Sensors on wheels – towards a zero-infrastructure solution for intelligent transportation systems

Samir Goel[1], Tomasz Imielinski[1], Kaan Ozbay[2], Badri Nath[1]
{gsamir@cs, imielins@cs, kaan@rci, badri@cs}.rutgers.edu

[1]Department of Computer Science
Rutgers, The State University of NJ
Piscataway, NJ 08854
Tel: (732)-445-2706
Fax: (732)-445-6722

[2]Department of Civil Engineering
Rutgers, The State University of NJ
Piscataway, NJ 08854
Tel: (732)-445-2792
Fax: (732)-445-0577

November 15, 2003

Word Count: 6600

ABSTRACT

Traffic congestion has become a part of daily life of most of us. One possible way of preventing congestion from building up is by disseminating traffic information. Today a number of commercial solutions exist for disseminating traffic information (e.g., Traffic.com, Metrocommute, EtakTraffic). However, these solutions are plagued by prohibitive deployment and maintenance cost that prevents widespread deployment. As a result, these solutions cover select highways while leaving out major fractions of the highways. In this paper we present our vision of a low cost, highly scalable zero-infrastructure system for collecting and disseminating information about traffic conditions. Our goal is to devise a system that can cover the entire road network of the US. We describe the proposed architecture for the system, and evaluate it using a micro-traffic simulator and a calibrated model of the South Jersey transportation network.

1 INTRODUCTION

"Congestion has become part of daily life in many places as traffic continues to increase on a relatively unchanging highway network. Highway congestion is not just a problem of recurring "rush hour" delay in major cities. More than half of all congestion is non-recurring, caused by crashes, disabled vehicles, adverse weather, work zones, special events and other temporary disruptions to the highway transportation system. Unless we manage highway congestion, our nation will continue to incur economic costs in forgone productivity, wasted fuel, and a reduced quality of life." (9) According to the Texas Transportation Institute (TTI), the total congestion "bill" for 75 urban areas in the United States in 2000 came to $67.5 billion, which was the value of 3.6 billion hours of delay and 5.7 billion gallons of excess fuel consumed (24). One possible way of controlling the extent of congestion is by disseminating traffic information. This gives an opportunity to the drivers to bypass congested routes thus preventing congestion from building up. Today a number of commercial systems exist for collecting and disseminating traffic information (e.g., Traffic.com (10), Metrocommute (6), EtakTraffic (5)). However, these systems tend to cover select highways while leaving out a major fraction of roadways, thereby creating a *"digital divide"*. In case of congestion, if any of alternate routes falls outside the coverage zone, drivers are left to guesswork, past experience, and their instincts in deciding on a faster route to their destination. The main factor that prevents these systems from covering the entire road network of the US is the cost involved. Each of these systems requires an infrastructure to be deployed (e.g., helicopters; static cameras at busy intersections and entrances to tunnels/bridges; traffic flow sensors along major highways). This represents a huge amount of money in one-time deployment cost, and a significant recurring cost in maintenance. In order to address the problem of traffic congestion, there is an urgent need for a solution that can be deployed on a large scale at a low cost.

Recent parallel projects (25, 26, 16, 22) have also proposed using vehicles as sensors of traffic information. In (16) authors propose using AVL-equipped transit vehicles as probe vehicles for estimating traffic information. In TRANSMIT (22) authors use Electronic Toll and Traffic Management (ETTM) equipment, compatible with E-ZPass System (an electronic toll collection system, currently in operation in NY/NJ/CT metropolitan areas) for traffic surveillance and incident detection. These projects, and the work by Brackstone et al (12), Chen et al (13), and Elango et al (18) demonstrate that it is feasible to use vehicles as traffic information sensors. Our work can be seen as extending these in two important ways: firstly, we attempt to use potentially **every vehicle** as a traffic sensor. We expect this to result in wider and better coverage (section 2). Secondly, we attempt to do away with the requirement of any centralized system and infrastructure (e.g., toll booths).

In (25, 26) authors propose equipping vehicles with short-range wireless link (e.g., IEEE 802.11 (11)) for spreading traffic information. Wischhof et al (25) suggest having each vehicle periodically transmit its position, velocity, and heading, along with aggregate traffic information (aggregated traffic information is computed based on periodic reports received by the vehicle). Ziliaskopoulos et al (26) suggest having vehicles exchange recent travel time information with vehicles traveling in the opposite direction.

As discussed in section 2, in order for the traffic information to be useful, drivers need to know it well in advance (a few miles, in some cases) to be able to avoid congested routes. With short-range wireless link, propagating traffic information a few miles would require many vehicle-to-vehicle communication hops, and would be feasible only in regions with sufficiently high vehicle density. This may not be true for all the regions (rural regions have low vehicle density) and at all times (vehicle density decreases substantially in the night). The proposed solutions cannot be carried over to wide-area wireless link (e.g., cellular network) where bandwidth is precious and expensive. In our work, we propose using wide-area wireless link and our emphasis is on controlling the bandwidth consumption while retaining requisite level of awareness of traffic conditions among vehicles.

The main goal of this paper is to examine various mechanisms for efficient dissemination of traffic information, and to access their effectiveness using well defined performance metrics. We use Paramics (8), a state-of-the-art micro-traffic simulator, along with a calibrated model of Southern New Jersey transportation network for this purpose. We identify the research issues that need to be resolved before a practical system can be built. The main contribution of this paper is that it tests various wireless communication concepts using well calibrated microscopic simulation. This is clearly an important step towards real life deployment of a large scale system where considerable number of vehicles participate, not just probe vehicles as in ADVANCE (3) and other related projects.

In the next section, we give an overview of the system. In section 3, we describe the proposed architecture for the system. In section 4 we enumerate the performance metrics. In section 7 we detail the mechanisms involved. In section 8, we evaluate the proposed mechanisms. In section 9 we outline our future work.

2 SYSTEM OVERVIEW

The design space for systems that can collect and disseminate traffic information is large. We envision a highly scalable system that exploit existing infrastructure and require minimal or zero additional infrastructure. Systems based on such designs will arguably have low deployment and maintenance cost. We believe such systems have the best chance of being deployed on a large scale. We envision that every participating vehicle would be equipped with a device that we call *TrafficRep*. This device is responsible for collecting and disseminating traffic information (we describe it in more detail in section 2.1). The TrafficRep device connects to the in-vehicle navigation system (e.g., 750NAV from Magellan (1)), supplying it with current traffic conditions. From a user's (driver) perspective, TrafficRep device appears as a black box. A user only interacts with the in-vehicle navigation system posing queries like, *"What is the fastest route from Busch Campus to Newark Airport?"* The in-vehicle navigation system, in turn, queries the TrafficRep device to obtain current traffic information on various road segments, computes the fastest route, and displays it to the user. Figure 1 shows the schematic diagram for the system. In response to a query, the TrafficRep device may either respond with the information that it already has, or it may pose the query to TrafficRep devices located in other vehicles. The driver has no direct control over this decision. This prevents a user from intentionally/unintentionally flooding the wireless channel with queries.

Assuming that "enough" vehicles in an area have TrafficRep device on them, these vehicles collectively store to-the-minute traffic information for the area. However traffic information of interest to a vehicle is potentially distributed across a number of highly mobile vehicles. In order to achieve the goal of traffic awareness, the system needs to support three basic operations in order to get the data from vehicles having desired traffic information to vehicles requesting traffic information:

- *Querying:* This operation allows a vehicle to query the traffic conditions on a road segment. Note that vehicles would typically issue a query for traffic conditions on a stretch of route. Such a query can be thought of as issuing multiple queries, one for each road segment on the queried route.
- *Monitoring:* This operation allows a vehicle to monitor the traffic conditions on a road segment. Whenever the travel time on a road segment changes appreciably, the vehicle receives an update.
- *Dissemination:* This operation entails broadcasting/multicasting traffic information in a specified ge-

ographic area. For frequently requested traffic information, it may be more bandwidth efficient to disseminate the information proactively.

