Neighborhood Changing Rate: An Unifying Parameter to charaterize and evaluate Data Dissemination scenarios

Jérôme Härri, Biao Zhou, Mario Gerla, Fethi Filali, and Christian Bonnet
Department of Mobile Communications
Institut Eurécom
06904 Sophia-Antipolis, France
Department of Computer Science
University of California
Los Angeles, CA 90095
{haerri,filali,bonnet}@eurecom.fr, {zhb,gerla}@cs.ucla.edu

Abstract— **Vehicular sensor networks are emerging as a new network paradigm particularly relevant to proactive monitoring of urban environment. In this mobile environment, each sensor can generate a large amount of data which must be reliably reported to actuator agents. Data dissemination in conjunction with efficient harvesting has proven to be very effective in this type of applications. One major property in data dissemination/harvesting is the ability of a node to discover new neighbors as it moves. The performance depends on many different parameters including speed, motion pattern, node density, data rate, transmission range. This multitude makes it difficult to accurately evaluate and compare data gathering protocols implemented, for example, on different simulation or testbed scenarios.**

In this paper, we introduce Neighborhood Changing Rate (NCR) - a parameter that cumulatively captures the essence of several other parameters relevant to data dissemination. By its intrinsic property, the NCR measure can well characterize a dissemination/harvesting scenario and allow to predict the performance of such scenario. We illustrate our approach by applying the NCR concept to Mobeyes, a lightweight data gathering protocol.

I. INTRODUCTION

Dissemination has shown to be the most efficient and non-intrusive way to transmit the seed of life, as even the most remote place on earth has been reached. Nature adapted itself to benefit from natural mobility in order to improve its dissemination efficiency. Every living form contributes to it, either actively or passively. As a matter of fact, humans must deal with dissemination on a daily basis, the best known example, and the most feared one, being virus dissemination. When a human is contaminated by a virus, it passively carries it within its body until it encounters another living form to which the virus is transmitted. This kind of epidemic transmission is very efficient (in fact, often catastrophic), as human mobility and density contribute to make it virtually unstoppable. The close similarity between human virus carriers and routers in a data network has became a source of inspiration for characterizing digital dissemination techniques.

Transmitting information in a mobile ad hoc network (MANET) has shown to be a very challenging subject, as mobility and sparseness put most routing protocols in difficulty. Vehicular ad hoc networks (VANETs), a special class of MANETs, have even more pronounced mobility rate and topology change characteristics. When a large volume of data needs to be transmitted to actuator nodes, as it may be the case of vehicular sources, those kinds of characteristics render data transmission unreliable. However, by analogy with epidemic dissemination, mobility can often be seen as an opportunity to improve data exchange, even over sparse networks.

A variety of approaches for efficient data dissemination have been studied. For geographic routing, data dissemination helps propagate nodes' geographic locations in the network [1], therefore relaxing the requirement for a centralized location server. In Delay Tolerant Networks (DTN) [2], data dissemination has shown to be particularly efficient to convey information over sparse networks with high reliability. In Vehicular Sensor Networks (VSN), where vehicles are equipped with on-board sensing devices, data dissemination also showed to be particularly efficient in maintaining a distributed index for data harvesting and forensic investigation [3].

Evaluation of the data dissemination efficiency is a key parameter for estimating the performance of protocols that rely on it. Studies in epidemic dissemination have shown that performance is closely linked to mobility pattern and to density. Whereas density is a rather straightforward measure to define, the mobility pattern is much more difficult to characterize, which is probably why the research community restricted its interpretation to velocity. However, authors in [4] showed that mobility patterns had a stronger impact on routing performance than velocity when modeling realistic mobility. The major drawback of mobility pattern is that it is composed of several parameters. And evaluating protocols based on multi-criteria is hard, often yielding in consisting results. Reducing the "control panel" to a small set of independent parameters helps facilitate the fair comparison of protocols.

In this paper, we introduce the Neighborhood Changing Rate (NCR), an aggregate metric for microscopic "motion pattern" which is based on the rate of neighbors entering and leaving a neighbor set. We will illustrate how NCR has a significant influence on the performance of data dissemination, much more so than speed or density. When NCR is provided jointly with with velocity and density, it will be shown to fully describe data dissemination.

The outline of the paper is as follows. Section II gives a short introduction on data dissemination in VANET, while Section III introduce the Neighborhood Changing Rate. In Section IV, we describe Mobeyes, a protocol based on data dissemination for urban monitoring. We will use Mobeyes in order to validate NCR. Finally, in Section V, we provide simulation results illustrating the significance of NCR, and in Section VI we conclude the paper.

II. DATA DISSEMINATION IN VEHICULAR AD HOC NETWORKS

Epidemic dissemination can also be applied to vehicular ad hoc networks. When a car has data to disseminate, it transmits it to other vehicles it encounters that have not yet received the data. In order to reduce the broadcast storm effect, cars do not relay this transmission. Then, each car that received the information shares it with any car it meets. The data dissemination is complete when all vehicles have received the information.

