Multimedia Data in Hybrid Vehicular Networks*

Bo Xu Ouri Wolfson Jie Lin
University of Illinois at Chicago, USA
{boxu,wolfson,janelin}@uic.edu

ABSTRACT

In this paper we study querying multimedia data such as video and voice clips in hybrid vehicular networks that consist of vehicles that are capable of both infrastructure-less short-range communication and infrastructure communication. We introduce a set of query processing strategies which differ from each other in terms of push versus pull, whether or not infrastructure communication is utilized, and whether metadata dissemination is separated from multimedia dissemination. We analyze these strategies theoretically and by simulations, and identify the one that is superior to the others.

Categories and Subject Descriptors

H.2.4 [**Database Management**]: Systems – *distributed databases, multimedia databases.*

General Terms

Algorithms, Performance, Design.

Keywords

Vehicular networks, multimedia, query processing, cellular communication, data dissemination, metadata.

1. INTRODUCTION

1.1 The contribution

A vehicular ad-hoc network (VANET) is a set of vehicles (mobile peers) that communicate with each other via short-range wireless technologies such as WiFi. An attractive application of VANETs is sharing of multimedia data such as songs and video clips (called reports in this paper) [3,7,8,9,14,15]. Existing studies explored two paradigms of sharing multimedia data in VANETs, namely push (data-to-query) and pull (query-to-data). In the push paradigm, multimedia reports are proactively disseminated [3]. In the pull paradigm, queries are proactively disseminated; multimedia reports are disseminated as responses to received queries [7,15]. VANETs have also been considered as an augment to the cellular communication [8,9]. In this case, the multimedia data sources reside on the fixed network. Some of the mobile peers download the multimedia reports via the cellular communication, and share the reports with the other peers via the short-range communication.

MoMM2010, 8-10 November, 2010, Paris, France.

* Research supported by NSF DGE-0549489 and IIS-0957394.

We refer to the VANETs in which the cellular communication is available as *hybrid vehicular networks*.

In this paper, we consider hybrid vehicular networks where multimedia reports are generated by the mobile peers, e.g., a 2-3 seconds video of the surroundings is generated by a peer every minute. In these networks, the data sources reside on mobile peers rather than in the fixed network. As in the network of shoppers in a mall, or passengers in an airport, or vehicles on the highway, a peer does not initially know the network-id's (i.e. cell-phone numbers) of the other peers in the network. However, a peer can communicate directly with other peers within its WiFi[1] transmission range without knowing their network-id. Furthermore, we do not require or assume a central server or any other form of directory/storage service in the fixed network.

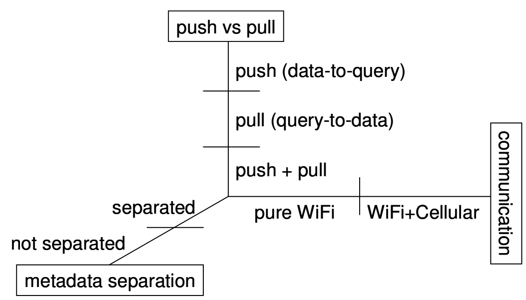

Figure 1.1. The design space for multimedia query processing in hybrid vehicular networks

An environment as described above renders a broad spectrum of possible query processing strategies, along the following three design dimensions. First, the peer-to-peer communication may use purely WiFi or it may use both WiFi and cellular communication (i.e., hybrid). Pure cellular communication is not an option because it requires knowledge of network id of the receiver, whereas a query originator initially only knows the description of the requested multimedia data but not the network-id of the peers containing this data. Second, as aforementioned, the query processing may adopt push or pull or the combination of the two. Third, observe that a multimedia report can be described by a much shorter *metadata* sub-report (e.g., time and location at which a multimedia clip was produced). The match (yes or no) between a query and a multimedia report can be determined solely based on the metadata sub-report. Due to size-differences, the metadata and multimedia sub-reports of a given report may be disseminated independently, and by different means (WiFi or cellular). The three design dimensions are illustrated by Figure 1.1.

[1] We use the term WiFi for simplicity, but most results of this paper apply to other short-range networking technologies such as DSRC.

In this paper we **conduct a comparative study of 13 query processing strategies** derived from a paradigm called **WiMaC** (WiFi-communication, Match, Communication). The WiMaC paradigm and the 13 strategies were introduced in a poster paper [6] based on the above three design dimensions. In the WiMac paradigm, since network id's are not known, query processing strategies start with a stage of WiFi dissemination to neighboring peers. The dissemination may be of the query, the multimedia reports, the metadata sub-reports, or some combination. When a match is found, it may be followed by a second stage of additional cellular or WiFi communication. In order to compare the WiMaC strategies we define the notion of <u>dominance</u> of one strategy by another. Then we prove analytically that **four strategies dominate the others** Finally, we **compare these four strategies by simulation** in a vehicular environment. It turns out that one is superior in terms of throughput. It is the one which disseminates queries and metadata sub-reports by WiFi; when a matching metadata and query meet at a peer, the corresponding multimedia report is transferred to the query originator by cellular communication.

1.2 Relevant Work

Multimedia sharing in mobile Peer-to-peer (P2P) networks. Many methods have been proposed for sharing multimedia data in mobile P2P networks, some using push [3] and the others using pull [7,15]. Most of these methods require that the network be connected so that multimedia reports/queries can be flooded or a communication path can be established. Our work does not require that the network be connected; we use an existing cooperative caching method (see [10]) to deal with disconnections. The method proposed in [15] (called V3) also uses cooperative caching as we do. V3 adopts pull and it uses only WiFi communication.