The important features of this system are:

1. The system consists of millions of highly mobile sensor nodes (TrafficRep devices on vehicles). This has two consequences:
    - The coverage provided by the vehicles is dynamic and changes with the location of vehicles.
    - Multiple vehicles may sense the travel time of each road segment at around the same time, introducing redundancy in information collected by the vehicles. This redundancy is critical to the resilience and reliability of the system.
2. Turning vehicles into traffic sensors has the advantage that zero additional infrastructure is required. However the disadvantage is that awareness of traffic information on a road segment is available only if there is relatively uniform flow of vehicles through that segment. The only exception to this limitation is when "enough" fraction of vehicles are instrumented. Then *"no news is good news"*.
3. For traffic information to be useful, a vehicle should receive it well in advance to be able to take an alternate faster route. This is especially true for highways, where the diversions points (exits) are typically separated by a "few" miles. This suggests that a typical querying, monitoring, or dissemination operation would involve communicating travel time information over a relatively long distance ("few" miles).
4. Density of participating vehicles in an area may vary dramatically with time (usually it drops substantially at night) and space (rural regions typically have lower vehicle density than urban ones).

2.1 TrafficRep device

A TrafficRep device is attached to three components: a GPS device, a static digital map database of the road network, and a wireless communication interface (Figure 1). We assume that the static digital map database is organized by road segments, where a road segment is a stretch of a road between two successive exit points (junction, exits, etc). For each road segment, the database stores three attributes:

1. GPS coordinates of its endpoints,
2. Length of the segment, and
3. Free flow traveling time (length of the segment divided by the free-flow speed limit).

TrafficRep uses the location and time information from the GPS unit and the static information about location of end-points of road-segments to calculate the travel time of vehicle for different road segments. Every time the vehicle travels on a road segment, TrafficRep records the corresponding travel time information as a travel log report (TLR). This includes identifier of the road segment, the travel time, and the time-stamp of the report. As TLRs get older, they are discarded by the TrafficRep device to create space for new ones.

The wireless communication interface may be a short-range communication link (e.g., 802.11 flavor) or a wide-area communication link (using cellular network (7) or two-way paging network). Our only requirement from the wireless communication interface is the capability to perform broadcast or multicast. TrafficRep device uses the wireless communication interface in order to support the basic operations: querying, monitoring, and dissemination.

The requirement that every vehicle be equipped with TrafficRep device along with GPS device and in-vehicle navigation system may seem like a deviation from our goal of zero additional infrastructure design.

There are two important reasons for this requirement: Firstly, we expect vehicles in future to come equipped with GPS device and in-vehicle navigation system. These two components have already started to appear even in low-end vehicles. Also, we expect TrafficRep device in itself to be quite cheap. This plus the fact that users may be willing to bear a part of its cost makes the deployment cost very small compared to that of existing solutions, for the same coverage. Secondly, equipping each vehicle with a device naturally makes the owner responsible for it. Thus, the cost of maintaining the system would be very small compared to that for equipment (cameras, helicopters, etc) needed in existing commercial solutions.

3 PROPOSED ARCHITECTURE

The "sensors-on-wheels" architecture has four entities: information *producers*, information *consumers*, information *aggregators*, and information *relayers*. Producers are vehicles that collect raw sensory data. Consumers are objects (e.g., vehicles) that query data. Aggregators collect and aggregate data from multiple producers. An aggregator is a logical entity that can reside on a stand-alone server or can even reside on a consumer. Relayers participate in forwarding messages to their intended destination.

In the transportation domain, with potentially millions of vehicles involved, the primary challenge in disseminating and querying dynamic sensory information is scalability. Both high scalability as well as zero additional infrastructure requirements makes *peer-to-peer* a natural architecture to investigate for us. In this architecture, vehicles communicate directly with other vehicles to disseminate traffic information and to process queries. This is in contrast to the centralized architecture, where all vehicles report traffic information to a central server, and the central server disseminates traffic information and answers queries from the vehicles. The centralized solution has three serious drawbacks: Firstly, the central server is single point of failure. Secondly, the central server may become a bottleneck for a system with millions of vehicles querying and updating it. Finally, having a central server brings in deployment and maintenance costs. We still simulate centralized solution in order to use its performance as a benchmark against which we compare the performance of our peer-to-peer solution.

4 PERFORMANCE METRICS

In the proposed peer-to-peer architecture, vehicles communicate directly with other vehicles. According to our back-of-the-envelope calculations, if every vehicle broadcast the collected TLR throughout US, every time it finishes traveling a road segment, it would require ˜83 Mbps bandwidth. Clearly it would not be possible to support this huge bandwidth requirement using the existing communication infrastructure. The important question is how to cut down on the bandwidth required while still maintaining a requisite level of accuracy in the knowledge of traffic conditions among vehicles. This suggests four performance metrics:

1. *Awareness:* This defines the minimum quality of monitoring possible in the system. For example, a vehicle is said to be aware of the query, *"What is the travel time on NJTP between exit 9 and 10?"*, with an error margin of 5%, if its knowledge of travel time is always within 5% of the true value.

2. *Coverage:* This is expressed in terms of the fraction of road network for which the system can support requisite level of awareness.

3. *Wireless bandwidth used:* It is important to keep the wireless bandwidth usage low. This is especially important when using commercial wide-area wireless link like cellular network, as every transmission has a cost associated with it. As a result the number of messages transmitted directly affects the cost of operation of the system.

An indirect way of measuring awareness is to consider the vehicles' travel time, and how close it is to the ideal one where the vehicle has complete traffic information. After all travel time is the metric that most users are directly concerned with.

5 SCOPE OF THE PAPER

As our first step, we assume that the system only provides support for dissemination mechanism. This simplifies our analysis of the proposed system. In this system a TrafficRep device maintains an estimate of the travel time on all the links. In the absence of any additional information, this estimate is set to the free-flow travel time on the link (obtained from static database of the road network). On receiving disseminated traffic reports, the vehicles update their travel time estimates. Since querying operation is not supported, the TrafficRep device assumes that the traffic information available locally is the accurate information. We investigate different dissemination mechanisms that would keep this local information close to the true value, and examine the bandwidth-usage versus accuracy tradeoff. Support for querying and monitoring operation is a subject of our future work.

