

Evaluation framework for adaptive context-aware routing in large scale mobile peer-to-peer systems

Ansar-UI-Haque Yasar · Davy Preuveneers · Yolande Berbers

Received: 1 November 2009 / Accepted: 23 July 2010
© Springer Science+Business Media, LLC 2010

Abstract The proliferation of novel wireless network technologies creates new opportunities for complex peer-to-peer information dissemination systems. A key challenge that remains in this area is how to select the best algorithms and protocols to communicate effectively on a large scale. In this paper, we focus in particular on large scale mobile networks where effectiveness depends on the requirements of the application at hand and on the context of the peers in the network. We propose a framework for context-aware adaptive information sharing that allows the evaluation of and comparison with alternative information routing protocols using network metrics to measure a variety of quality attributes of the information dissemination protocols. These metrics can then be used to verify which protocol is best suited for a particular application. We illustrate our evaluation framework with different information dissemination protocols in an inter-vehicle communication scenario.

Keywords Context-awareness · Adaptive routing · Peer-to-peer systems · Mobile networks · Scalability

1 Introduction

The embedding of technology in everyday objects has developed to such an extent that a variety of information can be collected by various commodity devices within our vicinity. By intelligently disseminating the information to each other in a p2p fashion not only the end-users but also other nearby devices in the area can make use of this information in order to perform context sensitive tasks. It enables these devices to share information they cannot sense themselves. Figure 1 shows a home network of domestic equipment and illustrates how an RFID scanner in a fridge can detect if we are about to run out of milk when you are making a nice cup of coffee. This information can be shared to collect items for a grocery list on the laptop. Other sensors can detect our presence in the house and play our favorite music on one of the multimedia systems. The previous network is rather small and static in nature which makes information sharing rather simple. Intelligent information sharing and retrieval in large scale evolving networks is not that straightforward, as processing the information to check where and when it is relevant and making sure it is routed in a p2p fashion to the right destination is a challenge that has not been fully addressed in the literature.

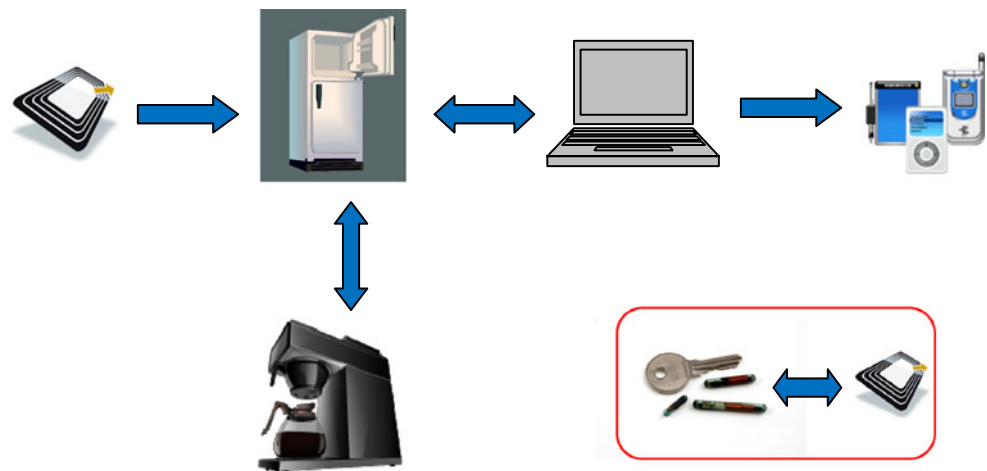
P2P data dissemination is an active research domain where various schemes, algorithms and mechanism have been proposed [1–3] to overcome the dynamicity and scalability of a network. Traditionally in peer-to-peer systems these concerns are handled by replication or redundancy in centralized and static networks [4]. But these solutions are not readily applicable to any type of distributed environment, such as vehicular networks which are highly mobile and large scale in nature. In our research we focus on intelligent retrieval and dynamic routing of

A.-U. Yasar (✉) · D. Preuveneers · Y. Berbers
Department of Computer Science,
Katholieke Universiteit Leuven,
Celestijnenlaan 200A bus 2402,
3001 Heverlee, Belgium
e-mail: ansarulhaque.yasar@cs.kuleuven.be

D. Preuveneers
e-mail: davy.preuveneers@cs.kuleuven.be

Y. Berbers
e-mail: yolande.berbers@cs.kuleuven.be

Fig. 1 Information embedded in the home environment



contextual information with adaptive information dissemination based on a feedback mechanism from immediate neighbors that expresses changes of interest. We will discuss the approach later on in more detail.

With a variety of potential adaptive routing schemes available for information dissemination in large scale mobile p2p systems, we need an evaluation framework to quantify various aspects of these routing schemes in order to compare them and select the most appropriate one for a given application. In this paper we list down a set of generic quality parameters to analyze and measure various adaptive routing schemes for large scale mobile networks. These quality parameters help us to identify potential strengths and weaknesses of a particular mechanism. We can also see the effects of each quality parameter in a particular mechanism creating a path for continuous improvement. In this paper, we take large scale mobile vehicular networks as one of the potential real life cases and analyze the evaluation framework in a simulated environment using OMNET++.

We will describe our motivating scenarios in Section 2. In Section 3, we will discuss about some of the adaptive routing schemes available with requirements for supporting large scale p2p systems. We describe our evaluation framework having a set of several generic quality metrics in Section 4. We end this paper with a comprehensive discussion and suggestions for future work in Section 5.

2 Motivating scenarios

In this section we describe two motivating scenarios enabling the development of a middleware for adaptive routing in the domain of mobile and scalable context-aware inter-vehicle communication. Vehicular networks are one of the modern realistic examples of scalable peer-to-peer systems interacting with each other to perform several

context sensitive tasks discussed in our scenarios and shown in Fig. 2.

