Unified Routing for Data Dissemination in Smart City Networks

Viet-Duc Le, Hans Scholten and Paul Havinga
Pervasive Systems (PS), Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS)
University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
{v.d.le, hans.scholten, p.j.m.havinga}@utwente.nl

***Abstract*—The Internet of Things continues to expand from traditional homogeneous technologies with low resources to increasingly heterogeneous and resource rich technologies. An emerging domain in this respect is the use of mobile phones to monitor and process events. Moreover, there is increasing interest in expanding the application domains, such as to smart cities. Existing routing algorithms for such technologies and application domains are still designed for homogeneous technologies, such as mobile phones, vehicles, and Road Side Units. In this paper, we propose a new routing scheme for a heterogeneous architecture that is composed of a heterogeneous set of nodes. Our proposed routing algorithm uses two parameters, namely the Delivery Capability and Number of Copies, to control the dissemination area and connection priority. By optimally choosing these two parameters according to characteristics of nodes, the proposed algorithm works well with various types of sensor node. In addition, a heterogeneous network architecture and several mobility models are introduced to obtain a realistic simulation scenario. The simulation results show that our proposed approach outperforms all other compared algorithms in terms of delivery ratio and latency.**

I. Introduction

Most existing sensor network research has focused on collecting and processing environmental data using a static topology, homogeneous devices, and an application-aware infrastructure, whereas opportunistic sensing involves collecting, storing, processing and fusing large volumes of data using any technology that is both available and suitable. Significant innovations in mobile phones nowadays inspire the use of mobile phones as sensor nodes in opportunistic sensing to both measure and transfer data. This creates a very detailed knowledge of what is happening in the world around us, and thus to make business processes more efficient and effective. In this paper, we focus on the use of opportunistic sensing to evaluate various services for smart cities, such as traffic control, safety, and environmental issues. Due to the mobility and sparse density of mobile phone nodes, end-to-end connectivity would be unstable and does not even exist. Therefore, the conventional routing protocols are not suitable for opportunistic city networks [1]. Recent literature has proposed novel opportunistic routing schemes [2]–[8], but none of these are suitable for a heterogeneous architecture with non-deterministic movement. Heterogeneity means that there are several kinds of sensor nodes, such as pedestrians, cars, busses, or Road Side Units (RSUs). Opportunistic network applications frequently need to have a heterogeneous architecture. The accuracy of data collected by mobile phones held by pedestrians or within cars might be further increased by powerful sensor nodes, such as RSUs. In addition, by forming a connected ad-hoc network, RSUs can be used to accelerate the speed of message delivery. This is particularly useful for public safety applications, such as fire detection, hazard detection, and law enforcement. Messages with high priority are transferred to a common sink node via the RSUs' network. Furthermore, the heterogeneous architecture increases the delivery ratio of routing algorithms.

In this paper, we propose a novel routing scheme to suit such compulsive architecture in an opportunistic mobile phone network. We term our proposal as *Unified* since it partly modifies and merges well-known existing opportunistic routing algorithms such as Direct Delivery [2], Epidemic [3], Spray and Wait [4], and oracle-based algorithms [6] into one unified routing scheme. Our protocol uses two main parameters, namely the Delivery Capability and Number of Copies, to optimize its performance in terms of delivery ratio, latency, and transmission cost. Depending on the physical characteristics of each type of node as analyzed by Le et al. [1], these parameters are optimally set to match each performance requirement. For example, the Number of Copies is set to be high for RSUs and low for pedestrians, cars, and busses. RSUs are set with the highest Delivery Capability owing to their high throughput and large storage.

Observations from the results of simulations suggest that our proposed routing scheme gives better performance in both terms of message delivery ratio, and latency than compared schemes do. Clearly, there is a penalty in increasing transmission cost. However, a high transmission cost, proportional to the number of message replicates, indicates that many messages are disseminated in the networks. Therefore, the high transmission cost is not the most demanding and relevant metric for data dissemination applications, such as public safety, intelligent transportation, and social networks.

The rest of this paper has the following structure. Related work is discussed in Section 2. Section 3 presents our Unified routing and optimization with a heterogeneous architecture. Simulations including settings for heterogeneous architecture and mobility models are described in Section 4, while Section 5 concludes this paper.

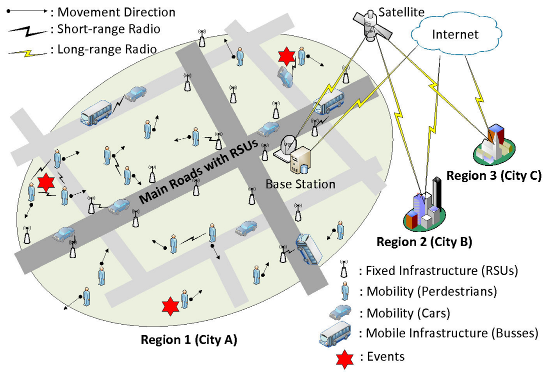


Fig. 1. Heterogeneous Architecture for Smart City Networks.

