

Towards a simple and efficient vehicular delay tolerant networks routing protocol for data collection in smart cities

Ngurah Indra Er¹, Kamal Deep Singh^{2,3}, Christophe Couturier^{4,5}, Jean-Marie Bonnin^{4,5}

¹Department of Electrical Engineering, Faculty of Engineering, Udayana University, Bali, Indonesia

²Data Intelligence Research Team, Laboratoire Hubert Curien, Saint-Etienne, France

³Department of Computer Science, Faculty of Science, Université Jean Monnet, Saint-Etienne, France

⁴Department of Networks, Telecommunication and Services, IRISA, Rennes, France

⁵Department of Network Systems, Internet Security and Digital Law, IMT Atlantique, Rennes, France

Article Info

Article history:

Received Dec 16, 2022

Revised Sep 9, 2023

Accepted Sep 26, 2023

Keywords:

Data collection

Intelligent transportation systems

Routing protocol

Smart city

Vehicle-to-everything

Vehicular delay tolerant networks

ABSTRACT

Smart cities today can utilize vehicular delay tolerant networks (VDTN) to collect data from connected-objects in the environment for various delay-tolerant applications. They can take advantage of the available intelligent transportation systems (ITS) infrastructures to deliver data to the central server. The system can also exploit multiple and diverse mobility patterns found in cities, such as privately owned cars, taxis, public buses, and trams, along with their vehicle-to-everything (V2X) communications capabilities. In the envisioned convergence between the ITS and V2X, we believe that a simple and efficient routing protocol can be deployed for the delay-tolerant data delivery, contrary to the implementation of optimized solutions that might be resource-demanding and difficult to standardize. In this paper, we analyzed the performances of four baseline VDTN routing protocols, namely: direct delivery, first contact, epidemic, and spray and wait, to understand their strengths and weaknesses. Our simulation results highlighted the trade-off between distinct approaches used by those protocols and pointed out some gaps that can be refined. This study provides new interesting ideas and arguments towards developing a simple, efficient, and high-performing routing protocol for data collection in smart cities.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Ngurah Indra Er

Department of Electrical Engineering, Faculty of Engineering, Udayana University

Bali, Indonesia

Email: indra@unud.ac.id

1. INTRODUCTION

The current and future smart cities will require data collection solutions from various connected-objects in their area. This in turn will help cities to manage natural resources intelligently, to ensure sustainable socio-economic development, and ultimately to enhance quality-of-life. The data collection's primary trend is to connect all sensors to a long-range operated network such as the cellular network and the low-power wide area network (LPWAN) [1], [2]. But as the number of sensors and the amount of data generated grows exponentially [3], [4], the utilization of a cellular network to collect all kinds of data might not be economically feasible. At the same time, the use of LPWAN might be impeded by insufficient bandwidth. Therefore, it is desirable to keep such networks for only collecting data which has delay constraints. Whereas, other economical solutions should be utilized to collect delay-tolerant data.

Simultaneously, advances in vehicular technology have enabled the vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), and vehicle-to-everything (V2X) communications to connected-objects in their surroundings [5]. Vehicles with current and future radio access networks will also have the capability to connect to the internet [6]. Furthermore, the regulatory body will soon make such communication capabilities compulsory for vehicles to support the safety-related applications [7]. This exciting new development will pave the way for vehicles to participate actively in the ecosystem of smart cities [8], [9].

In terms of networking, delay tolerant networks (DTN) enable communication even when there are connectivity issues. Such issues could be sparse and intermittent connectivity, long and variable delay, high latency, high error rates, highly asymmetric data rate, and even no end-to-end connectivity [10]. The extension of vehicular ad hoc networks (VANET) with DTN capabilities naturally leads to the concept of vehicular delay tolerant networks (VDTN), where the vehicular networks will be able to cope with the inherent intermittent connectivity during data delivery.

The intelligent transportation systems (ITS), on the other hand, are advanced applications aiming to provide innovative services related to different modes of transport and traffic management, through vehicular communication, to improve road safety and comfort for drivers and passengers [11]. Vehicles equipped with wireless devices can collect data from the environment and exchange traffic and road safety information with nearby vehicles, roadside units, and other connected objects. These functionalities are known as the vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), and the vehicle-to-everything (V2X) communications [5]. Concerning the implementation of VDTN in the ITS environment, we can expect an increasing number of ITS infrastructures available to assist in exchanging data within the networks. One crucial infrastructure element is the road side unit (RSU), which can connect to the internet. Such units will be installed throughout the city, and they can function as a point-of-presence (PoP) for accessing the internet and forwarding data from vehicles to the core network. In another scenario that we do not consider in this paper, the RSU might not have an internet connection, but it can still participate in data forwarding as long as it has some wireless communication capability.

The PoPs will be available in numbers and they will be strategically placed, such as at traffic lights, road intersections, bus stops, and road lighting posts. Therefore, there will be several locations in the city where vehicles can offload the data (which they carry from connected-objects) to such PoPs, instead of having only one specific destination. Consequently, this unique configuration would need an efficient data routing approach. This makes the routing mechanism an integral part of the data collection process, which can crucially determine the data collection performance. Therefore, routing needs to be studied extensively to make sure its suitability for the chosen implementation. Before implementing the VDTN data collection system in the smart city environment, as depicted in Figure 1, we need to recognize and analyze the specific nature of the data forwarding and routing process and assess the performance of some baseline VDTN routing protocols. Their performances over defined benchmarks can then be examined to develop a better routing protocol for this specific purpose.