The authors in [17] combine short-range communication and cellular communication to facilitate query processing in mobile P2P networks. Their method requires a central server for directory service. That is, the matching between a query and a multimedia report is performed by the central server. The WiMaC paradigm, on the other hand, does not require any central server.

Combination of Push and Pull. It has been demonstrated that the combination of push and pull is superior to pure push and pure pull, in a static environment (see e.g., [16]). In this case, the dissemination of reports/queries follows a geometric structure, such as line segments or trees. Such structure based methods do not work in mobile P2P networks due to mobility and disconnections. In our work we adopt existing mobile P2P data dissemination techniques including smart-flooding [1] and cooperative caching [10] for the dissemination of queries and multimedia reports. These techniques are able to handle mobility and disconnections.

Data dissemination in mobile P2P networks. Data dissemination in mobile P2P networks has been studied before (see e.g., [2,4,14]). Our present work is orthogonal to these papers in the sense that the methods proposed in these papers can be used to improve WiFi communication in the WiMaC strategies. For example, in [2] it is found that dissemination using only vehicles moving in the opposite direction has higher performance than using vehicles moving in both directions; [4] studies the tradeoff between the delay and communication overhead by properly choosing the flooding scope; [14] proposes a popularity based replacement method for cooperative caching.

The rest of this paper is structured as follows: Section 2 introduces the model. Section 3 introduces the WiMac paradigm, the derived query processing strategies, and provides the theoretical dominance analysis. Section 4 compares the non-dominated strategies by simulations in a vehicular environment.

2. THE MODEL

2.1 The Environment

The environment is a *system* consisting of a set of *mobile peers*, or *peers*. This set of peers may change over time. Each peer (e.g. a cell phone) is equipped with the following capabilities: (i) producing multimedia data such as video, voice, or multimedia clips; (ii) short-range wireless communication such as WiFi; and (iii) infrastructure based communication such as 3G cellular. Via the infrastructure, a peer is able to transmit messages to another peer by MMS (Multimedia Message Service) or TCP/IP communication. This is referred to as the *cellular-channel* or *cellular* communication. Each peer has a *network-id* that is used as its address for cellular communication, and this id is required in order to send a message to the peer via the cellular channel. The network-id can be a cell-phone number or an IP address. In addition, peers can communicate via the WiFi channel if they are within transmission range. Knowledge of the network id is not necessary for this purpose.

2.2 Reports and Reports Databases

Each peer periodically produces multimedia reports. Formally, a *multimedia report* R, is a couple <*Meta(R), Media(R)*>, where *Meta(R)* and *Media(R)* are the *metadata* and *multimedia sub-*reports, respectively. The metadata sub-report contains attributes describing the multimedia report such as Time when the report was produced, the Location at which it was produced, the Network-id of the producing peer, etc. *Media(R)* is the multimedia content itself, e.g., the music or video file.

A peer also produces queries that are stored and disseminated in the form of reports called *query-reports*. A query requests both sub-reports of each satisfying multimedia report, but it refers only to the metadata of the multimedia report. Thus, whether or not there is a match between a query-report and a multimedia report can be determined solely based on the query and the metadata of the multimedia report. For example, a query requests a song by its title, thus the match between the query and the song can be determined solely based on the metadata of the song-report, but the query asks for both the metadata and the song itself to be returned.

A peer is called the *producer* of the query- and multimedia reports that it produces. Each query and each metadata sub-report contains the network-id of the producer of the report. However, the number of peers in the system and their network-id's are unknown by a peer. Each peer maintains a reports database that consists of three relations, namely a metadata-reports relation, a multimedia-reports relation, and a query-reports relation. These relations store the metadata sub-reports, multimedia reports, and query reports produced by the peer or received from other peers. Each tuple in the multimedia reports relation contains both, the metadata- and the multimedia sub-reports of a report, whereas a tuple in the metadata reports relation contains only the metadata of such a multimedia report. If a match between the metadata and a query is found, then the network-id in the metadata can be used to access the multimedia report. The reason for maintaining separate relations for metadata and multimedia reports is that metadata may be disseminated separately from its multimedia counterpart, as will be explained in the next subsection. To deal with the storage limit, reports relations are managed by a cooperative-caching method such as the one introduced in [10].

2.3 WiFi Communication

WiFi communication at a peer is implemented by a *WiFi Communication Module (WCM)* that is invoked by the query processing strategy executing at the peer. Each invocation provides the WiFi communication method with a set of reports to be transmitted. As a consequence, at any point in time the WCM stores a *transmission set*, i.e. reports to be transmitted. Each transmission sends a subset of the transmission set to all the current *neighbors,* i.e. all peers that are currently within transmission range.

Addition of a report to the transmission set is an *idempotent operation*, i.e. adding the report again after its first addition does not change anything. The WCM may service multiple applications, and thus the WCM transmits the reports in the transmission set with a frequency that will optimize communication for all the applications. Furthermore, due to bandwidth limitations a transmission may not be able to send all the reports in the transmission set, thus the reports are prioritized by the WCM. The transmission frequency and priority are outside the control of the query processing strategy, and are executed by the WCM such that the WiFi communication efficiency is maximized. A possible prioritization scheme is introduced in [10].

