries. He has given lectures on wireless technologies in more than 30 countries.



Mohamed-Slim Alouini (Fellow, IEEE) was born in Tunis, Tunisia. He received the Ph.D. degree in electrical engineering from the California Institute of Technology, Pasadena, CA, USA, in 1998. He served as a Faculty Member with the University of Minnesota, Minneapolis, MN, USA, then with Texas A&M University at Qatar, Doha, Qatar, before joining the King Abdullah University of Science and Technology, Thuwal, Makkah, Saudi Arabia, as a Professor of Electrical Engineering in 2009. His current research interests include the modeling, design, and performance analysis of wireless communication systems.



Salah Abdeljabar (Graduate Student Member, IEEE) received the B.Sc. degree in electrical engineering from The University of Jordan, Amman, Jordan, in 2019, and the M.Sc. degree in electrical and computer engineering from the King Abdullah University of Science and Technology, Thuwal, Saudi Arabia, in 2023, where he is currently pursuing the Ph.D. degree. His research interests include long-range (LoRa) communication, Delay tolerant networking (DTN), and optical wireless communications systems.