6 GLOSSARY OF SYMBOLS

This section defines all the symbols commonly used in this paper:

L **:** set of recently visited links

L_{rep}**:** set of links that the vehicle has already disseminated information about

$e\hat{(l)}$**:** expected travel time for link l

$e(l)$**:** travel time actually experienced by the vehicle for link l

t_{last}**:** time of last dissemination for this vehicle

t_{cur}**:** current time

λ**:** maximum rate of dissemination (we measure it in terms of the minimum interval between two successive dissemination)

τ**:** minimum time interval between successive reports about the same link

Δ**:** minimum change threshold

ff(l)**:** travel time for link l based on its speed-limit (free-flow)

$h(l)$**:** last travel time report heard link for l. Reports heard more than k secs ago are discarded

rep(l)**:** last travel time reported for link l

7 DISSEMINATION MECHANISM

In order to disseminate information about traffic conditions, a vehicle needs to take three decisions: *when* to send the TLR, *what* to include in the TLR, and *where* to disseminate it. These three decisions are inter-related and they determine the bandwidth requirement of the system. In this paper, we explore two schemes for deciding *what* and *when* to disseminate:

1. *Naive Scheme:* In this scheme, vehicles transmit TLRs at a pre-defined maximum rate of dissemination. Whenever it is time to transmit a TLR, the vehicles looks at the set of recently traveled road segments and selects the one for which the difference between the expected travel time and the travel time actually experienced by this vehicle is the maximum. The pseudo-code for this scheme appears below.

```
function naive_scheme()
  if ($t_{cur} - t_{last} > \lambda$)
```

```
    $l_{max} = \arg\ max_{l \in L - L_{rep}}(\left|e(l) - e(\hat{l})\right|)$;
    disseminate_link_info($l_{max}$);
  end
```

2. *Smart Scheme:* In this scheme, vehicles send a TLR only when they have something "interesting" to report. Here, vehicles have a pre-defined maximum rate of dissemination. At this rate, vehicles select the most "interesting" road segment to report about. However, unlike in Naive scheme, a vehicle goes ahead with the transmission only when the difference between expected and actual travel time on the road segment exceeds a certain threshold. The pseudo-code for this scheme appears below.

```
function smart_scheme()
  disseminate = true;
  if ($t_{cur} - t_{last} < \lambda$)
    return;
  end

  $S = \{\forall l \in L - L_{rep} \mid h(l) = \texttt{null} \vee |h(l) - e(l)| > \Delta\}$;   ← Among the links which were never
                                               reported, exclude the ones for which the travel time recently
                                               heard is quite similar to what was experienced
  $l_{max} = \arg\ max_{l \in S}(\left|e(l) - e(\hat{l})\right|)$;
  if ($\left|e(\hat{l}_{max}) - e(l_{max})\right| < \Delta$)   ← check if the information is interesting
    if ($|\texttt{ff}(l_{max}) - e(l_{max})| < \Delta$)   ← check if travel time differs significantly from free flow
      disseminate = false;
    end
  end
  if (disseminate)
    disseminate_link_info($l_{max}$);
  end
```

In both the solutions, vehicles always disseminate traffic report about a road segment R in a geographic area centered on R. The size of the geographic area is determined by the dissemination scope. In our evaluation, we have chosen to keep the scope of dissemination as a parameter and evaluate the performance for different scope. In general, we would expect that larger the scope, the higher the chances that vehicles would know of traffic congestion and would be able to find a faster alternative route. The flip side is that the larger the scope, the more the number of broadcasts, resulting in higher bandwidth usage. This may translate into higher cost in case the link is a commercial wide-area wireless link (section 7.1).

We compare the performance of the proposed mechanisms based on peer-to-peer architecture with that for mechanisms based on centralized architecture. In the centralized solution, vehicles send their TLRs to the central server. They use the above two schemes, Naive and Smart, in order to decide *when* and *what* to report to the central server. The central server also needs to take decision on *what*, *when*, and *where* to disseminate. We discuss this in more detail in section 7.3.

7.1 Wireless link

We investigate the use of wide-area wireless link (e.g., two-way paging network, cellular network) in our system. We assume that the wide-area wireless network allows a vehicle to broadcast a message in the local cell (e.g., for cellular network this may entail sending a text message (SMS) to a specific cell-broadcast address). Note that this capability is already present in exiting cellular systems (17), though a normal user is not allowed to use it; only content providers (providing news, weather information, stock quotes) with tie-ups with cellular service providers are currently allowed to use the local broadcast feature (2). Every transmission in a cellular network has a cost associated with it. In order for the system to have low operational cost it is crucial that vehicles achieves requisite awareness with as few transmissions as possible. The investigation of a hybrid solution that makes use of both short-range but free communication and wide-area wireless link to achieve better cost versus awareness tradeoff is a subject of our future work.

As stated earlier, disseminating traffic information of general interest would reduce the number of queries. Here, a vehicle disseminates a TLR directly to other vehicles using cell broadcast. We associate a geographic scope with each TLR. We express this scope in terms of TTL (this is analogous to the concept of time-to-live (TTL) in wired networks). The TTL is expressed in terms of the number of hops of the cells in the cellular service. In this paper we consider an *idealized* scheme for implementing the concept of TTL, where the concept of TTL is supported by the cellular network. Such an *idealized* scheme simplifies the analysis by allowing us to concentrate on effect of other key parameters on performance. Investigation of how vehicles can collaborate and communicate to implement the concept of TTL is part of our future work.

7.2 Peer-to-peer solution

In the vehicle-to-vehicle solution, the TLR collected by a vehicle (information producer in this case) is sent directly to other vehicles via a wireless link. Aggregation of information is performed on the receiving vehicle – the information consumer. Each vehicle maintains an estimate of the travel time on all the road segments. On receiving a TLR, the vehicle updates its estimate of the travel time according to equation 1.

$$\widehat{e(l)} = \alpha \widehat{e(l)} + (1-\alpha)e(l) \tag{1}$$

where, α represents the decay factor and $e(l)$ refers to the travel time reported in TLR for link l.

Each TLR is disseminated in a geographic scope determined by its associated TTL. Consequently, each vehicle is aware of all information produced in the dissemination scope and can locally answer all queries with scope less than the dissemination scope.

7.3 Centralized solution

We simulate centralized solution in order to use its performance as a benchmark against which we compare the solutions of proposed schemes. In this solution, vehicles send TLR directly to the central server. Server maintains an estimate of travel time on all the road segments. Every time it receives a TLR containing a recent travel time experienced by a vehicle on a road segment, it updates this estimate (using equation 1). Also, if the server does not receive any TLR for a road segment for T_{age} for units of time, it resets its estimate to the free-flow travel time on the link. Central server continuously checks if it has any "interesting" traffic information to disseminate. For dissemination, it selects the link for which its estimate of travel time differs the most from what it disseminated last, and this difference is greater than a certain minimum threshold. The pseudo-code for the procedure appears below. When vehicles receive a traffic report from the central server, they simply set their estimate to the one reported.

```
function server__decide_link()
    $l_{max} = \arg\ max_{l \in L}(\left|\texttt{rep}(l) - \widehat{e(l)}\right|)$;
    if ($\left|\texttt{rep}(l_{max}) - e(\widehat{l_{max}})\right| \geq \triangle$)   $\leftarrow$ check if the information has changed since the last report
        return $l_{max}$;
    else
        $l_{max} = \arg\ max_{l \in L}(\left|\texttt{ff}(l) - \widehat{e(l)}\right|)$;
        // check if the information is interesting and last report about this
        // link was sent some time ago
        if ($\left|\texttt{ff}(l_{max}) - e(\widehat{l_{max}})\right| \geq \triangle) \wedge (t_{cur} - t_{last}(l_{max}) > \tau)$
```

```
            return $l_{max}$;
        end
    end
    return $\phi$;                                  ← return null
```

8 MICROSCOPIC SIMULATION MODEL OF SOUTH JERSEY NETWORK

To be meaningful, it is important to evaluate the performance of the proposed schemes using realistic data. To this end, we have chosen to simulate the peak afternoon traffic in Southern New Jersey area. Figure 2(a) shows the snapshot of the simulated transportation network. The transportation network model (23) used includes most of the highways in Southern New Jersey. The transportation network model consists of approximately 4000 road segments drawn to closely match reality. The parameters controlling the flow in our simulations were based on the data provided by the Delaware Valley Regional Planning Commission (DVRPC (4)) and calibrated in (23) to make sure that the traffic characteristics closely match the ones seen in reality.