2.1 Deployment of emergency response teams and emergency message dissemination to vehicles in a traffic incident

Deployment of Emergency Response Teams to a traffic incident has always been a crucial point with authorities. Traditionally in case of an incident information is sent to the concerned authorities about the type, location and time of incident over the cellular or wired telephone networks and there might be a road traffic jam at the same place. The problem with the current system is that vehicles are often informed too late and the message itself is usually broadcasted to all the vehicles, also to those that are not in the neighborhood of the accident.

Let us consider a scenario where an accident occurs between two vehicles travelling from Leuven to Brussels on the highway E40 causing a traffic jam. A vehicle owner on the same side of the road informs the emergency response teams using his cellular phone after the incident has occurred. The emergency response teams arrive on the location after 10 min to rescue the victims and to put up a sign informing about the incident 500 m away to inform the upcoming cars so that they can change their routes to avoid a traffic jam. But within this time frame quite a large number of vehicles are blocked on the road due to the accident which might take time to clear up. The types of interactions in this scenario can either be in the form of a query or a message, for example:

1. Notify the location of the accident
2. Query for a doctor or emergency response team near to the accident
3. Send emergency signal only to vehicles traveling towards the accident.

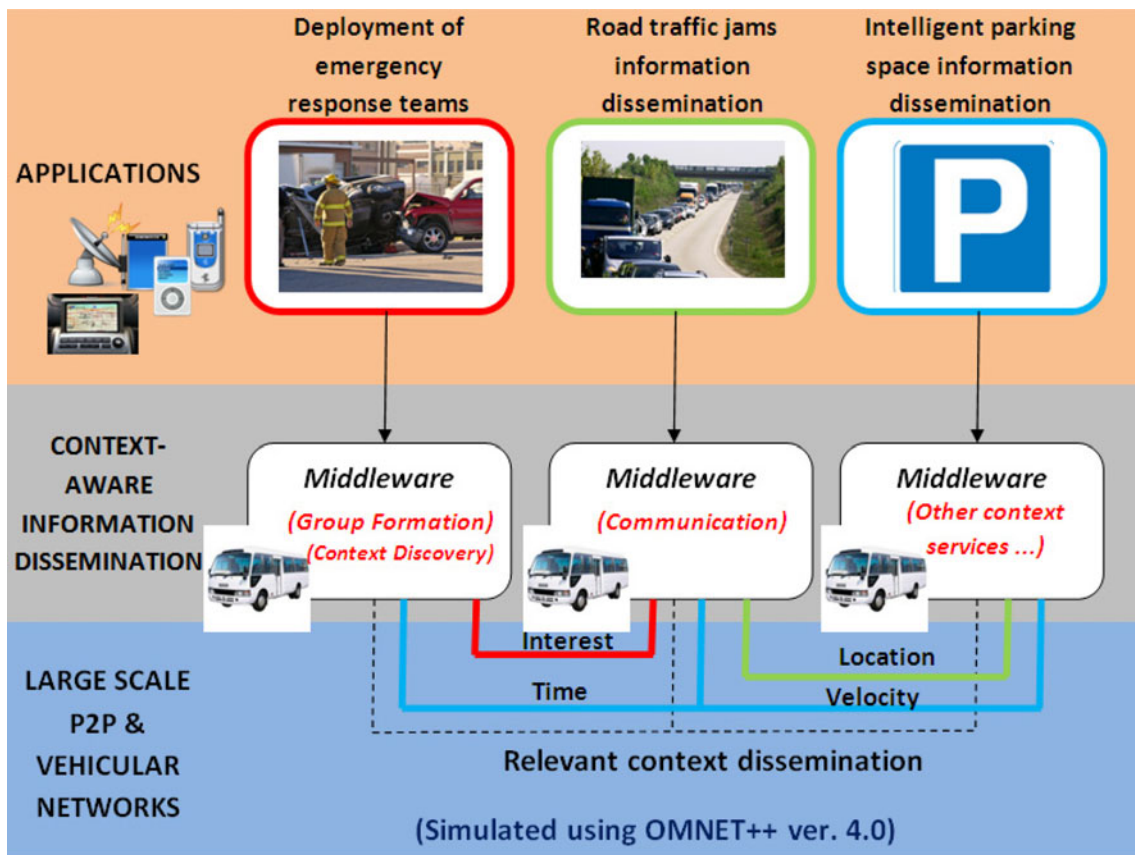


Fig. 2 Information dissemination in p2p systems for context-sensitive tasks

The participants in a vehicular network will have interactions in an ad hoc p2p fashion. This network may possibly span a large geographical area making a broadcasting algorithm to send all information to everybody less useful as in the home network in the introduction.

2.2 Intelligent parking space information dissemination to vehicles in a metropolitan city

Nowadays, most of the new vehicles have an embedded Global Positioning System (GPS) device to assist the drivers while driving from one location to another. Let us take a typical case of Brussels city during the rush hours when there are thousands of vehicles on the roads. Brussels is one of the most popular cities in Europe and a tourist attraction as well. At all times many vehicles are in search for a parking space near to their destination. Finding a free parking space is a hassle, and could be improved if information about free parking space is intelligently advertised to the vehicles. Even the installed GPS is of no use in this situation. Traditionally the parking information is displayed on electronic boards within the city for different parking spots. Timely delivery of up-to-date

information is important. An advertisement of a vacant parking space is no guarantee that the parking space is not already occupied by another vehicle if the information about that free parking space was too old and no longer relevant. In those cases vehicles park too far away from their destinations or waste a huge amount of time in search for a nearby parking space.

Some queries may have additional constraints, e.g. looking to park your car for several hours in Brussels city near the central station but with a maximum parking charge of 1 EUR/h. This would rule out the geographically nearest parking space with a higher parking fee of 1.50 EUR/h or another parking space with a maximum parking duration of 1 h. Such constraints can be taken into account in vehicular networks as well in order to make context-aware intelligent decisions. Different types of interactions in this scenario either in the form of a query or a message are listed below:

1. Look for a parking spot within 500 m
2. Look for a free parking space
3. How long can I park on that parking lot?
4. Advertise to other cars interested in a parking spot about a free space.