Due to WiFi communication errors and limited bandwidth, it is more likely that long reports (i.e. multimedia reports) get lost or delayed than short ones (i.e. queries and metadata sub-reports). Thus a short report may propagate differently from a long one, even if initially they are both broadcast simultaneously from the same peer.

3. The WiMaC Query Processing Strategies

As mentioned above, some or all reports that satisfy a query *Q* may reside on peers that are different than the query producer, *Qp*. Since *Qp* does not normally have the network id of such peers, and does not even know how many reports satisfy the query, all query processing strategies start with a stage of WiFi dissemination to neighboring peers. The dissemination may be of the query, the multimedia reports, the metadata sub-reports, or some combination. When a match is found, it may be followed by a second stage of additional cellular or WiFi communication. For example, assume that the match is between a query and a metadata report, and that the multimedia sub-report is located at another peer. Then the multimedia sub-report has to be transferred to the query producer by additional communication. Thus, this is the (WiFi-communication, Match, Communication) paradigm, called WiMaC, and all query processing strategies discussed in this paper are special cases of WiMaC.

The query processing strategy operating at a peer is not concerned with communication issues such as the location, distance, or direction of the destination peer of a report. If necessary, these are handled by the communication layer.

This section discusses the WiMac strategies, and is organized as follows. In §3.1 we present the structure of the design space. In §3.2 we describe 13 WiMaC query processing strategies. In §3.3 we prove that four strategies dominate the others.

3.1 Strategies Design-Space: General Guidelines

The structure of the design space is depicted in Table 3.1, and is explained as follows. As aforementioned, the first stage of the WiMaC paradigm is a WiFi dissemination which finds a match. The WiFi dissemination is possibly followed by a second stage to complete query processing. There are seven design choices for the first stage, depending on whether the query, the multimedia reports, the metadata sub-reports, or some combination thereof are disseminated (see left column of Table 3.1 which indicates the disseminated combination).

Table 3.1. Design space of the WiMaC paradigm
Notation: media = multimedia, meta = metadata sub-report, Q = query, cell = cellular. Strategy Names: 1 is (media), 2a is (meta)-WiFi, 3b is (Q)-cell, etc.

No.	Type of reports disseminated in the first stage (always via WiFi)	Communication medium in the second stage
1	(media)	No second stage
2	(meta)	WiFi (2a), cell (2b)
3	(Q)	WiFi (3a), cell (3b)
4	(media, meta)	WiFi (4a), cell (4b)
5	(media, Q)	WiFi (5a), cell (5b)
6	(meta, Q)	WiFi (6a), cell (6b)
7	(media, meta, Q)	WiFi (7a), cell (7b)

In the (media, meta) and (media, meta, Q) choices, multimedia reports and metadata sub-reports are disseminated independently. Observe that a multimedia report contains its metadata sub-report, and the reason for disseminating the metadata sub-report alone in addition to the multimedia report is as follows. As discussed in §2.3, multimedia reports and metadata sub-reports propagate differently by WiFi; the metadata report propagates faster, thus possibly meets more queries within a given time period.

The design choices for the second stage depend on whether the second stage follows at all, and if so whether WiFi or cellular communication is used for this stage (see right column of Table 3.1). Observe that when a query and matching metadata report meet at a peer, say V, then the network id's of both the query producer and the report producer are known to V since they are in the query and metadata report, respectively. Thus the second stage can be conducted via the cellular infrastructure.

Observe that if only multimedia reports are disseminated at the first stage (i.e., the (media) choice), then there is no need for the second stage. The reason is that in this case the query is not disseminated thus the match must have been found at the query producer, with the multimedia report constituting the answer.

Even if the strategy has a second stage, it does not mean that the second stage always executes after a match at a peer. For example, for the (media, meta) case both multimedia and metadata reports are disseminated in the first stage of WiMaC. If a match involving a metadata report occurs, then a second stage is necessary to get the actual multimedia report. But if the match involves a multimedia report, then the second stage is not necessary because the multimedia report is already at the query producer. Similarly, for the (media, Q), and (media, meta, Q) cases, a second stage may not occur after a match at a peer.

In the following subsection we present the 13 WiMaC strategies. Each strategy is denoted as follows. The denotation consists of the strategy number as defined in Table 3.1 and the strategy name. The strategy name is formed as follows. If there is not the second stage, then the strategy is named by the first stage, i.e., (media). If there is a second stage, then the strategy is named by the two stages connected by a "-". For example, 2b (meta)-cell denotes the 2b strategy which disseminates metadata sub-reports in the first stage and uses cellular communication in the second.

3.2 The WiMaC Strategies Description

In §3.2.1 we present the strategies that use WiFi communication in both stages, and in §3.2.2 we present the strategies that use cellular communication in the second stage.

3.2.1 WiFi-only Strategies

There are seven WiFi-only strategies, including 1(media), 2a (meta)-WiFi, 3a (Q)-WiFi, 4a (media,meta)-WiFi, 5a (media,Q)-WiFi, 6a (meta,Q)-WiFi, and 7a (media,meta,Q)-WiFi.

1 (media): In this strategy, in the first stage of WiMaC multimedia reports are disseminated via WiFi (see Figure 3.1). In other words, multimedia reports are simply pushed by WiFi. Queries are kept at the producer peer, and a match occurs when a disseminated multimedia report arrives at a matching query. There is no second stage.