We have performed extensive simulations using the *Paramics* simulator (8) to demonstrate the effectiveness of the proposed schemes for disseminating travel time information. Through its API, Paramics allows a programmer to maintain state on a per vehicle basis and alter the behavior of vehicles programmatically, including the path being followed by it. We extended the set of available APIs in two important ways:

- We added features for simulating vehicles that take fastest route using Dijkstra's shortest path algorithm (15).
- We added features for simulating cellular infrastructure that is geographically overlaid on the simulated road network. We divide the region encompassing the simulated road network into rectangular cells, and simulate a base station in each cell. Cars use cell-broadcast primitive to disseminate travel time information.

8.1 Simulation Methodology

Since the goal of this simulation is not to test dynamic traffic assignment techniques as discussed in (21, 14, 20), we employ a simple route choice mechanism described below. However, it is clear that the reliability of the simulation results can be improved using more sophisticated routing mechanisms as described in (21, 14, 20).

In order to simplify our analysis we consider only the vehicles traveling between a specific origin-destination zone pair. We classify the vehicles traveling between the target OD pair into two categories based on the routes they follow: *default* and *fastest*. The *default* vehicles always follow the fastest route based on the free-flow travel times in the network. This route is computed *a priori* and it does not change during the entire simulation. For all the simulations this default route remains the same. The *fastest* vehicles keep recomputing the fastest route to the destination based on the travel-log reports received. They alter their routes if an alternate route allows them to get to their destination faster. Only a fraction of the vehicles traveling between the target OD pair are categorized as *fastest*. This is to decouple the navigation problem that would arise from having all the vehicles choose the fastest route to the destination. In order to concentrate solely on the problem of effective dissemination of traffic information, we inject *fastest* vehicles at regular intervals into the network. These serve as probes in the network, continuously testing the accuracy of disseminated traffic information. The interval is chosen so as to keep the number of *fastest* vehicles small enough that they do not affect the travel time experienced by *default* vehicles, but large enough to continuously test the effectiveness of dissemination mechanism.

We measure awareness indirectly using total travel time. Total travel time for a class of vehicle is the sum of the travel time of all the vehicles of the same class traveling between target OD pair. Clearly, vehicles with accurate travel time information will experience smaller travel time. Thus this metric captures the effectiveness of the dissemination algorithm.

In order to evaluate the effectiveness of the dissemination mechanism, we simulate an incident on the default route in our simulations. This incident results in closure of lanes. The incident occurs 40 mins after the start of simulation, and lasts for 25 minutes. The incident site is shown in the zoomed-in section of the transportation network in figure 2(b).

We use the following parameters to control the behavior of the dissemination algorithms:

1. *Maximum rate of dissemination:* The different values of maximum dissemination rate used in the simulations are: once every 60 secs, once every 300 secs, and once every 600 secs.
2. *Geographic scope of each TLR:* In our simulations, we divide the physical region in rectangular cells of size 2 Km x 2Km. Cells of this size would come under the "small" category in a real cellular network (cell sizes typically vary in radius from 1Km to 35 Kms (17)). We chose smaller size in order to examine, at a finer level, how far the information needs to be propagated for it to positively affect travel time of vehicles. We express geographic scope in terms of TTL, which controls the number of hops a TLR travels. The TTL values used in the simulation are: 1, 2, 3, and 5 (dissemination scope of TTL=5, centered at the accident site, covers the entire simulated road network).
3. *Dissemination scheme:* We evaluate the performance of Naive scheme and Smart scheme for dissemination.
4. *Market Penetration:* This determines the percentage of vehicles participating in the dissemination mechanism. We vary the market penetration to explore its effect on performance.

8.2 Simulation Results

We start by examining the effect of above parameters on the performance of peer-to-peer scheme. The impact of these parameters remains the same, qualitatively for the centralized scheme. We then compare the performance of centralized scheme with peer-to-peer scheme.

8.2.1 Peer-to-peer solution

Effect of TTL

We examine the effect of TTL on the performance of Naive dissemination scheme. The maximum dissemination rate is set to once per 60 secs, and we simulate a market penetration of 10%. Figure 3(a) summarizes the result. The bar graph plots for different values of TTL, the percentage reduction in the travel time of *fastest* vehicles over that of *default* vehicles during the accident period. As expected the performance improves with increase in TTL. For TTL=1, the performance of *fastest* vehicles is very similar (slightly degraded) to that of *default* vehicles. This is to be expected because by the time the vehicles come to know of the higher trip times on road segments they are too close to the traffic congestion and do not have any better alternative route to avoid it. With larger value of TTL, the vehicles get to know of congestion state in advance and are able to avoid it by taking faster alternative route. This clearly illustrates the importance of disseminating traffic information in a wider geographic scope. The results show that beyond the value TTL=3, the performance improvement flattens out. This has important implications. This suggests that in real systems, for regions where cell size is greater than 6 Kms, we do not need any scheme for disseminating information beyond the local cell. Clearly, this number is specific to the transportation network used in the simulation. However, given the relatively large cell sizes in real cellular networks (1 km to 35 Kms), it seems that a local cell broadcast would suffice in majority of the cases. This suggests that schemes that are able to disseminating information with arbitrary TTL values may be an overkill, and simple schemes that work well for smaller values (e.g., 2) of TTL would be sufficient.

Note that the amount of improvement achieved is also a function of the duration of incident and the number of lanes affected by it. Clearly if an incident lasts longer and affects more lanes, potentially *fastest* vehicles will see greater reduction in travel time over *default* vehicles.

Effect of maximum rate of dissemination

We investigate the effect of the maximum rate of dissemination on performance. Consider a road segment, R. As stated earlier, the dissemination scope of every TLR reporting travel time on R has an associated geographic scope centered on R. For a fixed TTL, the geographic region in which TLRs for reports about R are disseminated remains constant. Vehicles outside the dissemination scope have very little idea about the traffic conditions on R. When a vehicle enters the dissemination scope for R and is intending to travel on this road segment, there is a limited window of time in which it should be warned about any congestion on R. As the vehicle gets closer to R, it has fewer alternate (faster) routes that avoid R. The rate of dissemination should be chosen such that with a very high probability, the vehicle hears at least one congestion report while it still has an option of avoiding it. The required rate of dissemination is directly proportional to the length of window of time available to vehicles. This would argue for a very high dissemination rate. However, the flip side is that higher the dissemination rate, the more the number of messages transmitted, resulting in higher cost of operation. In order to illustrate the impact of maximum dissemination rate, we performed simulation with TTL value of 3, market penetration of 10%, and with Naive dissemination scheme. Figures 3(b) and 3(c) shows the impact of change in maximum dissemination rate on the performance (the percentage improvement in travel time of *fastest* vehicles over *default* vehicles). TTL value was set to 3 because for this setup it gives close to peak performance (figure 3(a)). This helps us in isolating the impact of maximum dissemination rate.

As the scope increases, the window of time available to a vehicle to choose an alternate route increases. This implies that increase in TTL can compensate for decrease in rate of dissemination. The performance graph for TTL=5 shown in figure 3(c) illustrates this.

Naive versus Smart scheme

With respect to the number of messages transmitted the Smart scheme outperforms Naive scheme by a wide margin, an order an magnitude in certain cases. Figure 4(a) shows the average number of messages transmitted per cell for different values of TTL, for 10% market penetration and maximum dissemination rate of 60 secs. Figure 4(b) shows the average number of messages transmitted per cell for different values of maximum rate of dissemination for TTL of 3. These results are to be expected because in Smart scheme vehicles transmit a message only when they have something “interesting” to communicate.

Now consider the effect on travel time. Figure 4(c) compares the travel time for the two schemes for different values of TTL for the case when the maximum rate of dissemination is 60 secs and the market penetration is 10%. Figure 4(d) compares the performance of two schemes for different maximum dissemination rates. Somewhat counter-intuitively Smart scheme performs better with respect to travel time. This is because in Smart scheme vehicles transmit only “interesting” travel time information. As a result, the travel time estimates of vehicles are able to follow the true travel time more closely. This is especially true for road segments whose travel time information is changing quickly.