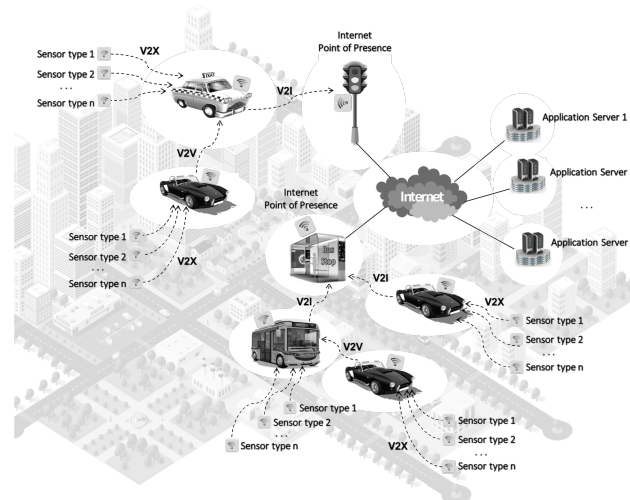


Figure 1. The VDTN-based data collection scheme in smart cities

In this paper, we investigate the networking performance of four baseline VDTN routing protocols, namely: direct delivery, first contact, epidemic, and spray and wait. These routing protocols implement the main routing concepts used in VDTN. We categorized them as ‘baseline’ based on the fact that they use only a few parameters in their data forwarding decision. While some other optimized routing solutions are available in the literature, the study of these baseline protocols will give us the main trends. Interestingly, these four baseline protocols have enough diversity of mechanism to forward data from the source to the destination, from different numbers of copies of data that they forward to varying sets of parameters that need to be considered for making the forwarding decision.

Our study aims to show how such VDTN solutions are efficient for collecting sensors’ data. To investigate each of the strategy’s potential, we devise scenarios where a sufficient number of cars are equipped with communications and networking capabilities. We implement scenarios where 37 static wireless sensors are almost evenly spread out in the city. There are also 5 PoPs where vehicles can offload the collected data from sensors and forward them to the central server. We also explore the impact of different mobility patterns on the networking performance of the data collection system. Two types of vehicles: cars and buses are used in the simulation. Cars represent random vehicle movement, while buses moving along their predetermined route represent a predictable mobility pattern commonly found in smart cities. Ideally, the routing protocol used in this kind of setting needs to consider these different mobility types to exploit them and get better performances. Due to their simplicity, the four baseline routing protocols investigated in this study do not recognize these factors. Yet, it is crucial to understand how they perform under this unique scenario for future refinement. The key performance indicators (KPIs) that we used for comparison are delivery probability (DP), average latency (AL), and overhead ratio (OR), as commonly used in other previous works [12]–[16].

We organized the rest of this paper as follows. Section 2 provides our methods that include baseline routing protocols, KPIs, and the simulation configuration. Section 3 presents the results and discussion of the performance evaluation. Lastly, section 4 delivers conclusions and possible future research.

2. METHOD

Baseline routing protocols are protocols that implement minimum number of parameters for their forwarding decision. Those protocols also require no prior (zero) knowledge of the network condition and nodes’ coordinates. Here, we define four baseline VDTN routing protocols [10], [11], [17]–[20] that are studied in detail later.

The direct delivery routing is a single copy forwarding approach at its simplest form, where a node that has the data forward it directly to its destination. We can use this routing protocol performance as a benchmark that should be exceeded if more complexities are added to the designed routing strategy. Due to its simplicity, we can expect minimal network and buffer usage from this protocol.

The first contact routing is a single copy forwarding approach where each node forwards messages randomly to the first node they encounter. This single copy of messages continues to hop randomly between in-range nodes until one reaches its destination. Nodes erase messages that they already relayed to another node, which means only a single copy of a message exists in the network. This strategy makes the routing protocol very efficient in occupying buffering spaces.

The epidemic routing is a multiple copy forwarding approach where each node keeps copies of every message while also forwarding them to every other node they encounter until the messages reach the destination. Each node receives messages that they do not already have, with their buffering capacity as the only limitation. This approach ensures that at least one copy of each message will reach its destination in the earliest possible time, with the expense of flooding the networks with redundancy. As a benchmark, we can expect this routing as an upper limit in terms of network and buffer utilization.

The spray and wait routing is a more controlled multiple copy forwarding approach, where the number L can be assigned to the protocol to specify the upper limit of copies that can be created per message by a node, as described in [19]. Furthermore, this routing protocol has two modes of operations: standard and binary. In binary mode, if a node is the origin of the data, it will logically hold L copies of the data specified initially (e.g., 6 is the default value). In standard mode, if a connection is established with another node that does not have a copy of the data, only a single copy is forwarded (the *spray* phase), and $L-1$ copies continue to be held by the originating node. The originating node then can forward the remaining copies of the data to each node that it encounters next. The process continues until the originating node only hold the last copy of the data,

when it stop forwarding copies to another node and only forward the data to its destination (the *wait* phase). In binary mode, on the other hand, $L/2$ (rounded up) copies are forwarded each time the originating node meets another node without the data copy, while it will keep $L/2$ (rounded down) copies. As in the standard mode, the forwarding mechanism continues until the originating node has only a single copy of the data when it goes to the wait phase, which is basically similar to the direct delivery routing.

The performances of DTN, and consecutively VDTN, can be measured by several KPIs. In this study, we utilize three commonly used KPIs in previous studies on this topic [12]–[16]. We discuss on each of them in the following.