The participants in such a scenario will have such meaningful interactions in a p2p (ad hoc) manner disseminating the relevant information intelligently.

There are many application domains in a large scale p2p system and each of this domain pose a different set of requirements to work efficiently. For example, information dissemination an application to provide assistance to patients in a hospital is time-critical as compared to an application for sharing a taxi at an airport. In order to satisfy these requirements we also have a number of context-aware algorithms and protocols discussed in detail later in the paper.

3 Adaptive routing schemes in large scale peer-to-peer systems

Peer-to-peer systems can have many interesting application dependant technical features like decentralized control, self-organization, adaptation, mobility and scalability [5]. These systems are sometimes highly dynamic in nature sharing massive amounts of information at a particular instance of time. As the context continuously evolves whenever situations change, we need adaptive routing schemes in order to support context-aware communication in a p2p system. Many such adaptive routing schemes are available for use in large scale p2p systems, including Pastry, gossip-based, directed diffusion, two-tier dissemination, gradient and plain broadcasting along with our relevance backpropagation algorithm which we have proposed earlier [6]. Which one is the most effective depends on the application at hand and what relevance criteria are important for the application's context-aware information delivery. In this section we will briefly discuss the requirements for these adaptive routing schemes in large scale p2p systems and some routing schemes with some potential problems associated with each routing scheme.

3.1 Requirements for adaptive routing in large scale p2p systems

In order to optimize context-aware information dissemination we have identified a set of requirements supporting such communication in our previous research work [6]. We briefly summarize them below.

R1: Location and direction-aware delivery of messages

It is always desirable to know the exact location of an incident for context-aware applications e.g. in the scenario 2.1 in case of an incident on the road the authorities should be notified about the exact location to react fast. Similarly, a context-aware application should be able to sense, manipulate and disseminate context information about direction and velocity of vehicles in the network to predict certain situations like traffic congestions or traffic accidents in

specific regions. Moreover, to optimize the routing and delivery of a message, location-specific traffic (direction) patterns have to be taken into account. These patterns allow predicting the likely movement of the traffic flow in a specific area and depending on the time of the day.

R2: Temporal relevance

Temporal relevance is the desired behavior of a context-aware application dealing with timeliness of information and routing efficiency. In a context-aware application on time arrival of information using an efficient route has always been a challenge. For example, if a road maintenance work is underway on 20th November 2009 between 10am and 5pm at Naamsestraat, Brusselsestraat and Lei in Leuven city, the information about traffic congestion or road condition is only relevant on this specific date and time.

It is required that only the relevant context information arrives at a particular node on the right time and place. Temporal relevance involves efficient filtering of irrelevant information at intermediate nodes for optimal routing and faster delivery of context information. Prediction techniques may enable the optimization of the message routing and timeliness.

It is desirable to test and analyze these requirements so this imposes a new requirement for our simulated framework for large scale vehicular networks.

R3: Analyze throughput, communication overhead and delivery efficiency

It is quite important to be able to quantify how much data transmitted over the network is actually used by network peers for any given communication protocol scheme on an application basis. This quantification will guide the researchers to properly analyze, improve and compare various algorithms and protocols based on the parameters like throughput, communication overhead and routing efficiency.

3.2 State-of-the-art routing schemes for large scale p2p systems

In this section we will briefly discuss some of the adaptive routing schemes currently in use for large scale p2p systems.

3.2.1 Pastry

Pastry is a scalable distributed object location and routing scheme for large scale p2p systems. It can perform application-level routing in a relatively large overlay network of nodes connected with each other. Each node in the Pastry network has a unique identifier. When presented with a message and a key, each node efficiently

routes the message to the node having the lowest key value. Each node also keeps track of its immediate neighbors with a way to notify each other about a node arriving or leaving the network [5].

3.2.2 Gossip-based schemes

In a similar fashion to an epidemic, rumors may be spread through a community: people gossip with their neighbours and friends, who in turn gossip with people in their proximity. In this way, the rumor is propagated through the population [7]. The term epidemic depicts how the information is disseminated through the network. Nodes pass information to a small number of other nodes, randomly selected from their neighborhood, in much the same way that an infection spreads from person to person. During each cycle of the epidemic an infected host may come into contact with other hosts. These target hosts are chosen randomly and may, or may not, already be infected. The newly infected nodes then, in turn, relay the infection to further nodes [7].

3.2.3 Directed diffusion

It is the data-centric communication technique widely used for wireless sensor networks. The vehicles request the context information by periodically broadcasting an interest for the required data. Each node will create a link with other nodes or a context provider from which it receives the context information of interest. The link also specifies the data rate and the direction towards which the context information should be sent. Once the link is created between the nodes and the context provider, the context provider will start sending information of interest to the nodes probably along multiple paths. As soon as the node wants to receive the context information, it will select a specific neighbour from which it will receive the information later on, thus defining a directed broadcast of the context information over a large scale network.

3.2.4 Two-tier data dissemination

It is a decentralized architecture where a grid structure is used to divide the network into cells. The context providers located at the boundary of the cell need to forward the context information to other cells. The context information is flooded within a cell. One tier is the cell at the nodes current location and the other one at the cell's boundary. The query is first propagated over the network to create a path between the node and the context provider and then the same path is used for the propagation of the context information. This technique involves a lot of intelligent routing mechanisms along with information storage overhead creating complexities in the libraries for large scale networks application development.

3.2.5 Gradient broadcast

This technique makes use of a cost variable during the transmission of context information. Initially the context providers set the cost to reach a node at infinity. The information is then broadcasted over multiple paths in the network where each of the intermediate context provider or a node calculates the cost of receiving the message [8]. At the end each of the context providers or nodes would have calculated the cost for it to send the context information to a particular node. The cost data is later on used to optimally transmit the context information over the network with a minimal cost. This is highly efficient for transmission of context information over a large scale vehicle network but at the same time it creates overhead for each node to calculate the cost information. This could be an issue given the limited processing capabilities of nodes in large scale mobile networks.