Effect of market penetration

The higher the market penetration, the closer we get to the ideal case where for every road segment there are always vehicles sensing traffic conditions on it and reporting if they detect anything unusual. Intuitively this should transform into greater improvement in travel time of the vehicles. We examined the performance for three different values of market penetration: 5%, 10%, and 20%. We considered two cases. In the first case, we kept the maximum dissemination rate constant at 60 secs and varied the scope of dissemination

(TTL). In the second case, we kept the TTL fixed at 3 and varied the maximum dissemination rate. Results for the two cases appear in figure 5. For both the cases we see that the performance improves as the market penetration increases. Smart scheme performs slightly better than Naive scheme because vehicles' travel time estimate is able to follow the change in true travel time very closely.

We find that for Smart Scheme, graph for number of messages increases versus market penetration increases in the beginning, but flattens out for higher values of market penetrations. This is to be expected because in Smart scheme vehicles only transmit "interesting" information (we omitted this graph for lack of space).

8.2.2 Centralized solution

In this section, we compare the results for peer-to-peer and centralized solution (in both cases, Smart scheme is used). When compared in terms of improvement in travel time, the two schemes have similar performance (figure 6) with peer-to-peer solution performing slightly better in some cases. However, when compared on the basis on number of messages received per hour, peer-to-peer scheme transmitted 4 to 6 times more messages than centralized scheme (figure 6). The reason is that in centralized solution the central server makes more efficient use of downlink bandwidth. Peer-to-peer solution does not fare well because vehicles take independent decisions on when to transmit a TLR. This illustrates that if an infrastructure already exist, and its maintenance cost is not an issue, it would be more resource efficient to make use of it. However, the system should not be dependent on existence of an infrastructure for its functioning. In our ongoing work, our goal is to bring the bandwidth usage of peer-to-peer solution as close to the centralized solution as possible.

9 CONCLUSIONS AND FUTURE WORK

In this paper we have argued that the only solution that can be deployed on large scale is one that requires zero additional infrastructure. We have presented such a solution, and analyzed it using microscopic simulations and well defined performance metrics. We found that for the simulated transportation network, the performance improvement flattened out beyond a dissemination scope of 6 Kms (TTL=3). Given the relatively large sizes of cells in cellular networks, this suggests that a local cell broadcast would be sufficient in most cases, and schemes for disseminating information with arbitrary TTL may be an overkill. We also compared the performance of proposed solution with an infrastructure based centralized solution. The results show that if the cost of infrastructure is not an issue, infrastructure based solution achieves the same performance while communicating 4 to 6 times less messages. As a result, an ideal solution should make use of any infrastructure that exists, while not relying on it for its functioning.

The proposed solution is a special case of a new architecture, called Grassroots (19), for collecting, disseminating, and querying information about the physical world. The Grassroots architecture consists of a massive number of individual agents called dataflies, which move around, collect, summarize, and query information about their immediate physical environments. The architecture needs highly scalable design for supporting dissemination of continuously changing information collected by millions of mobile dataflies, and supporting queries by potentially millions of objects. Apart from vehicles, such objects can include cell phones gathering information about network quality conditions, or environmental sensors (attached to vehicles, human beings, etc) that move and collect information about pollution or pollen levels.

10 ACKNOWLEDGMENT

This research was supported by the National Science Foundation's CAREER Award programs grant CAREER-CMS-9875296. This support is both acknowledged and appreciated. The contents of this paper reflect the views of the authors who are solely responsible for the facts and the accuracy of data presented herein. The contents do not necessarily reflect the official views of NSF.

We would also like to thank all the anonymous reviewers whose comments helped us in greatly improving the quality of this paper.

References

[1] 750NAV - Navigation system. Thales Navigation.
`http://www.magellangps.com`
`Accessed Aug 1, 2003.`

[2] 808 cell broadcasting. Tourist Info Online.
`http://touristinfoonline.turkcell.com.tr/tl/808.htm`
`Accessed Aug 1, 2003.`

[3] Advanced Driver and Vehicle Advisory Navigation Concept (ADVANCE).
`http://ais.its-program.anl.gov/advance/`
`Accessed Nov 1, 2003.`

[4] Delaware Valley Regional Planning Commission.
`http://www.dvrpc.org`
`Accessed Aug 1, 2003.`

[5] EtakTraffic - Real Time Traffic Information.
`http://www.etaktraffic.com`
`Accessed Aug 1, 2003.`

[6] Metrocommute Traffic Information.
`http://metrocommute.com`
`Accessed Aug 1, 2003.`

[7] Mp750 gps - wireless modem for global gsm/gprs networks. Sierra Wireless, Inc.
`http://www.sierrawireless.com`
`Accessed Aug 1, 2003.`

[8] Quadstone Paramics V4.0 - Microscopic traffic simulation system. Quadstone Limited.
`http://www.paramics-online.com`
`Accessed Aug 1, 2003.`

[9] Traffic Congestion and Sprawl. Federal Highway Administration, US Department of Transportation.
`http://www.fhwa.dot.gov/congestion/congpress.htm`
`Accessed Aug 1, 2003.`

[10] Traffic Information for US Cities.
`http://traffic.com`
`Accessed Aug 1, 2003.`

[11] IEEE Std. 802.11a - Wireless LAN Medium Access Control (MAC) and Physical layer (PHY) layer specifications: High Speed Physical Layer in the 5 GHz Band, 1999.

[12] M. Brackstone, G. Fisher, and M. McDonald. The Use of Probe Vehicles on Motorways, Some Empirical Observations. In *Proc. of 8th World Congress on ITS, Sydney, Australia*, October 2001.

[13] M. Chen and S.I.J. Chien. Transmit evaluation report. Research Report
`http://www.njtide.org/probe.doc`
`Accessed Aug 1, 2003.`

[14] P. Chen, K. Srinivasan, and H.S. Mahmassani. Effect of information quality on compliance behaviour of commuters under real-time information. *Transportation Research Record 1676*, pages 53–60, 1999.

[15] T.H. Cormen, C.E. Leiserson, and R.L. Rivest. *Introduction to Algorithms.* MIT Press, Cambridge, Massachusetts, 1990.

[16] D.J. Dailey and F.W. Cathey. AVL-Equipped Vehicles as Traffic Probe Sensors, January 2002. Research Report `http://www.its.washington.edu/pubs/probes_report.pdf`
`Accessed Aug 1, 2003.`

[17] Andy Dornan. *The Essential Guide to Wireless Communications Applications.* Prentice Hall, 2001.

[18] C. Elango and D.J. Dailey. Irregularly Sampled Transit Vehicles Used as a Probe Vehicle Traffic Sensor. *Transportation Research Record 1719*, pages 33–44, 2000.

[19] Samir Goel, Tomasz Imielinski, Kaan Ozbay, and Badri Nath. Grassroots – a scalable and robust information architecture. Technical Report DCS-TR-523, Rutgers University.

[20] P. Kachroo and K. Ozbay. Feedback control theory for dynamic traffic assignment. *Springer-Verlag Series on Advances in Industrial Control, Springer-Verlag*, 1999.

[21] H.S. Mahmassani. Dynamic network traffic assignment and simulation methodology for advanced systems management applications. *Network and Spatial Economics 1*, pages 267–292, 2001.