The DP, or also referred to as the delivery ratio, is the total number of messages successfully delivered to their destination divided by the total number of generated messages at the originating nodes. DP is defined in (1). The ideal condition is for the maximum DP value of 1, where all the generated messages are also successfully delivered to the destinations.

$$DP = \frac{\text{Number of delivered messages}}{\text{Number of generated messages}} \quad (1)$$

The AL, also referred to as the average delivery delay, is the average time it takes from the messages' creation at the sources to the time they are successfully delivered at the destination. AL is defined in (2). The desired condition is for messages to instantaneously reach their destination, i.e., the AL of close to zero, particularly for critical applications. Yet, in the intermittent connectivity of vehicular networks, the ideal condition of very low latency is almost impossible to achieve. Therefore, they are more suitable for delay-tolerant applications, whereas data with an AL in the order of seconds, minutes, or even hours, can still be useful. Nevertheless, the goal is to keep the AL as low as possible.

$$AL = \frac{\sum_{n=1}^N \text{Message arrival time} - \text{Message creation time}}{N \text{ delivered messages}}, \text{ if } N \geq 1 \quad (2)$$

cannot be defined, if } N = 0

The OR shows the ratio of the total number of transmitted messages in the entire network compared with the total number of delivered messages. OR is calculated as in (3). The ideal value for the OR is 0, which happens only when the source of a message or messages directly deliver them to their receiver. If intermediary nodes are involved in the delivery, relaying messages through those intermediaries is considered as excesses or overheads. Hence, relaying copies of messages that do not eventually reach the destination will also be counted as overheads, including copies being rejected by the receiver. In most cases, the receiver only accepts a unique message once, i.e., the first time it arrives. The parameter was also referred to as the network OR in some references because it directly affects the network's resource usage, such as energy consumption for processing and communications, as well as bandwidth allocation. A low Network OR is a vital characteristic of an efficient and scalable data collection system.

$$OR = \frac{T \text{ transmitted messages} - N \text{ delivered messages}}{N \text{ delivered messages}}, \text{ if } N \geq 1 \quad (3)$$

cannot be defined, if } N = 0

The opportunistic networking environment (ONE) simulator [21] was used to evaluate the VDTN-based data collections' networking performances. As we mentioned previously, we apply three KPIs for the performance evaluation, namely DP, AL, and OR [12]–[16]. We utilize a route-based movement for buses and random-waypoints movement for cars as mobility models in a real-world city map in the simulation. In this paper, our main goals are to have the performance comparisons between the baseline routing protocols and understand the trends as the density of the vehicular networks increases, which can be achieved even with these simplest mobility models. As discussed in [22], even though trace-based mobility models give realistic movement patterns, a complete set of traces for the newly investigated scenario, such as in our case, may not be available. Moreover, they produce a static scenario where the number of nodes stays the same, limiting our

possibility of observing the effect of a growing vehicular network. However, we are considering to conduct simulations using suitable location-based-traces for evaluations in our future work.

The city of Helsinki's map, a standard map in the simulator, with a land area size of approximately nine square kilometers, was used to set up the city data collection scenario. We proposed a delay-tolerant environmental monitoring system that acquires data from 37 wireless sensors which spaced almost evenly every 500 meters to each other in the city. Data from those stationary wireless sensors in the environment need to reach one of 5 available internet PoPs strategically placed along the bus routes. The mapping of the simulation is shown in Figure 2. We assume the wireless sensor to be limited in its power capability, i.e., battery-powered or energy harvesting. Therefore, it can only perform a basic data forwarding mechanism, such as the first-contact routing, to transmit data only to the first in-range vehicle.

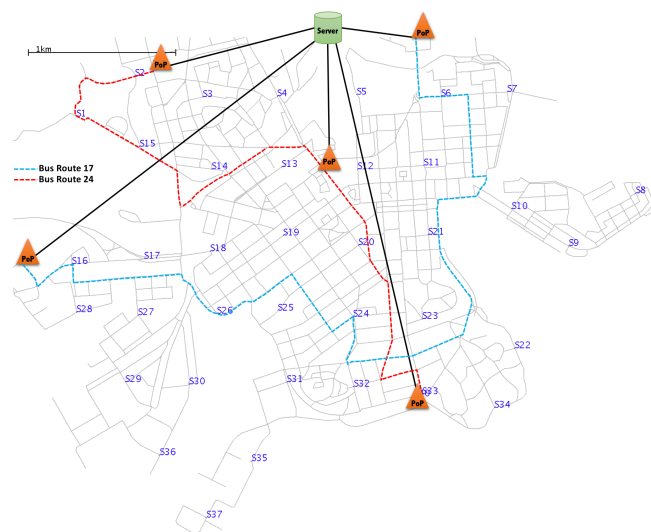


Figure 2. The simulation map and configuration

With several sets of simulation scenarios, we increase the number of cars involved in the data collection, while the number of buses remain the same. The cars involved are from 3 to 90, and 2 buses for each route make up the total of 4 buses assisting in the data collection. One of the PoP is positioned at the city center where most of the traffic converged, while four PoPs are positioned at each end of the two bus routes. In this configuration, a bus will pass at least two PoPs as they move along their route, while there are no guarantee that cars will encounter one. On the other hand, the benefit of utilizing cars is that they can go to various places in the city and opportunistically pick up data from in-range sensors, which is not the case for buses. With the understanding of the trade-off, we aim to observe some baseline routing protocols' networking performance in this setting and find the gap for future refinement.

Each sensor has a ZigBee wireless link profile with a 10 m communications range and data rate of 250 kbps. We choose 10 m as a conservative value for ZigBee's communication range to include the possibility of non-line-of-sight (NLOS) signal propagation condition between sensors and vehicles and even mimic indoor sensor placements. The sensors generate 10 bytes of data every 5 minutes over the first 7 hours of the 12 hours total simulation duration. The data have a time to live (TTL) of 5 hours. The 10 bytes of data is for a single environmental data measurement described in [23]. We choose the ZigBee communication technology as an archetype of a short-range, low-power sensor node operating on battery for months, even years. Each car, which can be connected wirelessly to the sensor via their identical ZigBee link profile, has a pseudo-random initial placement and mobility. For its V2X communications, the car also has an ITS-G5 wireless link profile, with a 300 m communications range and data rate of 6 Mbps. The PoPs have an identical ITS-G5 connection with the cars for data retrieval, and each also has a direct link to the destination server. We implement 5 MB of buffer size for cars and buses, which enable unconstrained evaluation of the networking performance. The number of cars in the scenario varied from 3 to 90, representing an average cars density of 0.33 to 10 per km².