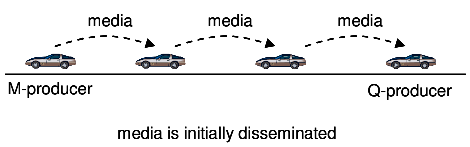

Figure 3.1. Illustration of the 1 (media) strategy.

2a (meta)-WiFi: In this strategy, in the first stage of the WiMaC paradigm metadata sub-reports are disseminated via WiFi. When a metadata sub-report M reaches a matching query Q producer, the Q-producer disseminates Q via WiFi. When the M(edia)-producer receives Q, the M-producer disseminates the multimedia report via WiFi, to reach the Q-producer and provide an answer to Q.

3a (Q)-WiFi: In this strategy, in the first stage of WiMaC queries are disseminated via WiFi (See Figure 3.2). When a query Q reaches the producer peer of a matching multimedia report, the M-producer disseminates the multimedia report via WiFi to reach the Q-producer. As stated in section 2.3, whether the dissemination is broadcast or directional towards the Q-producer is handled by the communication layer.

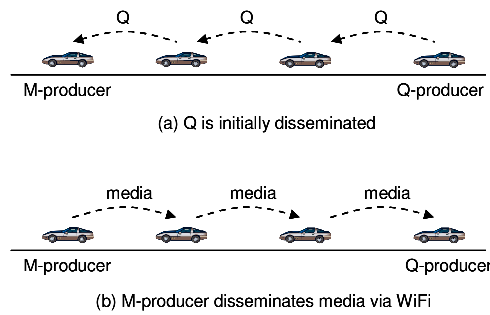

Figure 3.2. Illustration of the 3a (Q)-WiFi strategy.

4a (media,meta)-WiFi: In this strategy, in the first stage of WiMaC metadata and multimedia reports are disseminated separately via WiFi. If the producer of a matching query Q receives a multimedia report, then there is no second stage. If the producer of a matching query Q receives a metadata sub-report, the Q-producer disseminates Q via WiFi. When a peer Z that has a matching multimedia report M receives Q, Z disseminates M via WiFi to reach the Q-producer.

5a (media,Q)-WiFi: In this strategy, in the first stage of WiMaC multimedia and query reports are disseminated via WiFi. When a multimedia report and a matching query Q collocate at a peer Z, Z disseminates the multimedia report via WiFi to reach the query producer.

6a (meta,Q)-WiFi: In this strategy metadata sub-reports and queries are disseminated via WiFi. When a metadata sub-report M and a matching query Q collocate at a peer Z, Z disseminates Q via WiFi. When the M-producer receives Q, the M-producer disseminates the corresponding multimedia report via WiFi to reach the Q-producer.

7a (media,meta,Q)-WiFi: This strategy is a combination of (media,meta)-WiFi and (meta,Q)-WiFi in the sense that it does everything that (media,meta)-WiFi does and also everything that (meta,Q)-WiFi does.

3.2.2 WiFi-cellular Strategies

In the WiFi-cellular strategies, after a match is discovered, the answer M is communicated from a peer P to the query producer via the cellular channel. However, P first inquires via the cellular channel whether the producer has already received M (from other peers); if so, the transmission of M is suppressed.

There are six WiFi-cellular strategies, including 2b (meta)-cell, 3b (Q)-cell, 4b (media,meta)-cell, 5b (media,Q)-cell, 6b (meta,Q)-cell, and 7b (media,meta,Q)-cell.

2b (meta)-cell: In this strategy, in the first stage of WiMaC metadata sub-reports are disseminated via WiFi. When a metadata sub-report M reaches a matching query Q producer, the Q-producer peer sends Q to the M(edia)-producer via the cellular channel. In response, the M-producer sends M and all the other matching multimedia reports that it has to the Q-producer, via the cellular channel.

3b (Q)-cell: In this strategy, in the first stage of WiMaC queries are disseminated via WiFi. When a query Q reaches a matching multimedia report B, the M-producer sends M to the Q-producer via the cellular channel.

4b (media,meta)-cell: This strategy is identical to (meta)-cell (2b), with the following addition. Multimedia reports are disseminated as well in the first stage of WiMaC. There is no second stage if the Q-producer receives a multimedia report from the WiFi dissemination.

5b (media,Q)-cell: In this strategy multimedia and query reports are disseminated via WiFi in the first stage. When a multimedia report M and a matching query Q collocate at a peer Z, Z sends M to the Q-producer via the cellular channel.

6b (meta,Q)-cell: In this strategy metadata and query reports are disseminated via WiFi in the first stage (see Figure 3.3). When a metadata sub-report M and matching query Q collocate at a peer Z, Z sends Q to the M-producer via the cellular channel. In response, the M-producer sends the multimedia report M and all the other matching multimedia reports that it has to the Q-producer, via the cellular channel.

7b (media,meta,Q)-cell: This strategy is a combination of 5b (media,Q)-cell and 6b (meta,Q)-cell in the sense that it does everything that (media,Q)-cell does, and also everything that (meta,Q)-cell does.