3.2.6 Plain broadcast

The plain broadcasting technique the context information received by a node is stored locally and then the same information is forwarded using a re-broadcast to others. In a large scale, dynamic and mobile network like a vehicular network plain broadcasting may overload the network especially in the case of high traffic volumes thus violating some of the requirements mentioned earlier in Section 3.1. This technique is very useful in the case where the network under consideration is relatively small because efficiently routing the context information to specific nodes is more expensive in terms of network traffic, processing and communication overhead.

3.2.7 Relevance backpropagation

Our *relevance backpropagation* algorithm [6, 9] (as shown in Algorithm 1) is step-wise discussed as under;

- i. The context information is initially plain broadcasted to the adjacent nodes unless maximum number of hops is reached. Each node forwards the information to its immediate neighbours and waits for the feedback backpropagated to it.
- ii. Intermediate nodes will decide based on feedback backpropagated by the neighbouring nodes to reduce the number of peers to forward the information to. The feedback which is backpropagated contains information about the relevance of the context information received earlier as shown in Algorithm 1 lines 4, 9, 22 and 24.
- iii. Each forwarding node reduces the hop counter, adds its identification and marks the message relevancy tag if the information is relevant for its purpose as shown in Algorithm 1 lines 7, 8 and 16.

- iv. The feedback technique is based on context information like position, velocity, direction, time-to-live, interest etc. that helps decide whether the data that was received is relevant or not, and that also helps determine the information relevancy on the intermediate nodes. This way the irrelevant context information can be filtered out at an intermediate node.
- v. The feedback to the source node is propagated back if the context information is relevant, irrelevant, unused or duplicate information is received as shown in Algorithm 1 lines 4, 9, 22 and 24. This will reduce the information dissemination only to the interested nodes.
- vi. A vehicular network is highly dynamic in nature and application dependent. As the context information can be provided by the application itself the routing of the information is adapted accordingly and perhaps different for various applications.
- vii. All the nodes in the network send out a notification upon arrival in the network. So the network recalibrates itself if a new node sends an arrival notification or an old node no longer transmits the feedback after a certain period of time as shown in Algorithm 1 line 2 and 26.
- viii. In this mechanism the goal is to efficiently filter and intelligently route the relevant information as close to the source as possible in a dynamic network. The routing information also directly proportional to the relevancy of information.

Algorithm 1. Relevance Backpropagation (input: fromPeer, contextMessage)	
1	(messageRelevant, messageUnused, messageForwarded) = (false, false, false)
2	while (BeaconNewNode)
3	if (InFilterReceived(contextMessage.ID)) then
4	BackpropagateMessage(fromPeer, DUPLICATE, contextMessage.ID)
5	else
6	AddFilterReceived(fromPeer, context.Message.ID)
7	if (InFilterRelevant(contextMessage)) then
8	messageRelevant = true
9	BackpropagateMessage(fromPeer, RELEVANT, contextMessage.ID)
10	if (InFilterUnused(contextMessage)) then
11	messageUnused = true
12	LabelMessage(contextMessage, UNUSED)
13	else
14	LabelMessage(contextMessage, IRRELEVANT)
15	if (contextMessage.hopsLeft > 0) then
16	contextMessage.hopsLeft = contextMessage.hopsLeft - 1
17	for each Peer p in ForwardFilter (adjacentPeers, contextMessage.ID) do
18	messageForwarded = true
19	ForwardMessage (p, contextMessage)
20	if (not messageForwarded) then
21	if (not messageRelevant) then
22	BackpropagateMessage(fromPeer, IRRELEVANT, contextMessage.ID)
23	else if (messageUnused) then
24	BackpropagateMessage(fromPeer, UNUSED, contextMessage.ID)
25	end while
26	RecalibrateNetwork()

The conditions under which a node propagates back a feedback message are:

- Relevant context: The context information received by a node is either relevant for its own purpose or by one of its adjacent peers.
- Irrelevant context: The context information received by a node cannot be used for its own purpose and maximum number of hops has reached or it has no neighbouring nodes for which the context information might be relevant.
- Unused context: The context information received by a node is either not used or used infrequently. In this case the node will generate a message asking the transmitting peer to increase the transmission delay.
- Duplicate context: The context information received by a node has already been sent it by another node in the neighbourhood. In this case the node will ask the transmitting peer to stop sending this duplicate information in the future.

This algorithm is a best-effort algorithm which adapts itself according to the network configuration. The algorithm becomes intelligent with feedback information propagated in the network.

3.3 Potential strengths and weaknesses of the adapting routing scheme for p2p systems

Each adaptive routing scheme associated to the large scale p2p system has its own list of advantages and disadvantages listed briefly as under:

- i. In a large scale, dynamic and mobile network like a vehicular network plain broadcasting may overload the network especially in the case of high traffic volumes thus violating some of the requirements mentioned earlier in Section 3.1. This technique is very useful in the case where the network under consideration is relatively small because efficiently routing the context information to specific nodes is more expensive in terms of network traffic, processing and communication overhead [8].
- ii. State of the art protocols for p2p communication including Pastry, gossip-based and dissemination based schemes are not suited for context dissemination as the relevancy of context is not taken into account creating processing overheads in a potentially larger network [4].
- iii. The adaptive routing schemes like gossip-based schemes pose due to their nature a huge processing and communication overhead at each node in a potentially larger network [10].
- iv. Current peer-to-peer communication protocols are also inappropriate for scalable context-aware information dissemination as the relevancy of information cannot be determined at intermediate nodes during interaction between several nodes.
- v. For unstructured networks the algorithms like Gossip and Pastry are slow and impose high overhead taking into account the scalable and distributed nature of a network [1, 10]. Whereas, our relevance backpropagation algorithm handles the information dissemination based on the context relevancy which makes it scalable and better equipped for the distributed nature of p2p systems in general.