3. RESULTS AND DISCUSSION

In all the scenarios, the first in-range vehicle (car or bus) to the sensor will pick-up the available data. The vehicle then carries and forwards the data based on the routing protocol that they implement. We need to remember that the cars' possibility to reach the PoP is uncertain due to its random mobility setting, while the buses will always encounter at least 2 PoPs along their route. In terms of delivery probability, data carried by buses are desirable, but it may experience a longer waiting time at the sensor, which increases its latency. On the other hand, cars may pick up data from the sensor more often, as statistically more cars will be available in the city. But if the data is only carried by cars, then the delivery probability will suffer, as there is the uncertainty whether it will reach one of the PoP.

Furthermore, we need to emphasize on the dynamic proportion of vehicles involved in the data collection, as it crucially relates to the KPI results' interpretation. In our scenario, buses are considered the mobile network's backbone, as PoPs are placed along their routes. A constant number of buses, 4 in total, are deployed in all scenarios, while the number of cars increases from 3 to 90. Hence, more buses are involved in the network in the scenario with the least number of cars. The proportion of cars in the network then gradually increases while the number of buses remains the same, changing the ratio between vehicles with random and predetermined mobility. The effects of this dynamic, combined with the different mechanisms that each of the baseline routing protocol deploys, are explained in the following discussion.

We first evaluate each routing protocol's delivery probability and show their comparison in Figure 3. It shows that when the cars density was low, i.e., sparse network, the two single-copy approaches, the first contact and the direct delivery, achieved comparable delivery probability to the two multi-copy routing protocols, the spray and wait and the epidemic. But as the cars density increases, i.e., the network becomes denser, their delivery probability begins to lag. Initially, we expected that the epidemic routing protocol would always outperform the other three routings with its maximum copies approach. Yet, surprisingly, the spray and wait's delivery probabilities are comparable to the epidemic's in all network densities. In the larger part, it was because most of the undelivered messages were never transferred from sensors to in-range vehicles until they ran out of TTL and dropped from the sensor's buffer. In a smaller portion, the cause was the unsuccessful messages transfer from vehicles. It turned out that in the intermittent connection with moving vehicles (vehicles with vehicles or vehicles with stationary nodes), the limited contact duration and the number of messages queuing for transmission in the buffer diminished the advantage of having maximum copies of messages in the network. It is possible that when vehicles' buffer becomes flooded by epidemic messages: there is a lower probability that a particular message can be transferred in a brief window of contact between nodes (sensors, cars, buses, and PoPs). Moreover, the faster the vehicle move, the narrower the transferring window becomes, as previously studied in [24]. The dynamic highlighted that flooding the network with unlimited data copies seems to give minor gain, if not none, in terms of delivery probability in our scenario.

In the direct delivery routing, vehicles will keep all data they received from sensors and only do the forwarding if one of the PoP is in the communication range. Therefore, in our scenario, only data picked up by buses will have guaranteed delivery due to the PoPs placement, while data collected by cars will have none of the certainty. It explains why this routing protocol's DP are the lowest, and even more so as the cars density increases.

On the other hand, in the first contact routing, the vehicle with the data will always forward them to the first node (car, bus, or PoP) that it encounters. In this approach, two mechanisms with opposite effects can take place in the data collection. The first one is when the car forwards the data to an in-range bus, i.e., a condition of guaranteed delivery, which is desirable. The second one is when the bus needs to 'give up' its data to a nearby car, i.e., cancellation of guaranteed delivery, which is undesirable. Yet, despite this challenging dynamic, the results show that the first contact performs better than direct delivery in its probability of delivery. It turned out that data picked up by cars from sensors will have a better chance of reaching one of the PoP if it is opportunistically forwarded to the next in-range vehicle rather than to being kept for direct delivery. The advantage becomes more significant if the car is stationary for a long time; offloading the data would increase its delivery chance.

In the spray and wait and epidemic routing, with their multi-copies approach, the vehicle with the data will forward them to all in-range nodes while also keeping the original for direct delivery to the destination. It resulted in a higher delivery probability as more paths are possible for the data to reach the PoPs. The results also give one crucial point: the spray and wait's DP are comparable to those of the epidemic's, hinting to the advantage of forwarding limited copies and the inefficiency of deploying maximum copies in the scenario.