An example of this approach is given in Fig. 1, where $C1$, $C2$, $C3$ represent respectively Car 1, 2 and 3, and where t represent the time in seconds. The data is a security key which is only detained by $C1$ at time $t-3$. At time $t-2$, $C2$ encounter $C1$[1]. It shares the key with it and moves on. As $C2$ moves along its trajectory, it meets $C3$ at time $t-1$ and share the key with it. Eventually, the key will be disseminated to all cars in this example by simple encounters. The dissemination rate is defined as

$$rate^{diss} = \frac{\#cars^{new}}{\#cars^{old}}$$

where $cars^{new}$ is the number of cars that received the data, and $cars^{old}$ the cars that initially had the data. In Fig. 1, $rate = 2$.

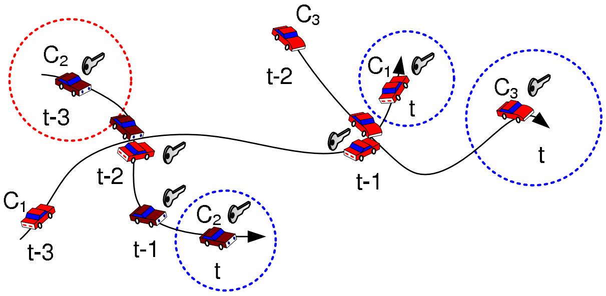


Fig. 1. Data dissemination with three vehicles.

What are the parameters controlling the dissemination rate ? Similarly to virus spreading, the larger set of vehicles a car meets per encounter point, the more efficient is the data dissemination. Indeed, in Fig. 2, $C2$ meets a larger set of cars at the first encounter point, thus increasing the rate to $rate^{diss} = 5$. In turn, group mobility is not a good factor for data dissemination, as a large group sharing the same data encounters other groups of potentially less or equal size. Then, as each car that gained the data is in turn able to spread it to other vehicles, the data dissemination may be increased even if the cars met at the encounter points do not follow a similar trajectory than the data bearer. Finally, each encounter point being an opportunity for data spreading, the rate a car encounters other neighbor cars also helps data dissemination. The data dissemination efficiency is therefore dependent to two major factors:

- The number of vehicles met per encounter point that do not follow a similar trajectory.
- The rate a car encounters other neighbors.

The major drawback for this formulation is that those two criteria are neither uncorrelated, nor atomic. In other words, they are both composed of, and potentially share, a multitude of parameters, such as velocity, density, or driving patterns.

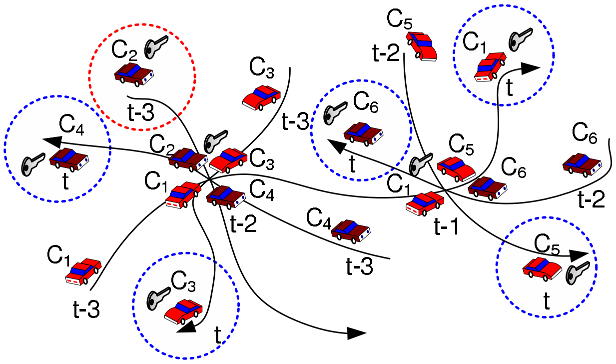


Fig. 2. Data dissemination with three vehicles.

III. NEIGHBORHOOD CHANGING RATE

Another intuitive formulation for the efficiency of data dissemination is the *Neighborhood Changing Rate (NCR)*. When sampled at appropriate time intervals, the interesting feature of this criterion is that it spans the previous parameters while at the same time it reduces to a simple abstraction of neighbors entering and leaving a neighbor set. In the rest of this paper, we will show how NCR is an efficient novel parameter describing data dissemination.

A. Definition

We define

$$NCR^i(t+\Delta t) = \frac{E[\#Nb^i_{new}](\Delta t) + E[\#Nb^i_{old}](\Delta t)}{2 \cdot E[\#Nb^i](t+\Delta t)} \quad (1)$$

as the NCR of node i at time $t + \Delta t$, with the following parameters

- Δt : Sampling time interval, which is the average time needed for a node to move a distance equal to its transmission range.
- $\#Nb^i_{new}$: Number of neighbors entering node i's neighborhood during the time interval Δt.

---

[1] We define an encounter as when two nodes are within mutual transmission range.

- $\#Nb^i_{old}$: Number of neighbors leaving node i's neighborhood during the time interval Δt.
- $\#Nb^i$: Node i's degree at time $t + \Delta t$.

Definition 1: the NCR has the following properties:

1) $0 \leq NCR(t) \leq 1$.
2) NCR $\perp$ $speed^{av}$.
3) NCR $\perp$ $density^{av}$.