[22] K. Mouskos, E. Niver, L.J. Pignataro, S. Lee, N. Antoniou, and L. Papadopoulos. Determining the Number of Probe Vehicles for Freeway Travel Time Estimation Using Microscopic Simulation, June 1998. Research Report `http://www.njtide.org/TRANSMITEv.pdf`
`Accessed Aug 1, 2003.`

[23] K. Ozbay and B. Bartin. South Jersey Real-time Motorist Information System, NJDOT Project Report. March 2003.

[24] David Schrank and Tim Lomax. *The 2002 Urban Mobility Report.* Texas A&M University: Texas Transportation Institute, 2002.

[25] L. Wischhof, A. Ebner, H. rohling, M. Lott, and R. Halfmann. SOTIS - A Self-Organizing Traffic Information System. In *Proc. of 57th IEEE Vehicular Technology Conference (VTC 03 Spring), Jeju, South Korea*, April 2003.

[26] A.K. Ziliaskopoulos and J. Zhang. A Zero Public Infrastructure Based Traffic Information System. In *Proc. of 82nd Annual Meeting of Transportation Research Board*, January 2003.

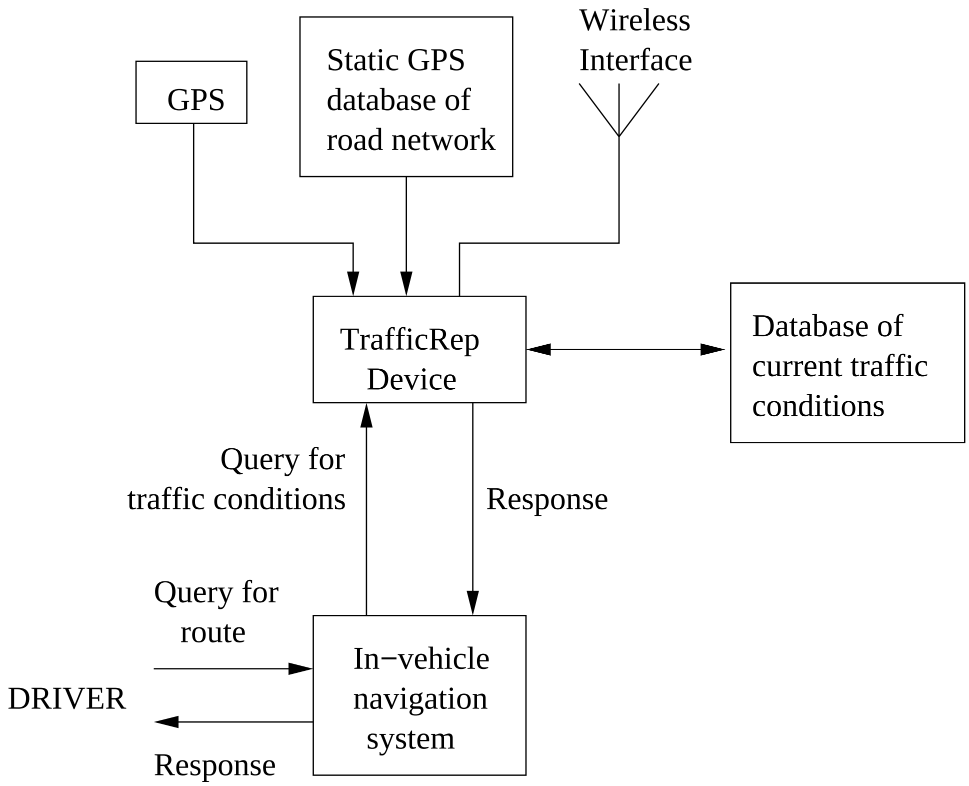
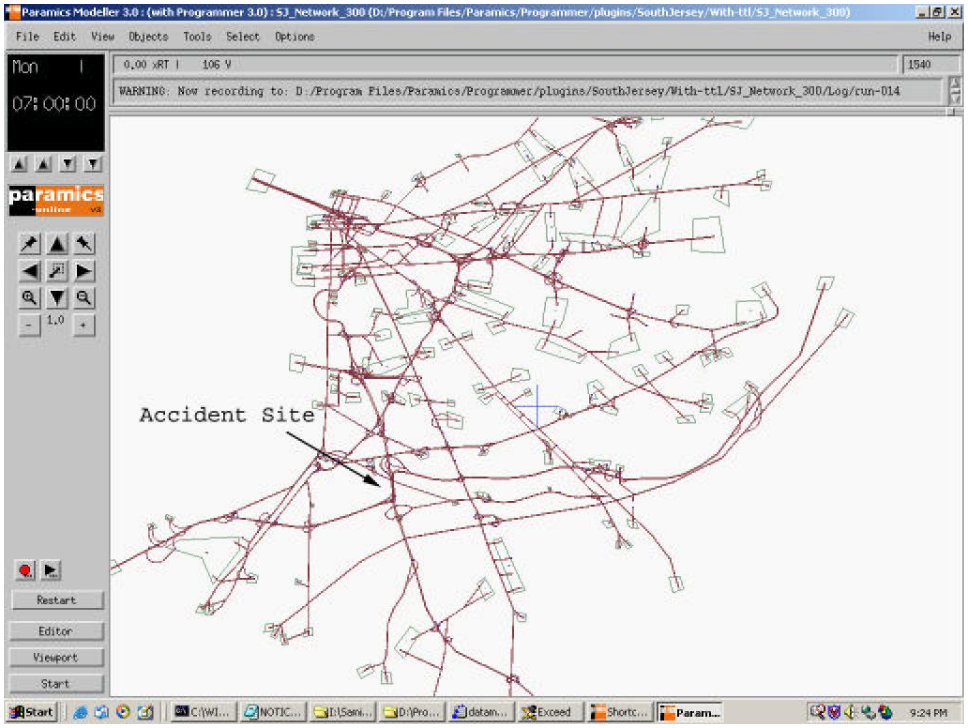


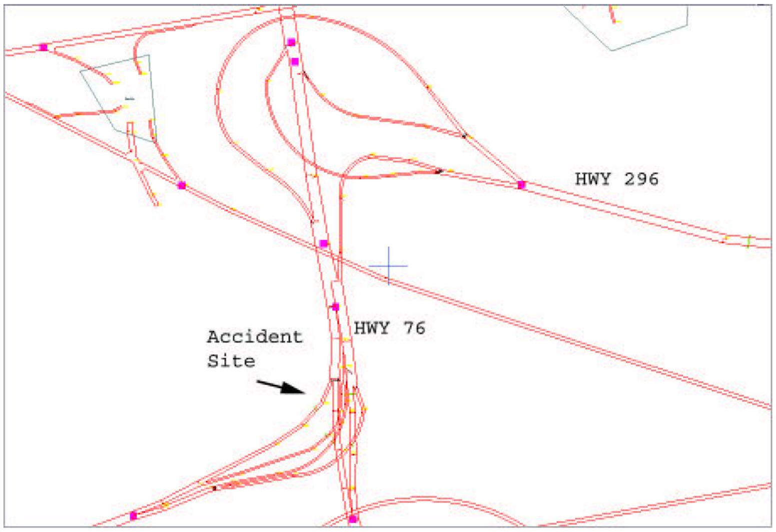
Figure 1: Schematic diagram of the components that will be present in each participating vehicle in the envisioned system



Paramics Modeller 3.0 : (with Programmer 3.0) : SJ_Network_300 (D:/Program Files/Paramics/Programmer/plugins/SouthJersey/With-ttl/SJ_Network_300)