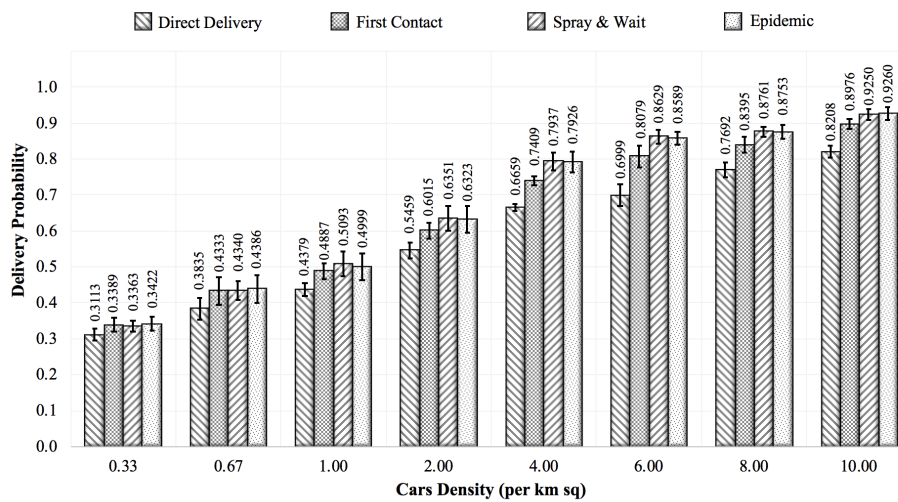


Figure 3. The probability of delivery comparison

Figure 4 compares the routing protocols in terms of the AL. It shows two opposite trends for majority of the routing protocols in relation to the available number of cars in the network. The direct delivery routing protocol is the only exception to the trend. For the direct delivery routing, increases in the cars density are also accompanied by increases in the AL until it reaches around 99 minutes at a cars density of 4 per km². At that point, the AL seems to come to its maximum value, where it remains below 99 minutes, even with the increase in the cars density. In the minimum car density of 0.33 per km², direct delivery's AL is also the lowest at 59.45 minutes. In interpreting this result, we need to relate to their low delivery probability shown in Figure 3, of around 0.31 only, where the AL is calculated only from the successfully delivered data. There are only 3 cars in the entire network in this cars density, less than the number of buses, which is 4. Therefore, there is a high probability that the delivered data were previously picked up by buses, which then carry them to one of the PoPs along their route, causing the lower AL. As the cars density and the delivery probability increase, more cars pick the data instead of buses. Consequently, the AL is also going up due to data being kept longer in cars with the direct delivery routing protocol. The results also suggest a particular maximum value of AL exists for the data collection with this routing protocol, although it will undoubtedly depend on the car's mobility. Yet, with about 0.82 delivery probability at the highest cars density in our simulation, there is merit for its implementation in denser networks for applications that can tolerate a longer latency and an incomplete set of time-series data.

For the first contact routing, although it has a different forwarding mechanism than the direct delivery, the AL trend is quite similar at the lower cars density, where increases in the density also lead to increases in the AL due to the same mechanism. The difference is that the trend reverses at a cars density of 6 per km² when the AL starts to decrease with an increase in the cars density. One important note in the minimum cars density of 0.33 per km²; the first contact AL of 64.52 minutes is higher than the direct delivery. The higher AL might be due to the undesirable effect of buses 'giving up' its data to in-range cars, which we explained earlier. A more advantageous mechanism would be for buses to keep their data until they reach one of the PoPs, a gap that can be refined for a better routing protocol.

The two multi-copies routing protocols, the spray and wait and epidemic, also possess a similar trend to the first contact routing. They differ in that they have lower AL and the reverse in the trend for them starts earlier at a cars density of 2 per km². Figure 4 also shows that the spray and wait only lagging the epidemic slightly in terms of its AL, particularly at the lower cars density, even though it only deploys up to 6 copies of each data in the forwarding process. It is a crucial advantage that will be highlighted further when discussing its OR comparison with the epidemic later on.

If we observe Figure 4 further and compare the AL of the first contact and spray and wait, we can find that the first contact's AL is always higher than spray and wait's in all of the cars densities. It is primarily due to the higher number of data copies that spray and wait allowed to forward, which increases the possibility that one of the copy will reach a PoP sooner. Interestingly, the dynamic is also influenced by the inclusion of

buses in the data collection network and how the spray and wait routing 'accidentally' harnesses their potential. At a specific condition where a bus received a copy or copies of the data, the two routing protocols behave differently. When a bus with first contact routing received a single copy of data, there is no guarantee that it will keep the data until it reaches one of the PoPs. As we pointed out earlier, there is the possibility that the bus will have to forward the data to an in-range car before it encounters any PoPs. Contrary, when a bus with spray and wait routing received a copy or copies of data, it will at least keep a copy for direct delivery, consequently elevating the delivery probability and potentially minimizing the AL. The 'accidental' behavior is a strength that needs to be embedded into a better routing strategy for the specific data collection purpose.

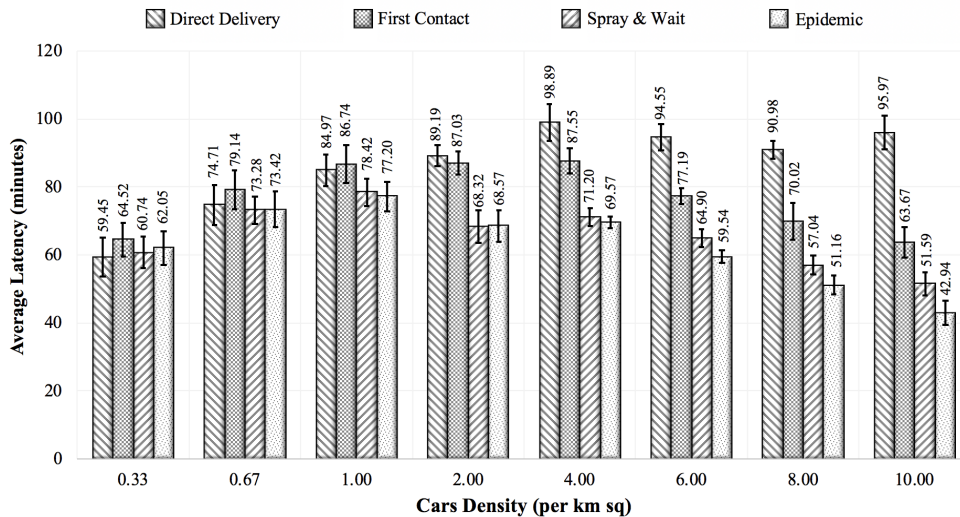


Figure 4. The average latency comparison

Figure 5 on the other hand, captures each routing protocol's networking cost in terms of their overhead ratio. It shows contrasts between each routing protocol in the single-copy and the multi-copies groups. In the single-copy group, the direct delivery routing produces the lowest OR in all cars densities, a constant value of 2, as the single-copy data forwarding only happens from sensors to cars or buses, and then from cars or buses to one of the PoPs. On the contrary, the first contact routing showed an increase of OR as the network becomes denser. The strategy of always forwarding the data to the next vehicle in-range makes the OR goes up as more vehicles become available to receive the data. There is no restriction on how many data forwarding processes can take place, until eventually the data reach one of the PoPs. There is also no distinction on which vehicles to forward the data to, which make the data transfer from buses to cars also possible. As described earlier, it would be more efficient if there is a distinction on which kind of vehicles to forward to, e.g., buses which have higher delivery probability should not forward data to cars. The added logic will lead to lower AL and OR, as well as higher delivery probability.