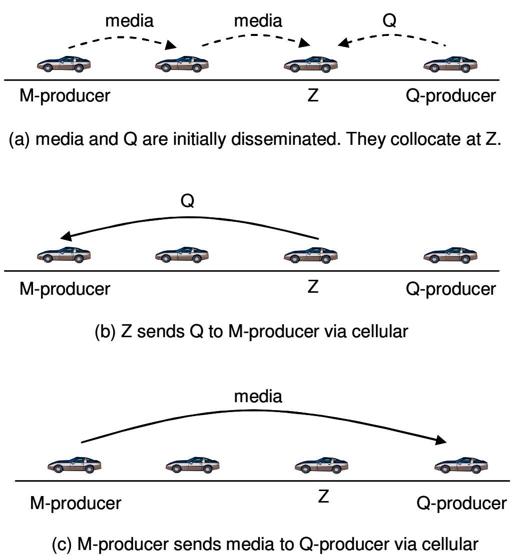

Figure 3.3. Illustration of the 6b (meta,Q)-cell strategy.

3.3 Strategy Dominance Analysis

3.3.1 Definition of Dominance

Let a peer receive an answer multimedia report at time t. The *response-time* of the answer is the length of the time period since the answer is produced until t.

We say that a strategy X is *dominated* by another strategy Y if the following 4 conditions are satisfied for every multimedia report M:

(1) For every query that M answers, if the answer is received in Y, it is also received in X;
(2) For every query that M answers, its response-time in Y is no higher than that in X.
(3) The WiFi communication cost of M in Y is not higher than that in X.
(4) The cellular communication cost of M in Y is not higher than that in X.

Intuitively, if X is dominated by Y, then the performance and the efficiency of X are no better than those of Y and therefore X is not worth further studying. In this subsection we identify the strategies that are dominated.

In the dominance analysis, the communication cost (but not the delay) of query-reports and metadata sub-reports is ignored for WiFi communication. Similarly, the communication cost of these reports is ignored for cellular communication. This is because query-reports and metadata sub-reports are very short. However, the simulations take into account the communication overhead of the query-reports and metadata sub-reports (see §4.1).

We say that strategy X is *weakly dominated* by strategy Y if the above dominance relationship only satisfies conditions 1-3, i.e. the cellular communication cost of Y may be higher. Weak dominance is appropriate for unlimited data plans offered by some cellular service providers.

3.3.2 Dominated Strategies

The dominance relationship is described by the following theorems.

Theorem 1: The strategies 4a (media,meta)-WiFi, 5a (media,Q)-WiFi, and 7a (media,meta,Q)-WiFi are dominated by strategy 1 (media).

Proof sketch: Compared to the 1 (media) strategy, 4a, 5a, and 7a disseminate more information via WiFi. Consider the implication of this extra data dissemination for a multimedia report M that answers a query Q. Since none of the strategies use the cellular channel, the implication is that a peer Z that has Q and M adds the answer M to the WiFi transmission set (see transmission set definition in §2.3). However, since Z has B, it would have done so anyway under strategy 1. Furthermore, since addition of a report to the transmission set is idempotent, even if the extra information disseminated in the first stage of WiMaC triggers more additions of M to the transmission set, these do not affect whether M is received by the Q-producer, and if so when. QED.

Theorem 2: The strategies 2a (meta)-WiFi and 6a (meta,Q)-WiFi are dominated by strategy 3a (Q)-WiFi.

The proof of Theorem 2 follows the same principle as that of Theorem 1 and is omitted due to space limitations.

Theorem 3: The strategies 3b (Q)-cell and 2b (meta)-cell are dominated by strategy 6b (meta,Q)-cell.

Proof sketch: We first prove that 3b (Q)-cell is dominated by 6b (meta,Q)-cell. In (Q)-cell, only queries are disseminated in the first stage of WiMaC, and a match occurs when the query Q reaches the producer of a matching multimedia report B. In (meta,Q)-cell, metadata sub-reports and query-reports are disseminated simultaneously, and thus the match between B and Q may be discovered at an intermediate peer at which B and Q co-locate. Thus matches are discovered no later in (meta,Q)-cell than in (Q)-cell. That is, a query-report meets a matching metadata sub-report no later in (meta,Q)-cell than in (Q)-cell. On the other hand, the communication overhead of (meta,Q)-cell is no higher than that of

(Q)-cell because in both strategies the matching multimedia reports are transferred via the cellular channel, and, as assumed in sec. 3.3.1, the communication overhead of metadata sub-reports is negligible. The proof for 2b (meta)-cell being dominated by 6b follows the same principle. QED.

Theorem 4: The strategies 5b (media,Q)-cell and 4b (media,meta)-cell are weakly dominated by strategy 7b (media,meta,Q)-cell.

The proof of Theorem 4 follows the same principle as that of Theorem 3 and is omitted due to space limitations. Observe that 7b only weakly dominates 5b and 4b because in 7b a match may be discovered earlier (i.e. when metadata and matching query co-locate), and the corresponding multimedia report B will be communicated by the cellular channel; whereas in 5b and 4b, B may reach the query producer later via the WiFi channel, avoiding the cellular communication of B.

Strategies 1 and 3a are incomparable because 3a disseminates only multimedia reports that answer queries, whereas 1 disseminates all multimedia reports thus its communication cost is higher; on the other hand, since 1 disseminates all multimedia reports as soon as they are produced, its response time is lower. Similarly, 7b and 6b are incomparable because the WiFi communication cost of 7b is higher, but its response time may be lower.

Observe that each dominated strategy is dominated by a strategy from the same group and thus is not worth further studying. Thus the next section focuses on the non-dominated strategies.