We proposed our relevance backpropagation algorithm in order to overcome these issues mentioned earlier optimizing the information dissemination in a large scale network to perform context-sensitive tasks.

4 Evaluation framework for information dissemination

In the previous section we discussed about some generic requirements for a large scale p2p system with a set of adaptive routing schemes fulfilling some or all of the mentioned requirements. But we still need to devise a mechanism to evaluate these routing schemes based on some quality metrics and later on fine-tune these schemes as per the demand of the applications. In this section we will discuss some of the generic quality parameters for a highly dynamic and mobile network. The quality parameters will help formulate and improve an intelligent information dissemination algorithm/mechanism based on the contextual properties of the nodes in a scalable and dynamic network. We have evaluated our relevance backpropagation algorithm with plain broadcasting and social filters for intelligent information dissemination in peer-to-peer systems as a case study of our framework based on the same quality parameters in a simulated environment over OMNET++. The following figures show several nodes having red and black colors in a dynamic and scalable vehicular network. The colors signify the contextual information interest of each node in the network based on the feedback provided. Figure 3 shows how node 1 disseminates information with repeated forwarding by nodes 2, 4 and 5. Nodes 4, 5, 6 and 8 have expressed an interest in the information node 1 can provide (color black), but due to the disconnection node 8 did not receive the information. Nodes 2, 3 and 7 have expressed no interest (color red), which means the processing and the forwarding of the message is overhead.

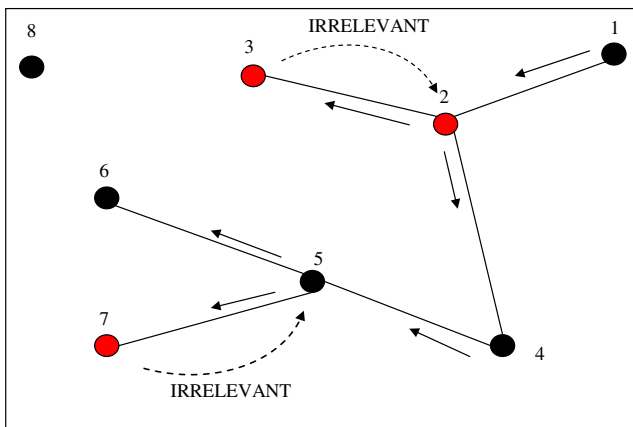


Fig. 3 Nodes in a mobile vehicular network

4.1 Generic quality metrics

In this section we will list and discuss several generic quality metrics for information dissemination in dynamic and scalable mobile networks.

4.1.1 Parameters to consider for a full network

Number of copies of messages (M_{copy}) This parameter directly deals with the total number of copies of a unique original message disseminated inside a network for a given time period. In Fig. 4, node 1 sends a unique message about road traffic congestion only interesting for node 6. Later during the transmission the copies of the same message is made 3 times in total by nodes 2, 4 and 5.

Number of dropped messages (M_{drop}) This parameter deals with the number of messages that were dropped by a node because they were neither interesting for the node itself nor for its neighbours for a given period of time. In Fig. 5, considering the same message transmission from node 1 for which node 6 has expressed an interest, the message also

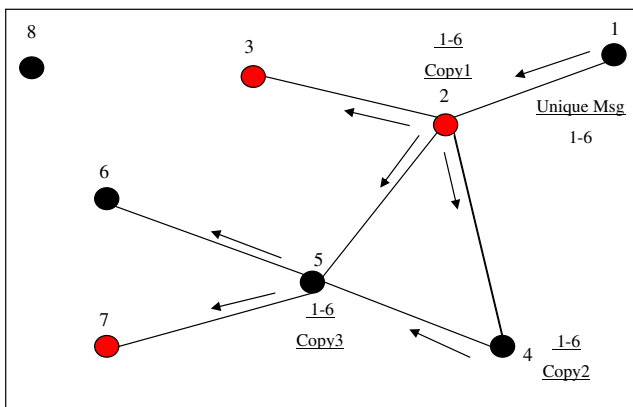


Fig. 4 Message transmission from node 1 to node 6, only relevant for the node 6

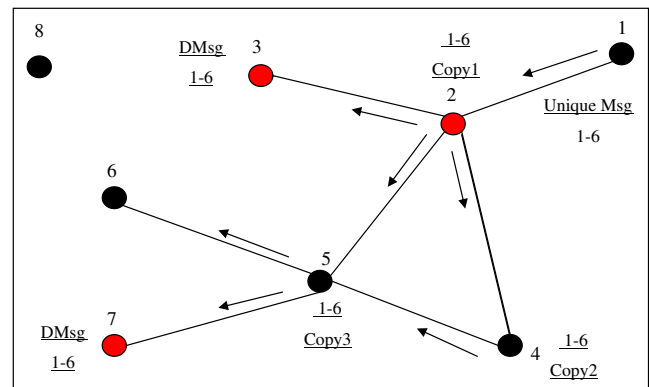


Fig. 5 Message transmission from node 1 to node 6, not interesting for nodes 3 and 7

arrives at nodes 3 and 7 which are not interested in the message. Thus, the unique message (1-6) is dropped twice being irrelevant for some of the nodes participating in the network.

Number of original unique messages (M_{us}) The total number of original unique messages generated by all the nodes in a network for a given time period. In Figs. 6, 5 unique messages are generated relevant for some of the nodes according to the contextual information being sent among the nodes.

Number of nodes (N) This parameter is a count of the total number of nodes present in a network for a given time period. This is an important parameter to measure considering the mobile and dynamic nature of a network.