II. Related Work

In this paper, we focus on data gathering and disseminating in a realistic sensor network as illustrated in Figure 1, which is fully described in work of [1]. The smart city network, which comprises the existing wireless sensor network and the mobile sensor network, has various components, for instance, Road Side Units, mobile phones carried by pedestrians, cars, and busses. Therefore, the network can be characterized as a mixture of connected ad-hoc networks and Delay-Tolerant Networks (DTN). Conventional wireless ad-hod network [9], [10], and even today DTN routing algorithms [2]–[4], [6], [11] cannot perform well in such a heterogeneous architecture. Wireless ad-hod network routing protocols certainly fail to discover a path for forwarding messages among mobile nodes when the network is sparse. Even if a such path is exposed, it will quickly change or disappear because of unpredictable mobility movement. Meanwhile, existing DTN protocols perform poorly on ad-hoc connected networks because a node, which is almost stationary, has no opportunity to contact other distant nodes. Moreover, it is also impossible to apply separate algorithms for different components because these algorithms cannot exchange messages between one another. Recently, several literatures [7], [8] show better performances with assistance of infrastructure, but they require movement information to be pre-known. Therefore, a new routing protocol is necessary for the emerging type of network that has a heterogeneous architecture with unpredictable movement of mobile nodes. In fact, our proposal is inspired by the following novel algorithms that cover most kinds of DTN routing: single-copy, multi-copy, stochastic, and context-based.

The Direct Delivery (DD) algorithm [2] allows a node to transfer a message only to the destination. Once completing the transmission, the node will delete the sent messages from its buffer. Thus, the DD algorithm is a kind of single-copy algorithm with very poor performance in term of delivery ratio. By contrast, Epidemic [3] diffuses messages similarly to the way in which virus and bacteria propagate. Once being in communication range, a node reduplicates messages that other nodes do not have and then transfers them. These messages will be continuously broadcast to other nodes during the following contact times. By flooding the network with multiple copies of messages, Epidemic can obtain a high delivery ratio, but rapidly exhausts available resources. To comprise the tradeoff between FC and Epidemic, Spyropoulos et. al proposed the Spray and Wait (SnW) algorithm [4], which estimates the optimal number of copies for each message. If that number reduces to one, the message can no longer be transferred or replicated. Note that FC, SnW, and Epidemic belongs to stochastic routing that delivers messages by simply disseminating them all over the network by using contact opportunities. Under such circumstances, most of the messages will be gradually delivered at the destination. On the other hand, the Probabilistic Routing Protocol using History of Encounter and Transitivity (ProPHET) [6] estimates delivery predictability for each known destination at each node before transferring a message. As using network oracles, ProPHET is categorized into the context-based scheme.

The ONE simulator [12] is used for our simulations since it possesses several implemented DTN algorithms as well as mobility models. Maps can be imported into the ONE simulator to make simulations more realistic. It is feasible to configure a simulation with specific settings by using parameters aplenty, such as, mobility velocity, message sizes, buffer sizes, and etc.

III. Unified Routing Protocol

As discussed in Section 2, current routing algorithms cannot perform well in the aforementioned scenarios. Therefore, it is necessary to have a unified routing algorithm that adapts most of the components of heterogeneous architecture such as lamp-posts, pedestrians, cars, and busses. RSUs are intentionally designed in order to improve not only the reliability of event detection, but also to play an important role in both gathering and disseminating data. Since the RSU network is a connected ad-hoc network with a few constraints on throughput, energy, and storage capacity, the approach should preferably transfer messages along the RSU lines rather than among mobile nodes to a common sink, such as the base station shown in Figure 1. The proposed scheme also must allow messages to be temporarily stored at RSUs so that other mobile nodes can directly retrieve the necessary information.

Moreover, the algorithm should avoid flooding the mobility network since mobile nodes have limited power and bandwidth, while still obtaining a high delivery ratio. It is necessary to disseminate a detected event to some specific areas or groups. Normally, the people surrounding an event are interested in what is happening nearby because it is likely that they are at risk. Therefore, the approach should not entirely flood the whole city with such information, but only neighboring areas. Finally, the algorithm must be light-weight and need as little network information as possible because mobile phones have limited energy and computation power, and many people are unwilling to share their personal information.

Following on from the above, we define the desirable objectives of our proposed routing algorithm for data dissemination in smart city networks:

- must work well with a heterogeneous architecture,
- maximize the use of the RSUs' network to improve performance in terms of delivery ratio and latency,
- control flooding phenomena in the mobility group,
- perform better than existing algorithms,
- possibly manipulate disseminated-message areas,
- need as little network oracle as possible to facilitate implementation in real-world applications.

To achieve these goals, we propose the Unified routing algorithm, combining the advantages of stochastic and oracle-based algorithms. By adding mechanisms to vary two parameters, namely the Delivery Capability and the Number of Copies, the algorithm can adapt itself to respond to various node and connection types.

A. Unified Parameters

Unified uses the Delivery Capability and the Number of Copies to decide which node and which message to start transferring, respectively. The Delivery Capability value is a constant, whereas the Number of Copies gradually decreases from an initial value.

1) Delivery Capability: The Delivery Capability, denoted by C, represents the possibility that a node can deliver information based on its hardware characteristics. For example, a RSU usually has unlimited power supply, high bandwidth, and thus possesses a high Delivery Capability value. Otherwise, a pedestrian carrying a mobile phone, which has battery constraints and low bandwidth, will be assigned a low Delivery Capability. Since hardware is almost the same for all nodes in the same component and remains constant over a long period, nodes in the same component have the same value of the Delivery Capability, and nodes from different components have different values of the Delivery Capability.