4. Comparison of Non-dominated Strategies by Simulations

In this section we compare by simulation the four non-dominated query processing strategies, namely: 1 (media), 3a (Q)-WiFi, 7b (media,meta,Q)-cell, and 6b (meta,Q)-cell (see Table 3.1). The comparisons are based on the application of delivering traffic multimedia clips among moving vehicles to warn drivers about traffic jams and dangers. §4.1 elaborates the application scenario. §4.2 describes the simulation environment. §4.3 introduces the performance measures. §4.4 presents the results.

4.1 Multimedia Traffic Information Application

In this application, each vehicle periodically captures short (2-seconds) multimedia clips consisting of audio and/or video of the traffic conditions surrounding it. We have implemented an experimental system via which automatically captured video clips from dashboard-mounted smart-phones are disseminated among vehicles using a WiMac strategy. A sample video clip can be viewed at [13]. The sample is 2 seconds long with the size of 65K bytes. The traffic conditions can be readily discerned by the viewer (the receiving driver) of the video. Each multimedia clip is encapsulated in a multimedia report R <Media(R), Meta(R)>. Meta(R), the metadata sub-report, is a 3-element tuple <producer, produce-time, produce-location>, where *producer* is the network-id of the vehicle that produced R; *produce-time* is the time when R starts to be captured; and *produce-location* is the (x,y) location where R starts to be captured. Media(R), the multimedia sub-report, is the multimedia clip encapsulated in R. Based on experiments with a smart-phone video camera, its size in the simulations is taken to be 65K bytes.

A query report is a 4-element tuple <producer, time, location, target-region>, where *producer* is the network-id of the query producer; *time* is the time at which the query is issued; *location* is the the location at which the query is issued; and *target-region* indicates that the query producer is interested in receiving multimedia clips that started to be captured in this region.

Each multimedia report and each query report has a *lifetime* which defines the period of time during which it is of interest. In the simulations all the reports have the same lifetime which is a system parameter. A report is dropped by the vehicles when its lifetime expires. A multimedia report R, or its metadata sub-report, *satisfies* a query report Q if (i) R is produced after the produce-time of Q; (ii) R.location falls within Q.target-region.

4.2 Simulation Environment

The simulated area is a part of the Chicago road network (see Figure 4.1). We used SWANS++ [5] as the simulation tool. We used smart-flooding [1] and cooperative caching [10] for reports dissemination via WiFi. Since these are existing techniques we do not elaborate on them further. The WiFi bandwidth is 2 Mbps, following the setup used in [3,8]. We augmented SWANS++ with cellular communication based on the typical parameters of a 3G cellular network (see [11, 12]).

In order to assure the robustness of the conclusions drawn from the SWANS++ communication model, we conducted simulations with the WiFi communication component of SWANS++ replaced by a simple model developed by our own. It turns out that our model leads to the same conclusions drawn from the SWANS++ model. Details are omitted due to space limitations.

Table 4.1 lists all the simulation parameters. We tested two traffic scenarios, namely light-congestion (4000 vehicles), and heavy-congestion (8000 vehicles with road construction on 50% of randomly selected road segments).

For each traffic scenario, only a fraction of the entire vehicle population generates multimedia clips and participates in the WiMac query processing. This fraction is called the *penetration ratio*. By varying the penetration ratio we varied the density of the WiFi network. The mapping between the penetration ratio and the average inter-vehicle distance is given in Table 4.2.

The query target region is at distance 1600 meters, and has a width of 500 meters. This means that a vehicle is interested in traffic multimedia clips captured in the area lying between 1350 meters and 1850 meters away from the current query producer's location.

Every 10 seconds, each vehicle produces a multimedia report with a probability that is in reverse proportion to the penetration ratio. It reflects the realistic fact that not all participating (in WiMac) vehicles are producing reports. Thus, in our simulations the density of participating peers varies, but the supply of multimedia reports is fixed. Particularly, the supply of multimedia reports is fixed at 4 per second. Queries are produced as follows. Every 300 seconds, each participating vehicle produces a query with a probability called the *query ratio*. The query ratio is a system parameter. Thus, for example, if the query ratio is 25%, then when the penetration ratio is 4%, every 300 seconds 1% of the vehicles produce queries.

Table 4.1. Simulation parameters and their values

Parameter	Values
Total road length, simulated area.	96 km, 24×31 sq. km
Traffic condition	Light-congestion: 4000 vehicles, 64km/hour avg speed over time among all road segments
	Heavy-congestion: 8000 vehicles, reduced speed-limit for 50% of road segments, such that avg speed over time among all road segments is 25km/hour
Penetration ratio (P_{ratio})	1% ~ 50%
WiFi transmission range, data transmission rate of WiFi channel	250 meters, 2 Mbps
Side-length, capacity of each cell	2.5 km, 30 users
Cellular data transmission rate	384 Kbps
Frequencies of query generation	Every 300 seconds with randomization
Query ratio	0.25, 0.5, 0.75, 1
Frequency of multimedia report generation	Every 10 seconds with randomization
Query distance, query width	1600 meters, 500 meters.
Query/clip report lifetime	60,120,180,240,300 seconds
Sizes of query, metadata, and multimedia-clip	40bytes, 28bytes, and 65Kbytes, respectively.
Reports database size	6 Mbytes
Length of a simulation run	3600 simulated seconds

Table 4.2. Mapping between the penetration ratio and the average distance between two neighboring participating vehicles. (distance unit: meter)

	1%	5%	12.5%	25%	37.5%	50%
Light congestion	2400	480	192	96	64	48
Heavy congestion	1200	240	96	48	32	24

4.3 Performance Measures

(Response-time bounded) answer throughput: The answer throughput is the number of distinct answers (i.e., matching multimedia reports) received for each query. An answer is counted towards the throughput only if both the answer and the query have not expired at the time when the answer is received. Thus the response-time is taken into account in the answer throughput measure.