We can also observe both similarity and disparity of the OR trend in the multi-copies routing pair. Spray and wait and epidemic both experience an increase in the OR as the number of cars grows in the network. The difference is that the spray and wait's OR seems to reach its saturation value of around 8 at a cars density of 2 per km² and only has minor increases as the cars density goes up. The dynamic occurs because the spray and wait routing implements a limited number of copies and, consequently, a limited number of forwarding strategy. Vehicles only forward a limited number of data copies to other vehicles. Each vehicle then does a direct delivery to one of the PoPs when only a single copy of the data remains in their possession. Therefore, the so-called 'wait' mode limits the communication hops and consequently reduces the OR. The opposite happens to the epidemic, where its OR goes up sharply as the cars density increases. The epidemic routing deploys unlimited data copies and communication hops strategy at the expense of inundating the network with more overheads, even though it manages to achieve the lowest AL in a denser network. Indeed, the trade-off will be a concern to its scalability in an ever-growing network and further emphasizes the advantage of spray and wait's limited-copies strategy.

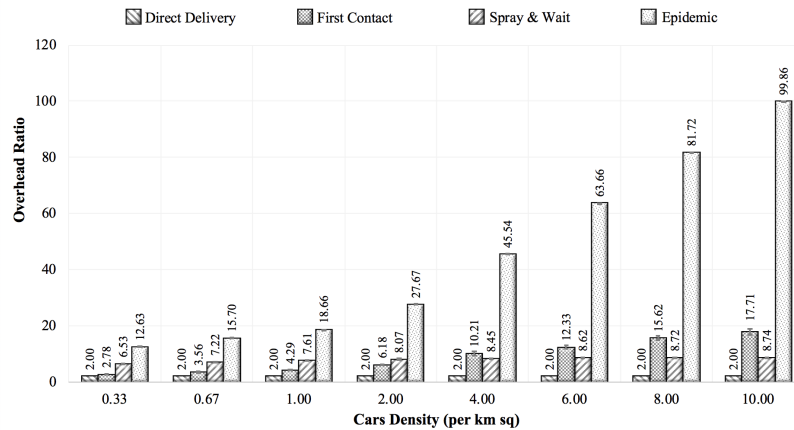


Figure 5. The overhead ratio comparison

To summarize, Table 1 shows the strengths and weaknesses of the four baseline routing protocols when applied to the proposed data collection scheme. All the results point out that the baseline routing protocols evaluated in this study do not have a mechanism to exploit the full potential of having vehicles with predetermined mobility, such as buses, in the data collection process. Therefore, the advantage of having ITS infrastructures to assist in the system is also overlooked. Our performance evaluation indicates the need to develop a routing protocol that recognizes and utilizes the advantage of different kinds of mobility in the field for data collection, such as the work on a hierarchical VDTN routing protocol named data collection for low energy devices (DC4LED) presented in [25], [26]. We believe that for the specific purpose of data collection that we already discussed, a simple and efficient routing protocol can be deployed for the delay-tolerant data delivery, contrary to the implementation of optimized solutions that might be resource-demanding and difficult to standardize.

Table 1. Summary of the baseline routing protocols' strengths and weaknesses

Routing protocol	Strengths	Weaknesses
Direct delivery	<ul style="list-style-type: none"> Fairly high delivery probability and showing an upper-limit of AL in a denser network; might be applicable for certain applications. Constant OR and the lowest among the baseline routings, offering the most efficient network usage. 	<ul style="list-style-type: none"> Lowest delivery probability and highest AL among the baseline routings.
First contact	<ul style="list-style-type: none"> Fairly high delivery probability and shows a decreasing AL as the network become denser. 	<ul style="list-style-type: none"> Continuous increases in OR as the network become denser. It does not have a mechanism to recognize vehicles with predetermined mobility; therefore, the single-copy of data might be offloaded to cars that have a lower probability of delivering them.
Spray and wait	<ul style="list-style-type: none"> Tied-highest delivery probability (with epidemic) among the baseline routings and shows a decreasing AL as the network become denser. Moderate OR with minor increases as the network grows. The <i>wait</i> phase, when a forwarding vehicle keeps at least a copy of the data, 'accidentally' ensures the direct delivery of the data by vehicles with predetermined movement. 	<ul style="list-style-type: none"> It does not have a mechanism to recognize vehicles with predetermined mobility; therefore, data that is already relayed to vehicles with predetermined mobility continue to be forwarded to cars with a lower probability of delivering them, consequently increasing the OR.
Epidemic	<ul style="list-style-type: none"> Tied-highest delivery probability (with spray and wait) and the lowest AL among the baseline routings. 	<ul style="list-style-type: none"> Extremely high OR, which will continue to increase in a growing network.