When there are several nodes in communication range, a node will be selected as a master node and the remaining nodes as slave nodes by applying some existing clustering mechanism, such as K-means clustering [13]. The master uses the value of the Delivery Capability to determine which slave node to transfer messages first. The slave node, of which the Delivery Capability is higher than that of others, will has the higher priority to exchange messages with the master. This connection selecting is essential to assure messages are handed to the most suitable nodes since the cluster will not last for long due to the fact that nodes move continuously.

2) Number of Copies: The Number of Copies, denoted by L, defines the constraint of message reduplication. Depending on the type of network and the number of sensor nodes, all messages originating from the same component have the same initial value of the Number of Copies. The values will be gradually decreased down to 1 by dividing by 2 after each successful transfer.

Unlike the Binary Spray and Wait mechanism proposed in [4], which does not allow a node to send out a message that has the Number of Copies equal to 1 and is consistent with the maximum instants of a message in the network, Unified still allows a node to transfer such a message to other nodes if these nodes belong to different components. Under such circumstances, a receiver will check the original component of the incoming message, from which component the message was created. If the message was generated by a node in the same component, the receiving node will assign 1 to the Number of Copies of the message in order to deter further broadcasting. If not, the initial Number of Copies of the receiving node will be attached to the message to encourage more dissemination. This idea is based on a hypothesis: Information is well known within a component but unpopular in other components.

In fact, the Number of Copies is used to optimize the delivery ratio of messages. If the Number of Copies is too small, there will be few opportunities to deliver messages to their destinations. However, if the Number of Copies is too large, many copies will flood the network and decrease the delivery performance.

B. Unified Routing

Each component k, such as RSUs or pedestrians, has its own predefined the Number of Copies, denoted by L_k, and the Delivery Capability, denoted by C_k. The current Number of Copies is gradually updated as described in Section 3.A. Clearly, the values of L_k and C_k rely heavily on the physical characteristics of each component. Moreover, these parameters can be flexibly set to match specific architectures and applications. To this end, we are led to the flowing pseudo code of the Unified routing algorithm for an arbitrary master node i. A master is the node is holding the communication channels as a cluster head, and trying to transmit its messages to the in-range neighboring nodes.

When node i encounters one or more nodes, connection is up, the master collects and sorts all active connections in descending order of the Delivery Capability C_k by $sortCapability(Connections)$. Then, for each connection, two nodes exchange their summary vectors to determine which messages the other node does not own, similarly to Epidemic [3]. These requested messages are sorted by a message sorting method, such as, random or First In First Out (FIFO), before being pushed into the outgoing buffer $outMesssages$. From now, node i can start sending out selected messages through connection by connection. Note that the current Number of Copies, $sendMessages.l$, is reduced if and only if the receiving node is in the same group. A slave node receiving messages will check again whether receiving messages exist in its buffer. The current Number of Copies l of $recevingMessage$ will be updated as the rule defined in Section 3.A. As long as a received message comes from a node in the same component, the slave node updates the current Number of Copies $recevingMessage.l$ by dividing by 2. Once the sending node belongs another component, the receiving node will check the origin of the message. If the message originates from a node in the same component, $recevingMessage.l$ will be set to 1. If not, $recevingMessage.l$ is set as L_k of the slave node.

Algorithm 1 <Unified Routing Algorithm - Master>

1: INITIALIZE: L_k, C_k
2: **if** Connection is up **then**
3: $Connections \leftarrow getConnections()$
4: $Connections \leftarrow sortCapability(Connections)$
5: **for** $connection : Connections$ **do**
6: $hostMessages \leftarrow getHostMessages()$
7: $otherMessages \leftarrow getOtherMessages()$
8: **for** $hostmessage : hostMessages$ **do**
9: $tempMessages \leftarrow \emptyset$
10: **if** $hostmessage \notin otherMessages$ **then**
11: $tempMessages \leftarrow tempMessages + \langle hostmessage, connection \rangle$
12: **end if**
13: **end for**
14: $tempMessages \leftarrow sortMessages(tempMessages)$
15: $outMessages \leftarrow outMessages + tempMessages$
16: **end for**
17: **for** $sendMessage : outMessages$ **do**
18: **if** $otherNode$ in same group **then**
19: **if** $sendMessage.l > 1$ **then**
20: send $sendMessage$ out
21: update $sendMessage.l$ if sending successfully
22: **end if**
23: **else**
24: send $sendMessage$ out
25: **end if**
26: **end for**
27: **end if**<end>

Algorithm 2 describes the pseudo code of messages received at a slave node.

In fact, the Unified algorithm also covers most novel existing schemes, both stochastic and oracle-based routings. When L is set as infinite for all components, the algorithm switches itself to an Epidemic-like scheme by entirely flooding the network with message copies. When L is finite, a node will spread messages like Spray and Wait ($+\infty > L > 1$) or Direct Delivery ($L = 1$) to nodes inside the component, and like Epidemic to nodes outside the component. Furthermore, the predictability of delivery, which can be estimated by an oracle-based routing algorithm, can be used to sort the requested messages of each connection at line 14 of Algorithm 1. Since messages probably have various destinations, the predictability of a node to deliver messages are not unique. Therefore, those messages that possess higher delivery probability should be transferred before others. In addition, since connections are sorted in advance at line 4, the Unified algorithm so far elaborates oracle-based algorithms by prioritizing connections. Due to the limited space, we will discuss the sorting schemes more detail in another work.