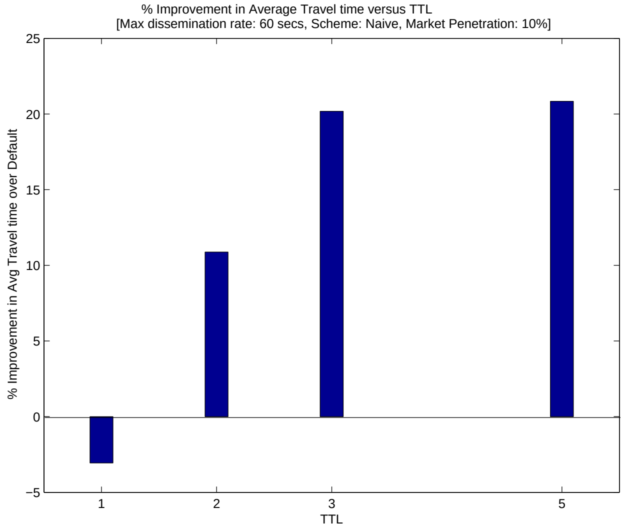
WARNING: Now recording to: D:/Program Files/Paramics/Programmer/plugins/SouthJersey/With-ttl/SJ_Network_300/Log/run-014

(a) Snapshot

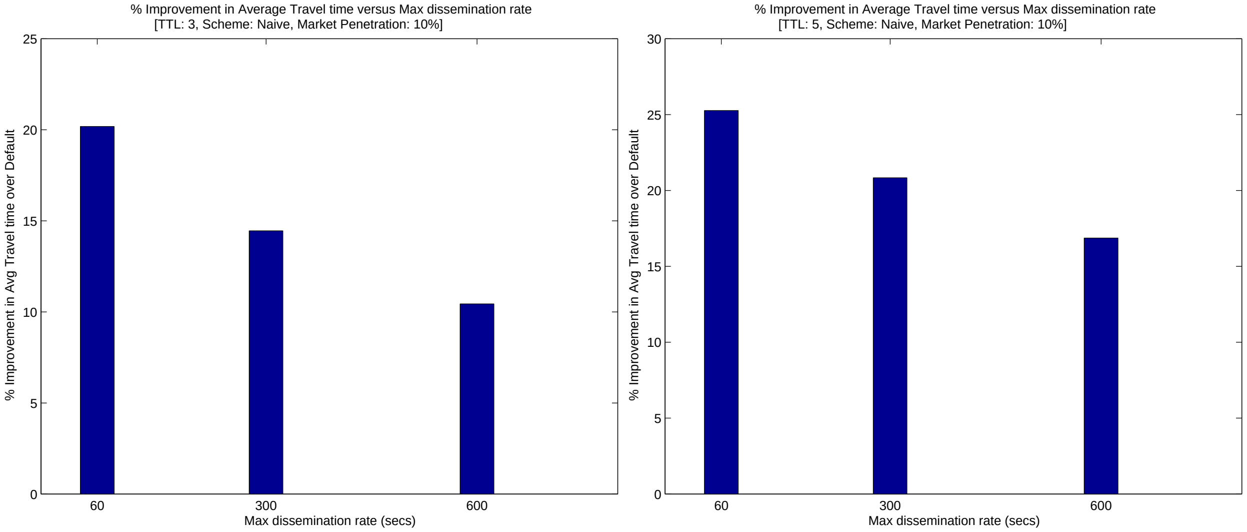


(b) Incident site

Figure 2: Snapshot and incident site



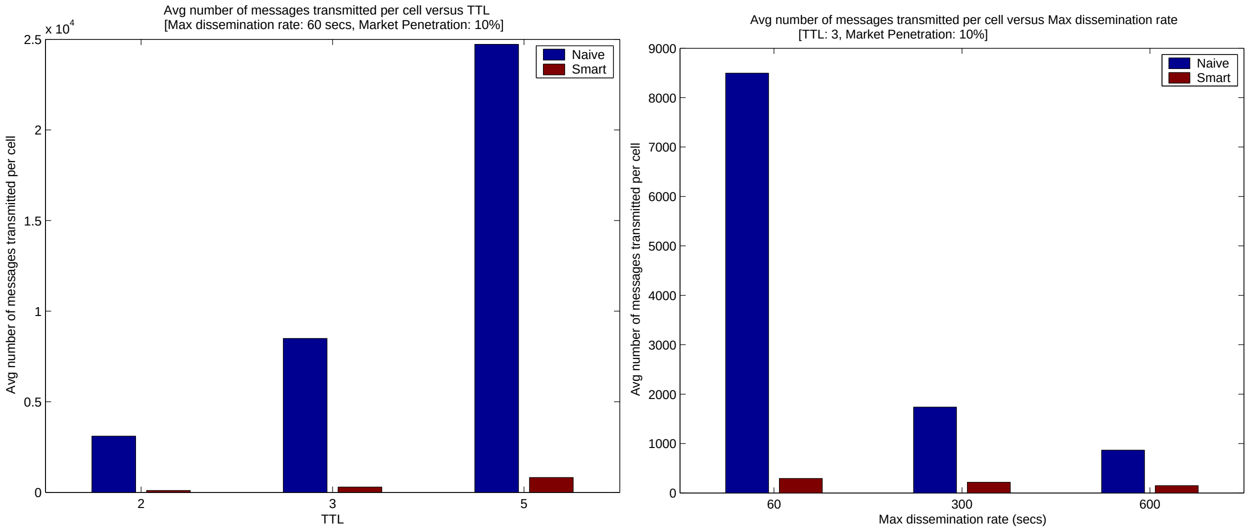
(a) Effect of TTL on performance



(b) Effect of Max dissemination rate on performance (TTL=3)

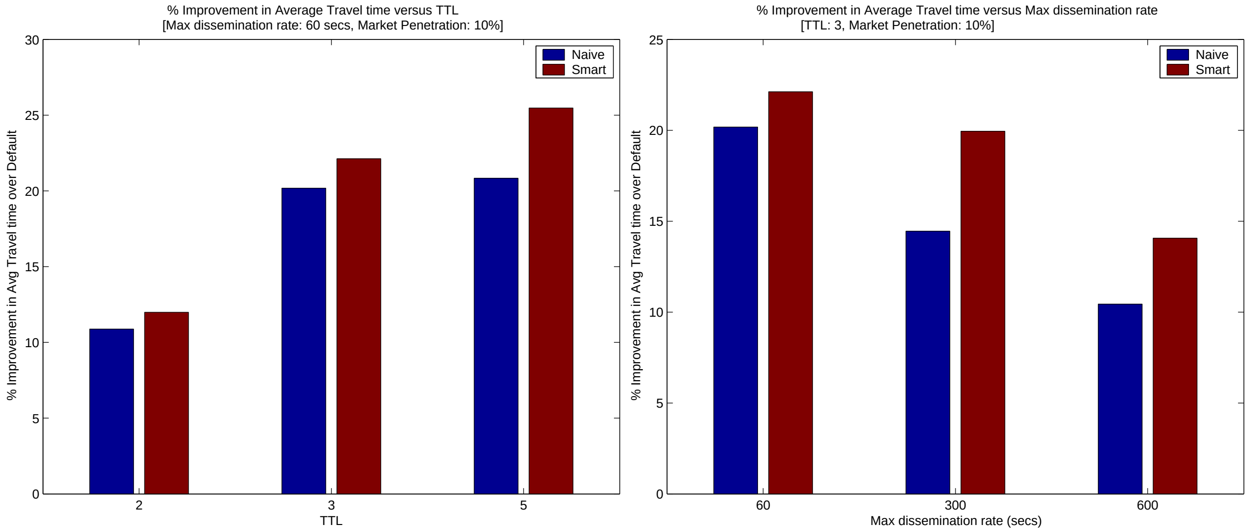
(c) Effect of Max dissemination rate on performance (TTL=5)