4. CONCLUSION AND FUTURE WORKS

This paper presented the performance comparison of four baseline VDTN routing protocols in a unique ITS-assisted data collection setting in smart cities. In the VDTN-based data collection scenario, several internet PoPs are available and strategically located in the city to assist in the data delivery from connected-objects to the application server. We assessed the four routing protocols in this specific scenario to see gaps that can be refined. These protocols have sufficient diversity of mechanisms to forward data, giving plenty of insight into their impact on the KPIs. Furthermore, the inclusion of two types of vehicles, each with their distinct mobility pattern, into the simulation also reveals some crucial forwarding dynamics. The comparison

between the baseline routing protocols showed that generally multi-copy strategies perform better by offering higher delivery probability and lower AL. Their drawback was on the high OR that will burden the network in a dense environment such as smart cities. Moreover, we also observed different trends in varying vehicles density, wherein a sparse network the single-copy approaches actually have a comparable performance with the multi-copy strategy. It was only when the network became denser that the advantage of forwarding limited multi-copies data became apparent, suggesting implementing a different routing strategy based on the vehicular network density. Our work also emphasizes some gaps in the routing strategies that can be refined for better performances. Our future works will include refining the DC4LED routing protocol to incorporate the advantage of the limited-copies approach highlighted by the spray and wait routing performances in this paper. The DC4LED routing protocol already consists of a mechanism that recognizes and utilizes specific mobility patterns from various vehicles available in the city. It forwards data hierarchically to maintain efficiency and scalability. The protocol will also need to be tested in a more realistic vehicle's mobility pattern provided by real-world global positioning systems (GPS) traces to increase the KPI values' accuracy.




REFERENCES

- [1] Y. Chen, Y. A. Sambo, O. Onireti, and M. A. Imran, "A Survey on LPWAN-5G Integration: Main Challenges and Potential Solutions," *IEEE Access*, vol. 10, pp. 32132–32149, 2022, doi: 10.1109/ACCESS.2022.3160193.
- [2] E. U. Ogbodo, A. M. Abu-Mahfouz, and A. M. Kurien, "A Survey on 5G and LPWAN-IoT for Improved Smart Cities and Remote Area Applications: From the Aspect of Architecture and Security," *Sensors*, vol. 22, no. 16, p. 6313, Aug. 2022, doi: 10.3390/s22166313.
- [3] R. Chataut, A. Phoummalayvane, and R. Akl, "Unleashing the Power of IoT: A Comprehensive Review of IoT Applications and Future Prospects in Healthcare, Agriculture, Smart Homes, Smart Cities, and Industry 4.0," *Sensors*, vol. 23, no. 16, p. 7194, Aug. 2023, doi: 10.3390/s23167194.
- [4] M. A. Jamshed, K. Ali, Q. H. Abbasi, M. A. Imran, and M. Ur-Rehman, "Challenges, Applications, and Future of Wireless Sensors in Internet of Things: A Review," *IEEE Sensors Journal*, vol. 22, no. 6, pp. 5482–5494, Mar. 2022, doi: 10.1109/JSEN.2022.3148128.
- [5] F. Arena and G. Pau, "An Overview of Vehicular Communications," *Future Internet*, vol. 11, no. 2, p. 27, Jan. 2019, doi: 10.3390/fi11020027.
- [6] G. Pau, A. Severino, and A. Canale, "Special Issue 'New Perspectives in Intelligent Transportation Systems and Mobile Communications towards a Smart Cities Context,'" *Future Internet*, vol. 11, no. 11, p. 228, Oct. 2019, doi: 10.3390/fi11110228.
- [7] Z. MacHardy, A. Khan, K. Obana, and S. Iwashina, "V2X Access Technologies: Regulation, Research, and Remaining Challenges," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 1858–1877, 2018, doi: 10.1109/COMST.2018.2808444.
- [8] L. Khoukhi, H. Xiong, S. Kumari, and N. Puech, "The Internet of vehicles and smart cities," *Annals of Telecommunications*, vol. 76, no. 9–10, pp. 545–546, Oct. 2021, doi: 10.1007/s12243-021-00891-7.
- [9] F. Arena, G. Pau, and A. Severino, "A Review on IEEE 802.11p for Intelligent Transportation Systems," *Journal of Sensor and Actuator Networks*, vol. 9, no. 2, p. 22, Apr. 2020, doi: 10.3390/jsan9020022.
- [10] P. R. Pereira, A. Casaca, J. J. P. C. Rodrigues, V. N. G. J. Soares, J. Triay, and C. Cervello-Pastor, "From Delay-Tolerant Networks to Vehicular Delay-Tolerant Networks," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 1166–1182, 2012, doi: 10.1109/SURV.2011.081611.00102.
- [11] N. Benamar, K. D. Singh, M. Benamar, D. El Ouadghiri, and J.-M. Bonnin, "Routing protocols in Vehicular Delay Tolerant Networks: A comprehensive survey," *Computer Communications*, vol. 48, pp. 141–158, Jul. 2014, doi: 10.1016/j.comcom.2014.03.024.
- [12] Agussalim and M. Tsuru, "Comparison of DTN Routing Protocols in Realistic Scenario," in *2014 International Conference on Intelligent Networking and Collaborative Systems*, IEEE, Sep. 2014, pp. 400–405, doi: 10.1109/INCoS.2014.80.
- [13] S. A. Hadiwardoyo and A. J. T. Santos, "Deploying public surface transit to forward messages in DTN," in *2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*, IEEE, Aug. 2015, pp. 1329–1335, doi: 10.1109/IWCMC.2015.7289275.
- [14] C. Giannini, P. Calegari, C. Buratti, and R. Verdone, "Delay Tolerant Network for smart city: Exploiting bus mobility," in *2016 AEIT International Annual Conference (AEIT)*, IEEE, 2016, pp. 1–6, doi: 10.23919/AEIT.2016.7892779.
- [15] E. Spaho, L. Barolli, V. Kolicic, and A. Lala, "Evaluation of Single-Copy and Multiple-Copy Routing Protocols in a Realistic VDTN Scenario," in *2016 10th International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS)*, IEEE, Jul. 2016, pp. 284–289, doi: 10.1109/CISIS.2016.43.
- [16] R. Hernández-Jiménez, C. Cardenas, and D. Muñoz Rodríguez, "Modeling and Solution of the Routing Problem in Vehicular Delay-Tolerant Networks: A Dual, Deep Learning Perspective," *Applied Sciences*, vol. 9, no. 23, p. 5254, Dec. 2019, doi: 10.3390/app9235254.
- [17] S. Jain, K. Fall, and R. Patra, "Routing in a delay tolerant network," in *Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications*, New York, NY, USA: ACM, Aug. 2004, pp. 145–158, doi: 10.1145/1015467.1015484.
- [18] A. Vahdat and D. Becker, "Epidemic routing for partially-connected ad hoc networks," CS-2000-06, 2000.
- [19] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and wait: An efficient routing scheme for intermittently connected mobile networks," in *Proceedings of the ACM SIGCOMM 2005 Workshop on Delay-Tolerant Networking, WDTN 2005*, New York, New York, USA: ACM Press, 2005, pp. 252–259, doi: 10.1145/1080139.1080143.
- [20] T. Abdelkader, K. Naik, A. Nayak, N. Goel, and V. Srivastava, "A performance comparison of delay-tolerant network routing protocols," *IEEE Network*, vol. 30, no. 2, pp. 46–53, Mar. 2016, doi: 10.1109/MNET.2016.7437024.
- [21] A. Keränen, J. Ott, and T. Kärkkäinen, "The ONE simulator for DTN protocol evaluation," in *Proceedings of the Second Interna-*