C. Optimizing C and L

Without loss of generality, we analyze our proposal with one single city only for convenience of understanding. It is possible to extend the approach for the multiple-city scenario by adding

Algorithm 2 <Unified Routing Algorithm - Slave>

INITIALIZE: L_k, C_k
if Connection is up **then**
$hostMessages \leftarrow getHostMessages()$
$receivingMessages \leftarrow getRecevingMessages()$
for $recevingMessage : recevingMessages$ **do**
if $recevingMessage \notin hostMessages$ **then**
$inMessages \leftarrow inMessages + recevingMessage$
if $otherNode$ in the same group **then**
update $recevingMessage.l$
else
if $recevingMessage$ originated from the same group **then**
$recevingMessage.l \leftarrow 1$
else
$recevingMessage.l \leftarrow L_k$
end if
end if
end if
end for
end if<end>

the Message Ferries algorithm for the bus component. Thus, busses are categorized into the mobility component, and we consider optimal values of C and L for only two components in our proposed algorithm: Road Side Units and Mobility.

1) Road Side Units: RSU networks are a kind of wireless ad-hod network (WAN) so that one might think conventional routing algorithms would be the best options. In fact, the aim of conventional routing is mainly to find the shortest or least-cost path for message delivery. However, the prime objective of data dissemination is not only delivering messages to a specific sink, but also broadcasting them as widely as possible to warn other civilians about imminent dangers. Under such circumstances, conventional routing algorithms fail to spread information over broader areas. Conversely, the proposed Unified with setting high C and large L can work well.

On the one hand, the Unified approach transfers messages along RSUs to a sink quickly since it is not necessary to find a shortest path. Since RSUs are implemented in linear formation, a RSU can only communicate with at most two neighbors. Once a message first reaches a RSU, no more than two copies of a message are created. Afterwards, the algorithm replicates the message copy and send it to the next RSU. In this way, the message is rapidly delivered to the sink. The progress of the Unified routing does not take much longer than that of conventional algorithms because there are only a few branches or spanning trees that lead to the sink.

On the other hand, this approach causes the RSU network to be entirely flooded with messages. Once a RSU receives a message from outside of its network, for instance, from a pedestrian, Unified speedily diffuses the message to most of the RSUs. Other pedestrians, cars, or even busses can retrieve the information directly from RSUs they encounter. Therefore, it is unnecessary to send a request message to the base station to demand information. In other words, RSUs can be either sink or source nodes for the mobile sensor network.

To this end, the optimal value of C for RSUs should be set to the highest value ($C = 1$) so that messages can take the advantage of the RSU network. In addition, the optimal value of L should be infinite ($L = +\infty$) in order that messages can reach as many RSUs as possible.

2) Mobility: We consider a spare mobility network in which pedestrians are the majority. Limited battery, narrow throughput, low speed and unpredictable movement patterns self-evidently imply that the mobility component has a quite low delivery capability. Therefore, the optimal value of C should be the lowest: $C \simeq 0$. In addition, the L of mobility can be in the range from 1 to $+\infty$ depending on the expected compromise between delivery ratios and network-flooding effects. However, one of the simplest ways to find a near optimal value of L for mobility is to solve the equation of compromise between optimal and expected latency, described in [4], by letting $ED_{sw} = aED_{opt}$, where a is a factor for the delay constraint dictated by the application. As the numerical example in [4], we set a to 5. This means the expected delay is allowed to be up to five times of the optimal delay.

The expected delay of the optimal case, when $L = +\infty$:

$$ED_{opt} = \frac{H_{M-1}}{M-1} ED_{dt}, \qquad (1)$$

where $H_n = \sum_{i=1}^{n} \frac{1}{i}$ and M is the number of mobile sensor nodes.

The upper-bound of the expected delay of normal case with $1 < L < +\infty$ message copies is:

$$ED_{sw} \leq (H_{M-1} - H_{M-L} ED_{dt} + \frac{M-L}{M-1} \frac{ED_{dt}}{L}). \qquad (2)$$

Suppose that $1 < L \ll M$, we have $H_{M-1} \approx H_{M-L} \approx H_M$ and $M-1 \approx M-L \approx M$. Therefore, L in Equations 1 and 2 can be simplified as:

$$L \approx \frac{M}{aH_M}, \qquad (3)$$

Analogously, we can set other values of C and L for a sub-mobility group, for example, cars. Cars usually have high-speed and long-range communication so they certainly possess higher C than pedestrians do. In fact, the more appropriate value setting of these two parameters, the better performance of networks.

We remarked earlier that there is another solution to find the optimal Number of Copies for the mobility component in a heterogeneous architecture as Figure 1. In fact, Equations 1 and 2 are derived on the assumption that the network has only mobile nodes and the destination of a message is a random mobile node. However, our considered network is heterogeneous with various types of components, including stationary and mobile nodes. Furthermore, the comment destination is a common sink fixed in the center of the map. Therefore, the expected delay of $L = +\infty$ and $1 < L < +\infty$ cases will be slightly different. We are going to investigate this aspect in a journal since it takes a lot of space to describe the solution. However, the estimated value of L for mobility calculated by applying Equation 3 is almost consistent with our simulation results.