Number of duplicate messages ($M_{duplicate}$) This parameter is a subset of the number of dropped messages in a network. It is the count of messages which were received more than once per unit time by a particular node in the network and later on dropped because it was received before and hence irrelevant. In Fig. 4, node 6 will have the

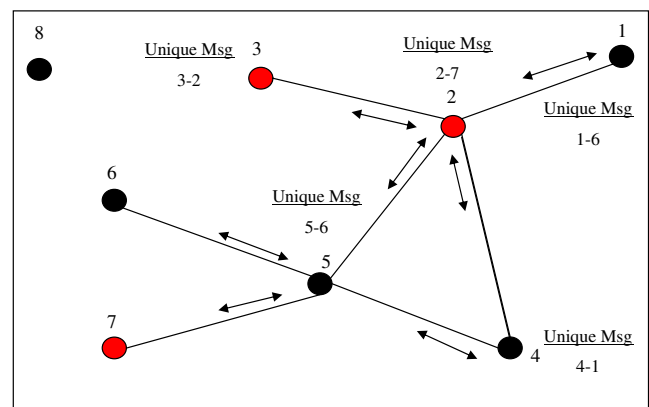


Fig. 6 Unique messages transmission

same message three times arriving from distinct routes in a network. So the receiving node will discard the duplicate messages.

Number of edges (E) The total number of edges present between all the nodes in a network at a given point in time. Considering Fig. 4, there are seven edges present between the nodes as all the nodes are not in direct or indirect communication range depending on the type of network being used.

Average number of connections (C/N) This parameter is a measure of the average number of connections present between the nodes in a network for a given time period.

Number of messages sent (M_{sent}) This parameter is the sum of all the sent messages by each node in a network for a given time period. For example in Fig. 4, node 1 sent one unique message for node 6.

Number of messages forwarded ($M_{forwarded}$) This parameter is the sum of all the forwarded messages by each node in a network for a given time period. For example in Fig. 4, node 1 sent one unique message for node 6 and the same message was forwarded six times by other nodes to others in the network.

Number of connections (C) The sum of all the unique connections established in a network for a given time period.

Number of disconnections (D) The sum of all the unique disconnections occurred in a network for a given time period.

Number of unique messages received by a node (M_{ur}) This parameter is the sum of the total number of unique message received by a node for a given time period. This parameter is related to the number of unique message nodes being sent. For example, in Fig. 4, node 6 receives only one unique message and two duplicates.

4.1.2 Generic metrics per nodes basis

Relevancy (R)

$$R = \frac{\sum_n ((M_{ur} + M_{duplicate}) - M_{drop})}{\sum_n (M_{ur} + M_{duplicate})}$$

If a node in a network receives a message from another node and the message is either relevant for itself or one of its (physically connected) neighbouring nodes then this message is considered to be relevant. For example, if in the

Fig. 3 node 2 receives a message from node 1 which is interesting for nodes 4, 5 and 6 then that message is considered to be relevant for node 2.

Sharing ratio (SR)

$$SR = \frac{\sum_n (M_{ur} + M_{duplicate})}{\sum_n M_{forwarded}}$$

Sharing ratio is the measure which tells us about the active participation of a node in a network in terms of the messages it receives and forwards to other nodes for a given time period.

Network traffic (NT)

$$NT = \sum_n (M_{sent} + M_{forwarded})$$

Network traffic is an important aspect to measure for a particular mobile network. This measure can tell us about all the messages disseminated inside a network for a given time period.

Message distance (MD)

$$MD = \frac{\sum_t E}{\sum_t N}$$

Message distance tells us about the route being taken by a message between a sending and a receiving node in terms of the hops count for a given time period.

Availability (A)

$$A = \frac{\sum_n M_{us}}{\sum_n M_{ur}}$$

Availability is the measure of the number of unique messages received by a node in the network with respect to the number of unique messages sent in the same network over a given period of time.

Communication overhead (CO) (at a sender)

$$CO = \frac{\sum_n M_{us}}{\sum_n (M_{ur} + M_{duplicate})}$$

Communication overhead is a type of overhead which is generated by a sending node in a network for a given time period. For example, in Fig. 4 if node 1 sends a message only interesting for nodes 4, 5 and 6 then node 2 will have a communication overhead as it has to forward this message to the nodes 3 and 7.

Processing overhead (PO) (at a receiver)

$$PO = \frac{\sum_n M_{ur}}{\sum_n (M_{ur} + M_{duplicate})}$$

Processing overhead is a type of overhead which is due to an additional processing step a node has to perform when it receives a message in order to determine its relevancy for a given time period. For example, in Fig. 4 if node 1 sends a message only interesting for nodes 4, 5 and 6 then the nodes 3 and 7 will have a processing overhead as they have to check the message relevancy and drop it later on if not relevant.

Filtering (F)

$$F = \left(1 - \left(M_{\text{forwarded}} / \sum_n (M_{\text{ur}} + M_{\text{duplicate}}) \right) \right)$$

This parameter is measured at each node where a node receives a number of messages but only forwards those messages to other nodes which are relevant for them for a given time period. There could also be some special kind of filters discussed later.

Social filters (static) using friend's information (SF) If we involve social networks with the mobile networks they act as special additional filters on top of the mobile network to limit information dissemination only between trusted entities using pre-defined static information. Using such a filter we can make sure that we only disseminate the information between the trusted entities in the network based on the social profiles.

Adaptive filters (dynamic) using relevance backpropagation (AF) Relevance backpropagation algorithm is a kind of adaptive filter which is dynamic in nature and which filters out unnecessary information flow between nodes at runtime in a network based on the feedback provided by the nodes. More details can be found in Section 3.

4.2 Simulated performance analysis

Evaluating and analyzing different adaptive routing schemes in a realistic large scale network is quite difficult in terms of resources and costs associated with it. In our earlier work [6], we have found our approach better in comparison with plain broadcasting currently popular in p2p system and later analyzed it using the same quality metrics. These quality metrics help in identifying different optimal routing schemes in a large scale p2p system. For example in Fig. 5, the node 6 requires some relevant unique information from node 1. In this case the unique information from node 1 arrives at node 6 via nodes 2 and 5 instead of nodes 2, 4 and 5 due to the lower usage of the overall network traffic, on-time arrival of message and various other factors. The choices for optimally routing the information in a large scale p2p system are made based on these quality metrics. Using the same metrics and

measures an adaptive routing scheme could be optimized as per the demand of an application. We can also analyze and eliminate various bottlenecks in a large scale p2p system, if any, found using these quality metrics.