- tional ICST Conference on Simulation Tools and Techniques*, ICST, 2009, doi: 10.4108/ICST.SIMUTOOLS2009.5674.
- [22] J. Dede et al., "Simulating Opportunistic Networks: Survey and Future Directions," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 1547–1573, 2018, doi: 10.1109/COMST.2017.2782182.
- [23] L. Kang, S. Poslad, W. Wang, X. Li, Y. Zhang, and C. Wang, "A Public Transport Bus as a Flexible Mobile Smart Environment Sensing Platform for IoT," in *2016 12th International Conference on Intelligent Environments (IE)*, IEEE, Sep. 2016, pp. 1–8, doi: 10.1109/IE.2016.10.
- [24] N. I. Er, K. D. Singh, and J. M. Bonnin, "Contact Strategy for VDTN Data Collection in Smart Cities," in *2018 International Conference on Smart Green Technology in Electrical and Information Systems: Smart Green Technology for Sustainable Living, ICSGTEIS 2018 - Proceeding*, IEEE, Oct. 2018, pp. 61–64, doi: 10.1109/ICSGTEIS.2018.8709131.
- [25] N. I. Er, K. D. Singh, and J.-M. Bonnin, "DC4LED: A Hierarchical VDTN Routing for Data Collection in Smart Cities," in *2019 16th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, IEEE, Jan. 2019, pp. 1–4, doi: 10.1109/CCNC.2019.8651757.
- [26] N. I. Er, K. D. Singh, and J. M. Bonnin, "CityView: A Vehicle Based Image and Data Collection for Smart Cities," in *International Conference on Wireless and Mobile Computing, Networking and Communications*, IEEE, Oct. 2019, pp. 1–7, doi: 10.1109/WIMOB.2019.8923547.

BIOGRAPHIES OF AUTHORS






Ngurah Indra Er    is a lecturer in the Department of Electrical Engineering, Faculty of Engineering, Udayana University (UNUD), Bali, Indonesia, since 2002. He received his Ph.D. from IMT Atlantique, Rennes, France. He completed his M.Sc. degrees in Communications Engineering from University of Birmingham, U.K., and his B.Eng. in Electrical Engineering from Sepuluh Nopember Institute of Technology (ITS) Surabaya, Indonesia. His current research interests include internet of things, internet of vehicles, vehicular networks, opportunistic networks, and smart city data collection. He can be contacted at email: indra@unud.ac.id.






Kamal Deep Singh    received the B.Tech. degree in electrical engineering from the IIT Delhi, New Delhi, India, in 2002, and the Ph.D. degree in computer science from the University of Rennes 1, Rennes, France, in 2007. He then joined the Dionysos Group, National Research Institute in Computer Science (INRIA), as a Postdoctoral Researcher, where he co-developed many components of quality-of-experience estimation tools and worked mainly on the analysis of video-based applications. He was a Postdoctoral Researcher with the Telecom Bretagne, Rennes, where he worked on the internet of things and cognitive radio. He is currently an Associate Professor with the Université de Saint-Étienne, Saint-Étienne, France, where he is a part of the research team Data Intelligence, Laboratoire Hubert Curien. His research interests include the internet of things, mobile networks, AI, quality of experience, and software-defined networking. He can be contacted at email: kamal.singh@univ-st-etienne.fr.



Christophe Couturier    joined IMT Atlantique, France in 2014 where he holds a position of Assistant Professor. After an engineering degree in electronics at INSA of Rennes in 2001, he worked for 13 years as a technical manager for innovative telecommunication projects in the field of transportation (rail and road). Currently, his main areas of interest are intelligent transportation systems (ITS) and mobile networks. He can be contacted at email: christophe.couturier@irisa.fr.



Jean-Marie Bonnin    is currently a Professor (HDR) in Networks and Computer Science at IMT Atlantique, France and the leader of the Inria/EASE research team. After a Ph.D. degree in computer science at the university of Strasbourg in 1998, he came to ENST Bretagne for a research position. Since 2001, He has been mainly interested in the convergence between IP networks and mobile telephony networks, and therefore in heterogeneous handover management. More recently, he has been involved in projects dealing with the mobility of the networks and its application to intelligent transportation system (ITS). His current research interest is on how to provide pervasive applications with a good perception of their environment through localized interactions, specifically in cooperative autonomous vehicles. He is also a co-founder and scientific advisor of the YoGoKo startup which provides communication solutions for cooperative ITS including autonomous vehicles. He can be contacted at email: jean-marie.bonnin@irisa.fr.