Communication overhead: The average number of bytes per vehicle submitted to the MAC level during the simulation, by the WiFi channel and the cellular channel, respectively. In other words, this overhead is the amount of attempted communication, the amount communicated successfully is lower.

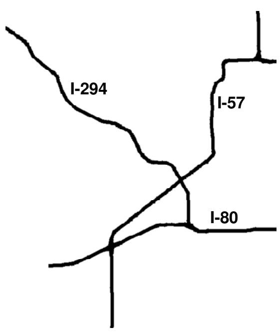

Figure 4.1. Simulated road network: portion of the highway system in a US city. Network size: 24×31 km^2, total road length: 96 km.

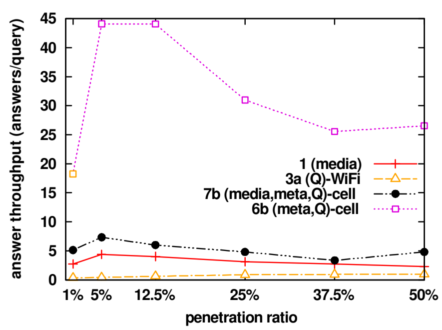

Figure 4.2. Answer throughput versus penetration ratio, light-congestion, report lifetime = 300 seconds, query ratio = 0.25.

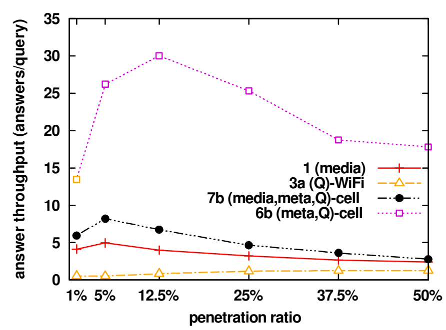

Figure 4.3. Answer throughput versus penetration ratio, heavy-congestion, report lifetime = 300 seconds, query ratio = 0.25.

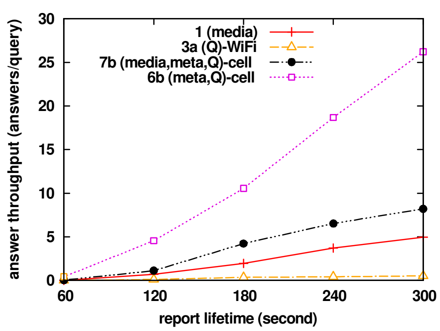

Figure 4.4. Answer throughput versus report lifetime, heavy-congestion, SWANS++, penetration ratio = 0.05, query ratio = 0.25.

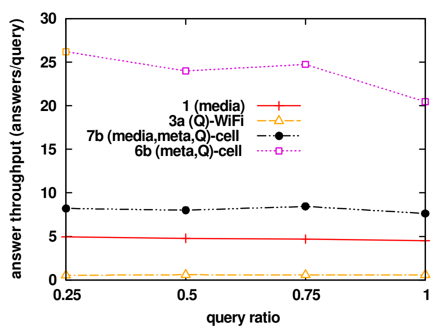

Figure 4.5. Answer throughput versus query ratio, heavy-congestion, penetration ratio = 0.05, report lifetime = 300 seconds.

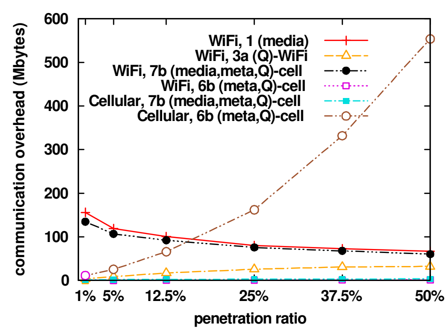

Figure 4.6. Communication overhead versus penetration ratio, heavy-congestion, SWANS++, penetration ratio = 25%, report lifetime = 300 seconds, query ratio = 0.25.

4.4 Simulation Results

In §4.4.1 we present the results with regard to the answer throughput measure. In §4.4.2 we present the results with regard to the communication overhead measure.

4.4.1 Answer throughput

Figures 4.2 and 4.3 show the answer throughput as a function of the penetration ratio, for the light-congestion and heavy-congestion scenarios, respectively. Figure 4.4 shows the answer throughput as a function of the report lifetime for the heavy-congestion scenario. Figure 4.5 shows the answer throughput as a function of the query ratio for the heavy-congestion scenario. The figures for the report lifetime and the query ratio for the light-congestion scenario are omitted due to space limitations. It can be seen that in all the figures, ranking of the strategies based on throughput is 6b (meta,Q)-cell > 7b (media,meta,Q)-cell > 1 (media) > 3a (Q)-WiFi.

Best strategy. Strategy 6b (meta,Q)-cell is the clear winner. The advantage of 6b increases as the penetration ratio increases. In some cases, the answer throughput of 6b is seven times higher than those of the other strategies. It is surprising that strategy 7b (media,meta,Q)-cell is much worse than 6b. Compared to 6b, strategy 7b also disseminates multimedia reports via WiFi in the first WiMaC stage, thus vehicles have a chance to receive answers from the WiFi dissemination directly. Presumably the performance of 7b should be close to if not higher than that 6b. The poor performance of 7b is probably due to the fact that the WiFi dissemination of multimedia reports occupies a lot of WiFi bandwidth, which creates contention and collisions in the dissemination of metadata sub-reports and query reports. This interference significantly slows down the discovery of matches.