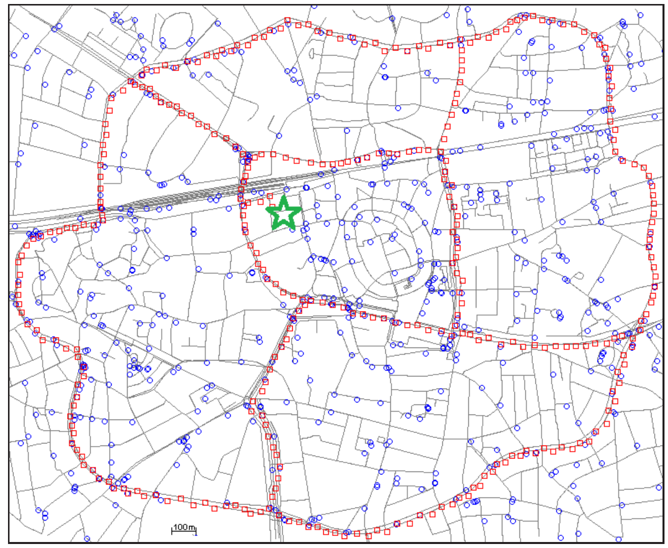


Fig. 2. Simulation on the city map of Enschede.

IV. Evaluation

A. Simulation Setup

We reuse the simulation set up by Le et al. [1] except for a few changes in parameter configuration, such as the net bit rate, message and buffer sizes. For a realistic setting, the map of the city of Enschede, measuring approximately 3000 by 3000 meters, is imported into our simulation. We create several submaps for RSUs, roads for cars, paths for pedestrians, and routes for busses. There are 336 RSUs deployed in linear formation at the outer ringroads, inner ringroads, and four radiating roads leading to the center of the city. Pedestrians are able to ramble everywhere, but cars can only run on roads. Moreover, four distinct bus lines in the city are mapped onto the simulation. For each bus line, two busses are assumed to serve the route.

Figure 2 illustrates a screen shot of the simulation. The red square dots □ are RSUs fixed on main roads, and the blue circle dots o represent mobile nodes including pedestrians, cars, and busses. The green star ★ at the center is the common sink.

Since our main concern is the contribution of pedestrians to data dissemination, the number of pedestrians will be varied between 258 and 1058. One thousand of pedestrians seems too few but indeed realistic because the population density of Enschede is just about $1000/km^2$ and we assume roughly one tenth of civilians walking in streets at a certain time. The simulation also includes 50 cars. The initial positions of cars and pedestrians are randomly distributed. The speed of cars and pedestrians are randomly generated in ranges of from 10 to 40 km/hr and 1.8 to 5.4 km/hr, respectively. Since pedestrians walking at almost uniform speed form the majority of the

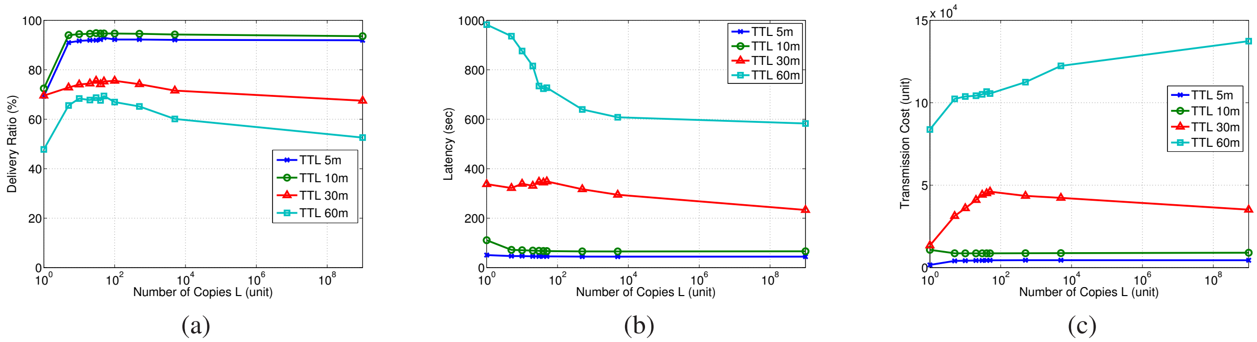


Fig. 3. Unified performance with varying the number of copies

network, the effect of the mobility velocity on performance results is not studied in this simulation. As mobile internet and mobile services are expensive, we only consider built-in short range interfaces, such as Bluetooth and WiFi. We assume that all mobile phones possess Bluetooth Version 2.0 at 2 Mbit/s net bit rate with a 10 m radio range, while only smartphones have a WiFi interface at net bit rate of 11 Mbit/s with a 60 m radio range. According to Eriksson et al. [14], the 11 Mbit/s rate overall performs much better than the other $802.11b/g$ rates for vehicles. Fifty percent of pedestrians own smartphones, and the rest uses regular phones without the WiFi interface. RSUs have both interfaces as do smart phones. However, the remaining nodes, cars and busses, possess WiFi only because Bluetooth is unrealistic at high speeds.