In order to overcome this problem we can consider a realistic simulated environment like OMNET++ to run our evaluation. OMNET++ is a discrete event simulation environment. Its primary application area is the simulation of communication networks, but because of its generic and flexible architecture, it has been successfully used in other areas like the simulation of complex IT systems, queuing networks and hardware architectures. It can easily be adapted to simulate other scenarios including large scale, dynamic and mobile networks of vehicles. With the use of the simulated environment we can continuously test and improve the adaptive routing schemes, more details are discussed in Section 5.

We have used realistic vehicular mobility data [6] to perform the simulations using two variants of our relevance backpropagation algorithm one with social filter and the other one without them. We let the nodes move around like cars and let connections appear and disappear according to the range to other nodes. Several nodes acted as context providers whereas other nodes acted as context receivers. All nodes forward the information to their peers as long as the maximum TTL has not been reached and all context constraints are met. Figure 7 shows a visualization of the experiment with 27 nodes. We carried out the experiment for a period of 1 h.

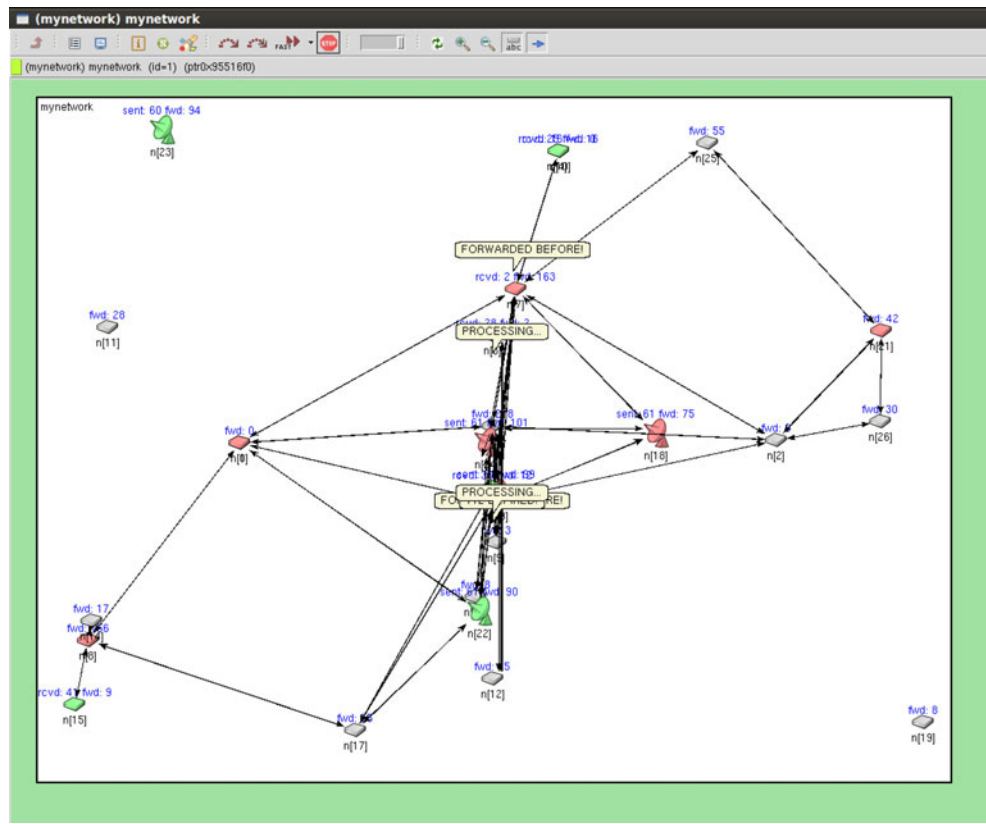
5 Discussion and future work

In this paper we present a layered approach as shown earlier in Fig. 2 where our reusable framework lies between the applications and the physical layer in a large scale p2p network. In this section we will discuss about our evaluation framework described earlier with reference to the results (as shown in Fig. 8) obtained from the simulated experimentation using OMNET++.

In addition to the context requirements for improving an adaptive routing scheme suitable for large scale mobile p2p systems, we also need to have a way to quantify some quality metrics discussed earlier in Section 4. In Fig. 8, we performed two simulated experiments; (i) using plain relevance backpropagation algorithm and (ii) using relevance backpropagation algorithm improved with the social filter to limit the information dissemination only between the trusted nodes.

Using our quality metrics we can analyze that by inducing a social filter we can limit the overall network traffic as assumed without affecting the relevancy of information in a large scale vehicular network. Similarly,

Fig. 7 Information dissemination in a vehicular network using OMNET++

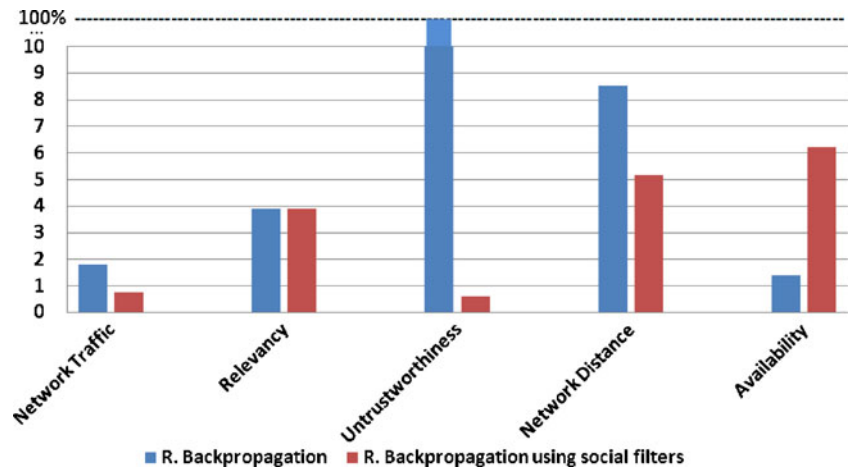


the availability of the reliable information in a potentially large vehicular network has enhanced significantly for each node to perform context sensitive computation. Our evaluation framework helps us to analyze and fine tune the information dissemination per node and per network on an application basis in a large scale mobile p2p system. We can easily quantify each metric according to the requirement of the p2p application, compare varieties of information dissemination protocols, and also modify the behavior

of an adaptive routing scheme by fine-tune the way messages are routed.