Comparison of WiFi-only strategies. 1 (media) is better than 3a (Q)-WiFi. 1 and 3a represent two paradigms of query processing, i.e., 1 represents push (data-to-query) and 3a represents pull (query-to-data). The simulation results show that push is better than pull for the considered environment. Intuitively, the pull strategy requires a round-trip dissemination in order for a query originator to receive an answer: the query has to travel from the query originator to the answer producer and then the answer has to travel back from the answer producer to the query originator. If either way does not go through or experiences a long delay, then the answer does not reach the query originator within the lifetime; and this scenario is likely in a highly mobile environment.

4.4.2 Communication Overhead

Figure 4.6 shows the communication overhead as a function of the penetration ratio for the heavy-congestion scenario. The throughput for the same configuration is presented in Figure 4.3. It can be seen that the winning strategy in terms of throughput, 6b (meta,Q)-cell, has the lowest WiFi communication volume, because only metadata reports and query reports (which are short) are disseminated via WiFi. 6b has the highest cellular communication volume. When the penetration ratio is low (12.5% or below), the communication overhead of 6b (including WiFi and cellular) is lower than those of the other three strategies. When the penetration ratio is high (25% or above), the communication overhead of 6b is higher than those of 1 (media) and 7b (media,meta,Q)-cell. In this case, the 6-fold increase in throughput comes with a higher communication cost.

5. Conclusion

We compared WiMac strategies introduced in [6] for querying multimedia reports in VANETs, where cellular communication is also available. We proved analytically that 4 of the 13 WiMac strategies dominate the others. Intuitively, strategy A dominates strategy B when each query returns in A a superset of the set of answers it returns in B, each with a response time that is not higher in A than in B; additionally, the communication cost of A is not higher than that of B.

Finally, we compared the four non-dominated strategies by simulations in a transportation application. The simulations revealed that 6b (meta,Q)-cell has by far a higher throughput than the other three (up to 6-fold). It also often approaches an ideal benchmark that despite of the challenging environment delivers all the existing answers to each query. However, the communication cost of strategy 6b is also higher than that of the others. Intuitively, strategy 6b operates as follows. It separates metadata dissemination from its multimedia report, it combines push of metadata and pull by queries, and uses the cellular infrastructure to communicate multimedia reports.

References

[1] T. Osafune, L. Lin, and M. Lenardi. Multi-hop vehicular broadcast (MHVB). ITST, 2006.

[2] T. Nadeem, P. Shankar, L. Iftode. A Comparative Study of Data Dissemination Models for VANETs. *MobiQuitous*, 2006.

[3] V. Gau, C. Huang, J. Hwang. Reliable Multimedia Broadcasting over Dense Wireless Ad-Hoc Networks. Journal of Communications, 4(9):614-627, 2009.

[4] U. Lee, E. Magistretti, M. Gerla, P. Bellavista, and A. Corradi. Dissemination and Harvesting of Urban Data Using Vehicular Sensing Platforms. *IEEE Transactions on Vehicular Technology*, (58)2, 2009, pp. 882-901.

[5] http://www.aqualab.cs.northwestern.edu/projects/swans++/

[6] O. Wolfson, B. Xu, H. Cho. Multimedia Traffic Information in Vehicular Networks (poster paper). *ACM GIS* 2009, pp. 480-483.

[7] H. Wu, H. Peng, Q. Zhou, M. Yang, B. Sun, and B. Yu. P2P Multimedia Sharing over MANET. *MMM* 2007.

[8] S. Al-Chikhani, L. Al-Kanj, and Zaher Dawy. Video Distribution over Wireless Networks with Mobile-to-Mobile Cooperation. ACTEA 2009.

[9] M. Leung and S.-H. Chan. Broadcast-Based Peer-to-Peer Collaborative Video Streaming Among Mobiles. IEEE Trans. on Broadcasting, (53)1: 350-361, 2007.

[10] T. Zhong, B. Xu, P. Szczurek, and O. Wolfson. Trafficinfo: An Algorithm for VANET Dissemination of Real-Time Traffic Information. *World Congress on Intelligent Transport Systems*, 2008.

[11] http://wireless.agilent.com/rfcomms/refdocs/wcdma/e6703t_user_documentation.php

[12] P. A. Hosein. Capacity of packetized voice services over time-shared wireless packet data channels. *INFOCOM*, 2005.

[13] http://www.cs.uic.edu/~boxu/video_clips/video_13126071744_11-21-24.asf

[14] Y. Zhang, J. Zhao, and G. Cao. Roadcast: A Popularity Aware Content Sharing Scheme in VANETs. *ICDCS*, 2009.

[15] M. Guo, M. Ammar, E. Zegura. V3: A vehicle-to-vehicle live video streaming architecture. *PerCom*, 2005.

[16] X. Liu, Q. Huang, Y. Zhang. Balancing Push and Pull for Efficient Information Discovery in Large-Scale Sensor Networks. *IEEE Trans on Mobile Computing*, 6(3): 241-251, 2007.

[17] S. Ghandeharizadeh, S. Kapadia, and B. Krishnamachari. An evaluation of availability latency in carrier-based vehicular ad-hoc networks. *MOBIDE'06*.