The Random Shortest Path Map Based Movement model (RSPMBM) [1] is applied for most pedestrians and all cars with travel distances ranging from 50 to 500 m and 500 to 5000 m, respectively. Cars certainly can go farther than 5000 m, but they would be out of the simulation area of 3000 m by 3000 m. In the RSPMBM model, a pedestrian would walk for a short distance, typically $50-500$ m, but not usually for long distances such as $500-5000$ m. For a long distance, he or she will tend to drive a car. In addition, we assume that there are always 100 pedestrians who just stroll around the city and modeled by the Map Based Movement model (MBM) [12]. These pedestrians simply follow Random Walks along map paths. Modeled by the Bus Movement model, busses shuttle between the bus stops along fixed routes. Finally, among pedestrians, we assume that there are 100, who prefer taking the bus and are modeled with the Bus Traveler Movement model [12]. They walk to a nearby bus stop and wait to catch a bus to take them to their destinations.

During interval time, which is randomly drawn from 25 to 35 seconds, an event happens at a random place in the city. Nearby mobile phone sensors, which are carried by pedestrians or cars, will measure data together. A node will be assigned as a cluster head that is responsible for distributed data processing to detect such event. A message, which contains inferred information about the event, is created at the cluster head and ready for dissemination. In other word, there will be 732 unique messages created during simulation of 6 hours in real world. In our study, RSUs do not generate messages, but act as a communication backbone. In fact, we intend to use RSUs to further improve the reliability and accuracy of detection in future. For message size setting, as [15] shows that the optimal message size for MAC layer is approximate 500 bytes, a random size between $0.5-1$ Kbytes is assigned to each message. We remark that 500 bytes suffice for most kinds of data in a sensor network, such as, humidity, temperature, or toxic chemical intensity. Since the common buffer length for the MAC layer is from 25 to 50 packets, the message buffer is set at 25 Kbytes for all nodes in most of the simulations in this paper, except when studying the effects of varying the buffer size.

Data dissemination is simulated using our proposed algorithm. As discussed in Section 3.C, the Delivery Capability C of RSUs is set as one, and that of mobility is set as zero. The Number of Copies L of mobile nodes are varied from $1..1e+9$. Meanwhile, L of RSUs are set as infinite for all cases. Furthermore, a comparison with First Contact, Spray and Wait, Epidemic, and ProPHET is also made. For the same reason as mentioned in [1], Message Ferries (MF) is not used for busses in our simulation since MF is more suitable for carrying messages from city to city.

B. Evaluation Metrics

We evaluate the performance of the Unified algorithm against three metrics: delivery ratio, latency, and transmission cost.

- *Delivery Ratio*: The number of messages that convey information of detected events and successfully arrives the common sink is divided by the total number of events. Note that each unique message contains information about only one event.
- *Latency*: The delay between the moment that an event message is created at a mobile node and the time the message reaches the common sink.
- *Transmission Cost*: The total number of messages, including copies, is divided by the number of successfully delivered messages at the common sink.

C. Results and Discussion

Using above configured scenario, we evaluate performance of the Unified algorithm in terms of average delivery ratio, latency, transmission cost. FIFO is used to sort messages, in line 14 of Algorithm 1. The compared algorithms also used FIFO to sort message buffers to make a fair comparison.