Figure 3: Simulation results showing effect of TTL and maximum dissemination rate on performance of Naive scheme



(a) Number of messages transmitted versus TTL (Max rate of dissemination: 60 secs)

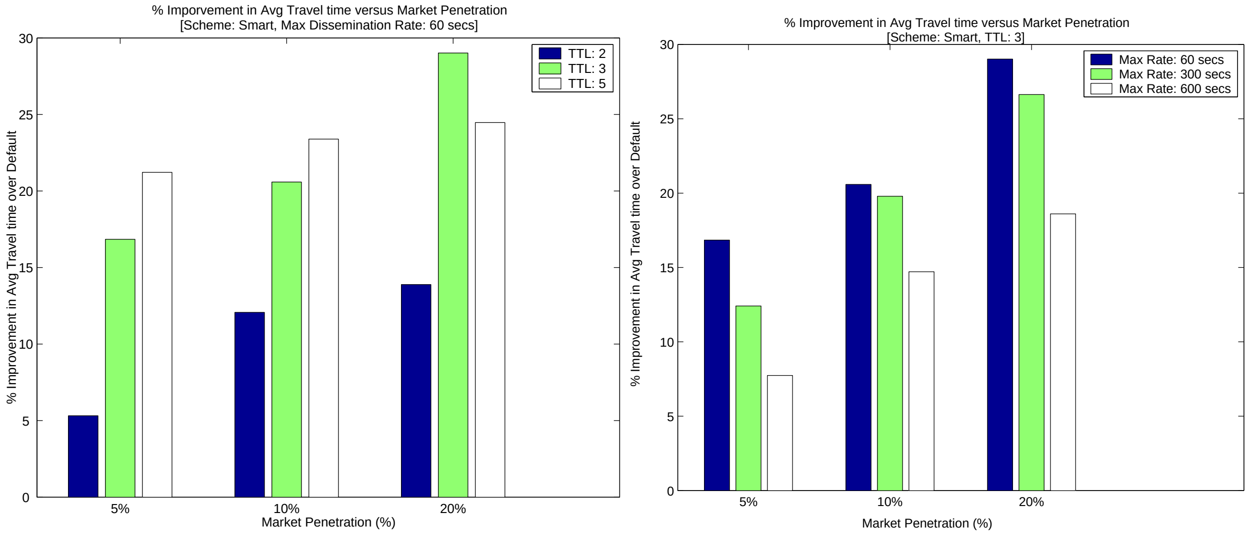
(b) Number of messages transmitted versus Max dissemination rate (TTL=3)



(c) Improvement in travel time versus TTL (Max dissemination rate = 60 secs)

(d) Improvement in travel time versus Max dissemination rate (TTL=3)

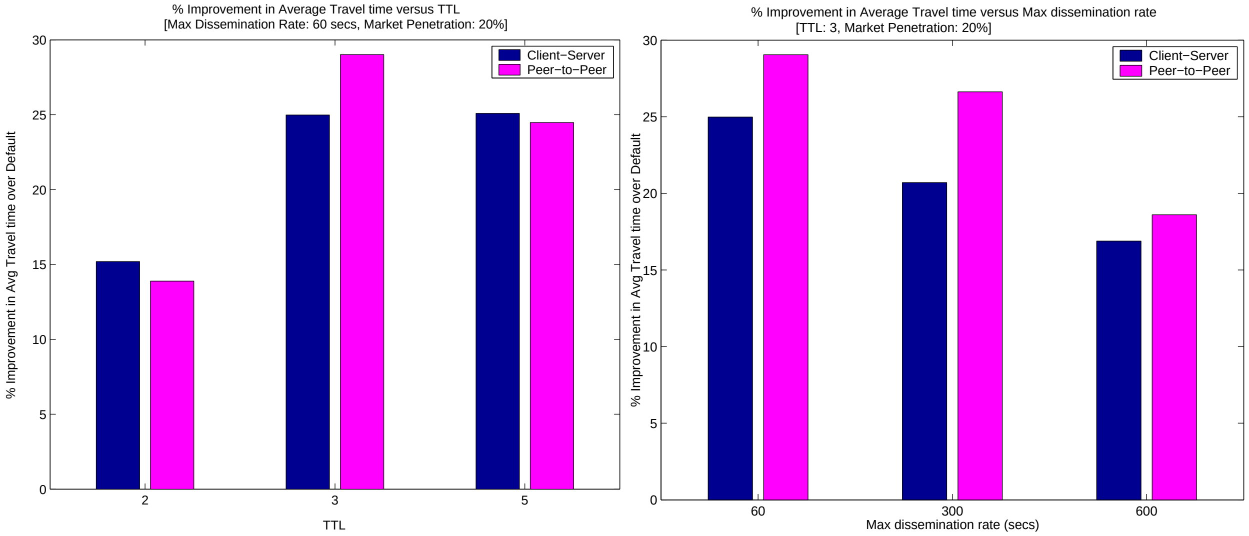
Figure 4: Simulations results comparing the performance of Naive and Smart scheme (market penetration: 10%)



(a) Maximum dissemination rate = 60 secs

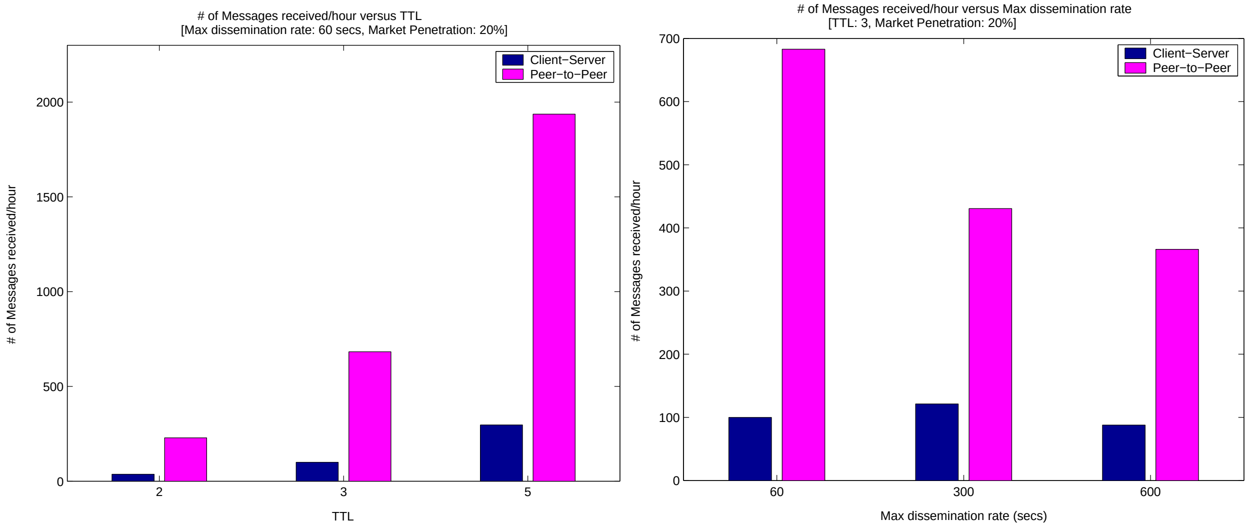
(b) TTL = 3

Figure 5: Effect of market penetration on performance of Smart scheme



(a) Improvement in travel time versus TTL (Max dissemination rate: 60 secs)

(b) Improvement in travel time versus Max dissemination rate (TTL = 3)



(c) Number of messages transmitted versus TTL (Max dissemination rate: 60 secs)

(d) Number of messages transmitted versus Max dissemination rate (TTL: 3)

Figure 6: Simulation results comparing the performance of peer-to-peer and centralized solution