The biggest advantage of our framework compared to pure network simulators is that we can go beyond measuring network specific properties, such as bandwidth consumption, message delays, delivery failures, etc. We can now analyze the effects of context-aware information dissemination. Previously, we had to rely on information agnostic p2p routing protocols which did

Fig. 8 Simulated experimental results (%)



not take the relevance of the information at a particular into account for the routing of the information. In some cases this meant that information was disseminated through the network that had become irrelevant in the meantime. As the content of the message was only processed at the receiving nodes, the relevance could only be checked after it had already been delivered. In our relevance backpropagation algorithm, we take into account en route any context that is relevant for the relaying or filtering of the message, with a feedback mechanism to dynamically update the interests of nodes in the network to their direct peers. Our framework helps to analyze how information was routed and how it can be further optimized to improve one of the quality parameters or attributes that are of particular interest for an application.

As a part of future work we plan to extend our framework with the ability of having multiple applications—each with their own specific information dissemination requirements and protocols—being analyzed in parallel as if they are being executed and sharing information in the network concurrently. At this point our evaluation framework only supports investigating different algorithms for a single application at a time. Furthermore, to truly validate our evaluation framework, we plan to reproduce the simulations of our evaluation framework in a realistic environment using a set of adaptive routing schemes with a variety of scenarios to cross verify whether the ranking of the adaptive routing schemes in the simulated evaluation framework will be similar to the one in a real life experiment. It will be interesting to see if real networks with external influences have other side effects that the evaluation framework cannot take into account and whether it will alter any decisions made based on our previous outcomes.

Acknowledgements This research is partially funded by the Interuniversity Attraction Poles Programme Belgian State, Belgian Science Policy, and by the Research Fund K.U.Leuven.

References

1. Lindemann C, Waldhorst OP (2004) Exploiting epidemic data dissemination for consistent lookup operations in mobile applications. *SIGMOBILE Mobile Comput Commun Rev* 8(3):44–56

2. Medidi M, Ding J, Medidi S (2005) Data dissemination using limited flooding in wireless sensor networks. In: Yang LT, Arabnia HT, Wang L-C (eds) *ICWN. CSREA, Las Vegas*, pp 434–440
3. Landsiedel O, Gotz S, Wehrle K (2006) A churn and mobility resistant approach for dhts. In *MobiShare'06: proceedings of the 1st international workshop on decentralized resource sharing in mobile computing and networking*. New York, USA, ACM, pp 42–47
4. Roczniak A, Saddik AE (2009) Improving robustness of P2P applications in mobile environments. In *Peer-to-Peer Networking Applications*, Springer, pp 217–229
5. Rowstron A, Druschel P (2001) Pastry: scalable, decentralized object location, and routing for large-scale peer-to-peer systems. In *Lecture notes in computer science*, vol. 2218, pp 329–350
6. Yasar A, Preuveneers D, Berbers Y (2008) Adaptive context mediation in dynamic and large scale vehicular networks using relevance backpropagation. In *Proceedings of Mobility conference, ACM, Taiwan*
7. Williamson G, Stevenson G, Neely S, Coyle L, Nixon P (2006) Scalable information dissemination for pervasive systems: implementation and evaluation. In *Proceedings of the 4th international Workshop on Middleware For Pervasive and Ad-Hoc Computing (MPAC 2006)*. ACM, New York
8. Bokareva T, Bulusu N, Jha S (2004) A performance comparison of data dissemination protocols for wireless sensor networks. *Proceedings of IEEE Globecom Wireless Ad Hoc and Sensor Networks Workshop, IEEE, Dallas*
9. Preuveneers D, Berbers Y (2007) Architectural backpropagation support for managing ambiguous context in smart environments. *Proceedings of C. Stephanidis, editor, HCI, volume 4555 of Lecture Notes in Computer Science*, pp 178–187, Springer, Berlin
10. Risson J, Moors T (2006) Survey of research towards robust peer-to-peer networks: search methods. *Comput Netw* 50(17):3485–3521



Ansar Yasar is a Doctoral researcher on a scholarship at the Department of Computer Science of Katholieke Universiteit Leuven. He is a member of the Embedded Systems Task Force of the DistriNet Research Group. His other research interests include; wireless Networks: Location Determination of Wireless Nodes in WLAN, Software Engineering: Bayesian Networks and Secure Software Engineering and Multimedia

Networks: Real Time Multimedia Streaming support for Mobile devices.



Davy Preuveneers is a post-doctoral researcher at the DistriNet research group of the Department of Computer Science at the Katholieke Universiteit Leuven in Belgium. He received his PhD degree in 2009, his MSc in artificial intelligence in 2003, and his MSc degree in computer science in 2002 from the Katholieke Universiteit Leuven. His research interests include middleware and service oriented architectures for context-aware service

interaction and application adaptation, and software engineering for embedded software in particular for mobile and ubiquitous computing.



Yolande Berbers is a professor at the Department of Computer Science of the Katholieke Universiteit Leuven (Belgium), where she is a member of the research group DistriNet (Distributed Systems and Computer Networks). In general her research interests include software engineering for embedded and real-time systems, model-driven architecture, middleware infrastructure, context-aware computing, ambient intelligence and ubiquitous computing and system support for distributed applications.