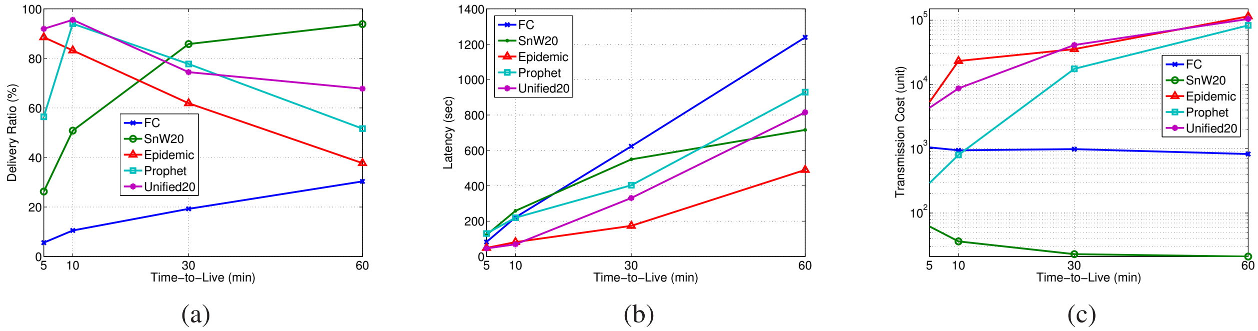


Fig. 4. Algorithms comparing

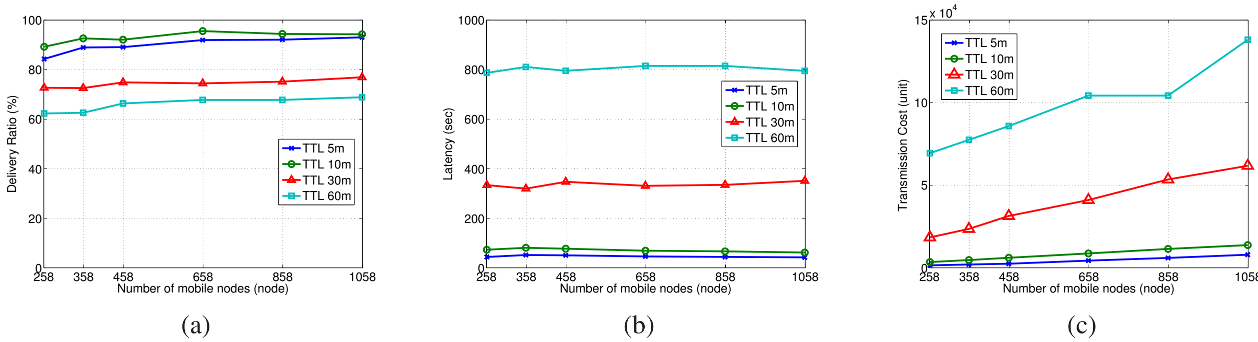


Fig. 5. Unified performance with varying the number of mobile nodes

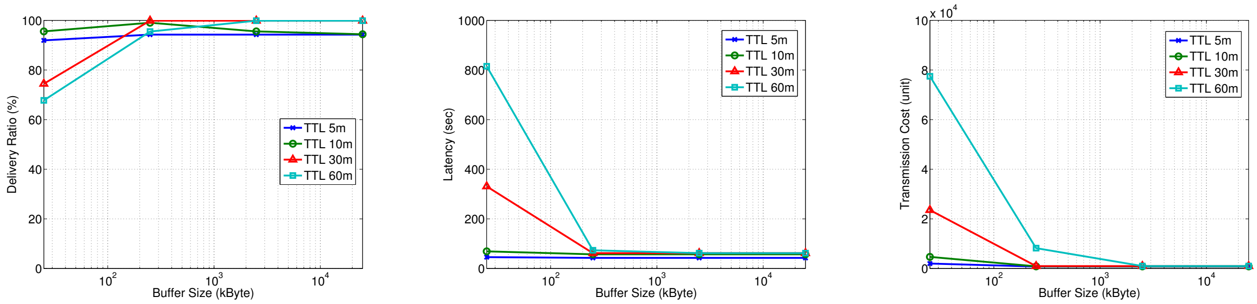
The results demonstrate that the Unified algorithm performs significantly better others in most cases.

Figure 3.a plots the overall delivery ratio of the Unified algorithm when the Number of Copies L of mobile nodes increases from 1 to $1e + 9$. There are 658 mobile nodes moving as models described in Section IV.A. By benefiting from the RSU network, Unified can obtain quite high delivery ratios even when $L = 1$, at least 47.81% with $TTL = 60$ minutes. The delivery ratios are significantly improved when L changes from 1 to 5 since there are more copies broadcast. However, the delivery ratios remain constant and even decrease a little if L increases continuously, especially with large TTL. Large L and TTL cause the flooding effect similarly to that for Epidemic. In addition, Unified gives high delivery ratios with short TTL messages because the algorithm can drop more expired messages, which probably have already been delivered, allowing it to receive new ones.

Figure 3.b shows the average latency of message delivery performed by the Unified algorithm. The algorithm can obtain good delivery delay for messages with short TTL even with a small Number of Copies L. However, for large TTL, Unified has to run with $L = 20$ at least to have good latency for $TTL = 60$ minutes. By looking at the delivery cost shown in Figure 3.c, the total number of messages including replicates over that of delivered messages, we can see that setting L correctly minimizes the tradeoff between latency and transmission cost. It seems that $L = 20$ is an optimal value so that the Unified algorithm obtains a high delivery ratio, low latency, and acceptable transmission cost for various values of TTL. By considering Equation 3, $L = 20$ is consistent with $L \approx 19$ ($a = 5$ and $M = 658$).

We also compare the Unified algorithm with First Contact (FC), Spray and Wait with the number of copies is 20 (SnW20), Epidemic, and ProPHET in Figure 4.a. All algorithms run with the same 658 mobile nodes and 25 Kbytes buffer size. Delivery ratios obtained by applying Epidemic and ProPHET significantly decreases when TTL increases, but SnW20. Observation shows that Unified takes advantage of Epidemic of Epidemic when TTL is short, and that of Spray and Wait when TTL is long. First Contact, of course, gives a very poor delivery ratio because it has only one copy of messages. Figure 4.b shows that the Unified algorithm has the second lowest average latency because it can take advantage of the RSU network. Clearly, by flooding the network with multiple copies of messages, Epidemic has optimal latency. Overall, larger TTL results in a longer delay since messages can stay longer in the buffer before being delivered. In addition, transmission costs of algorithms are shown in Figure 4.c. The transmission cost of Unified is as high as that of Epidemic when TTL is greater than 30 minutes. Since one of our objectives is disseminating information, high transmission cost, which means that more messages are broadcast, is not the most demanding.

To investigate the effect of the network size, we vary the number of nodes but keep the buffer size constant at 25 Kbyte. Results show that the network size does not significantly affect the performance of the Unified algorithm, as shown in Figure 5. Even the number of nodes vary widely from 258 to 1058, the existence of the RSU network makes that delivery ratios and latency almost remain constant in Figures 5.a and 5.b for various $TTLs$, respectively. Only transmission costs monotonically increase with the number of mobile nodes.



(a) (b) (c)

Fig. 6. Unified performance with varying the buffer sizes

This is expected since the more nodes, the more transmissions.

Figure 6 shows that the performances of Unified increase significantly in term of delivery ratio, latency, and transmission cost when the buffer sizes are bigger since more messages can be exchanged during a limit contact time. However, when the buffer size is greater than 250 Kbyte, holding approximately 375 packets, saturation is reached so that increasing the buffer size hardly improves the performance, as expected. Even with a finite buffer size, a pair of nodes will not be able to exchange all messages in buffers due to limited contact duration. We remark that the optimal value of buffer sizes depends on the movement model as long as different models have different contact times. As discussed in [1], the Random Shortest Path Map Based Movement is quite realistic. Therefore, the results shown in Figure 6 are realistic and give a good prediction for real experiments in future. Note that we set the number of mobile nodes as 658 for all simulations in Figure 6.

V. Conclusion

This paper is motivated by a real-world sensor network system, which comprises various kinds of sensor nodes such as mobile phones carried by pedestrians, cars, and busses. We observed that existing routing algorithms cannot perform well on such a heterogeneous architecture since algorithms were originally designed to serve either connected ad-hoc networks or delay-tolerant sensor networks. Even an attempt to use specific algorithms for a variety of architecture components would fail because it is hard to exchange messages between the distinguished algorithms. Therefore, this work proposes the Unified routing algorithm that builds on the advantages of existing algorithms to obtain better performance. Particularly, the Delivery Capability and Number of Copies are used as two major parameters to match the algorithm with various types of sensor components. Fixed infrastructure, such as a RSU network is designedly set with a high Delivery Capability and a large Number of Copies. Meanwhile, the mobile phones are set for low Delivery Capability and a small Number of Copies.

The results from simulations show that the performance of our proposed algorithm significantly exceeds that of existing algorithms, in particular, the delivery ratio overall is from 20% to 300% better than existing algorithms. Our future work in this area will involve a new message sorting mechanism, based on message priority and time-to-live, to further improve the deliver ratio and delivery speed. The implementation of the Unified protocol on a testbed is also planned.

Acknowledgment

This work is supported by the SenSafety project in the Dutch Commit program, www.sensafety.nl.

References

[1] V.-D. Le, H. Scholten, and P. Havinga, "Towards opportunistic data dissemination in mobile phone sensor networks," in *Proc. The Eleventh International Conference on Networks (ICN 2012)*, 2012.

[2] T. Spyropoulos, K. Psounis, and C. Raghavendra, "Single-copy routing in intermittently connected mobile networks," in *Proc. Sensor and Ad Hoc Communications and Networks (SECON)*, 2004, pp. 235–244.

[3] A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks," Department of Computer Science, Duke Univeristy, Durham, NC, Tech. Rep., 2000.

[4] T. Spyropoulos, K. Psounis, and C. Raghavendra, "Spray and wait: An efficient routing scheme for intermittently connected mobile networks," in *Proc. of ACM SIGCOMM Workshop on Delay-Tolerant Networking (WDTN)*, 2005.

[5] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and focus: Efficient mobility-assisted routing for heterogeneous and correlated mobility," in *Proceedings of the Fifth IEEE International Conference on Pervasive Computing and Communications Workshops*, 2007, pp. 79–85.

[6] A. Lindgren and A. Droia, "Probabilistic routing protocol for intermittently connected networks," *Internet Draft draft-lindgren-dtnrg-prophet-09, Work in Progress*, 2011.

[7] Y. Xian, C. Huang, and J. Cobb, "Look-ahead routing and message scheduling in delay-tolerant networks," in *Proc. IEEE Conference on Local Computer Networks (LCN)*, 2010.

[8] B. K. Polat, P. Sachdeva, M. H. Ammar, and E. W. Zegura, "Message ferries as generalized dominating sets in intermittently connected mobile networks," *Pervasive and Mobile Computing*, vol. 7, pp. 189–205, 2011.

[9] C. E. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (dsdv) for mobile computers," in *Proc. of the SIGCOMM 1994 Conference on Communications Architectures, Protocols and Applications*, 1994, pp. 234–244.

[10] C. E. Perkins and E. M. Royer, "Ad-hoc on-demand distance vector routing," in *IEEE WORKSHOP ON MOBILE COMPUTING SYSTEMS AND APPLICATIONS*, 1999, pp. 90–100.

[11] S. Jain, K. Fall, and R. Patra, "Routing in a delay tolerant network," in *Proc. of ACM SIGCOMM on Wireless and Delay-Tolerant Networks*, 2004.

[12] A. Keranen, J. Ott, and T. Karkkainen, "The one simulator for dtn protocol evaluation," in *Proc. of the 2nd International Conference on Simulation Tools and Techniques(SIMUTools)*, 2009.

[13] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley, 2001.

[14] J. Eriksson, H. Balakrishnan, and S. Madden, "Cabernet: Vehicular Content Delivery Using WiFi," in *14th ACM MOBICOM*, San Francisco, CA, September 2008.

[15] V. Mitlin, "Optimal mac packet size in networks without cut-through routing," *IEEE Trans. on Wireless Communications*, vol. 2, pp. 901 – 910, 2003.