Proof:

1) By contradiction, let $NCR(t) > 1$. This means $2 \cdot Deg(t)$ is smaller than the total number of neighbors which entered and left the neighborhood. In a network of n nodes, at any time instant, if the neighbor set degree is m, the maximum number of new neighbors that can enter between time $t-1$ and t is $n-m-1$. And the maximum number of neighbors that can leave during the same time interval is m. Therefore, the nominator becomes $n-1 \nleq 2 \cdot Deg(t) = 2(n-1)$.
2) By intuitive explanation, $\Delta t = \frac{Tx\_range}{speed}$. If $speed_1 > speed_2$, $\Rightarrow \Delta t_1 > \Delta t_2$. In other terms, if nodes have a faster mobility, the sampling interval will be shortened, canceling the effect of $speed^{av}$ (average speed) on NCR.
3) Also by intuitive explanation, a larger density may be related to a larger number of neighbors entering and leaving the neighborhood at the same time interval. However, as we normalize NCR, we cancel the effect of $density^{av}$ (average density) on NCR.

■

The last two properties may be intuitively illustrated by the possibility of obtaining different NCRs for the same average velocity or density. However, as we will show in Section V, data dissemination with same average speed or density have different performance, fostering the justification for NCR as a appropriate parameter modeling temporal and spatial dependencies and other mobility patterns that cannot be described by average speeds or velocities.

We depict in Fig. 3 the evolution of the NCR computed for the car $C1$. At time $NCR(t-3) = 0$ as the vehicle did not encounter other cars. At time $t-2$, $NCR(t-2) = \frac{1}{2}$, as $C1$ met $C2$ as new neighbor. And at time $t-1$, $NCR(t-1) = 1$, as $C1$ lost and gained a neighbor. Although we might argue that at time $t-1$, $C1$ had only one encounter and that the data dissemination capacity should be similar to the previous case. However, it also released $C2$ which will spread the information to other nodes, a situation that would not be possible if $C2$ had stayed in $C1$ neighborhood, thus the difference of NCRs.

B. Discussion

The performance of protocols for vehicular ad hoc networks using data dissemination depends on a multitude of factors, such as speed, density, topology, mobility patterns. And evaluating protocols based on multi-criteria is hard and often yield to arguable results. Reducing the panel to a short set of independent parameters helps having a fair comparison between protocols.

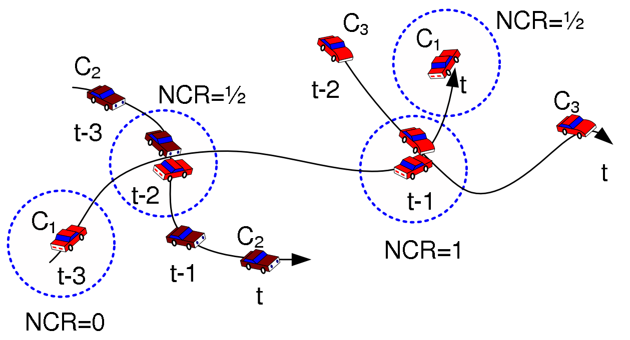


Fig. 3. Example of the evolution of the NCR for car 1.

A similar situation also exists in Transportation Planning. How to represent traffic flows in transportation that depend on multi-parameters, such as speed, density, volume/capacity ?

The community defined the *Level of Service (LOS)*. It works like an American report card grade, using the letters A through F, with A being best and F being worst.

By using LOS classification and referring to a traffic situation as having a particular LOS , engineers can have a global knowledge of traffic condition in a particular area.

NCR is designed to have the same usage. Indeed, by referring to data dissemination as having a particular NCR, we can have an intuitive vision of its efficiency, and thus evaluate accurately VANET protocols using this feature.

IV. MOBEYES

Data mining of some sensed events potentially monitored by distributed mobile sensors, require the collection, storage, and retrieval of massive amounts of sensed data. This is a major departure from conventional sensor networks where data are usually collected, examined, and dispatched to a "sink" under predefined conditions, such as alarm thresholds. Then, this becomes the problem of searching in a massive, mobile, and completely decentralized storage of sensed data, by ensuring low intrusiveness, good scalability, and disruption tolerance against sensor mobility via completely decentralized cooperation.

Mobeyes has been precisely designed to tackle this issue, as it is an efficient lightweight support for proactive urban monitoring based on the primary idea of exploiting vehicle mobility to opportunistically diffuse summaries about sensed data. We decided to use this protocol to illustrate the effect of NCR, as Mobeyes is a typical example of protocols mainly based on the use of efficient data dissemination for data retrieval. In the following section, we shortly introduce the functionalities of Mobeyes. For a full description and performance evaluation, the reader is referred to [3].

A. Mobeyes Architecture

Mobeyes is composed of three main components:

- Mobeyes Sensing Interface (MSI) : It represents the sensor control layer and is responsible for the access to sensors or GPS.

- Mobeyes Data Processor (MDP): It is in charge of reading raw sensed data obtained from MSI, and of generating summaries.
- Mobeyes Diffusion/Harvesting Processor: This is the core of Mobeyes data dissemination. MDHP works by opportunistically disseminating/harvesting summaries produced by MDP.

The first two components are beyond the scope of this paper and we will not further describe it. In the next section, we will focus instead on MDHP.

B. MDHP Protocols

The MDHP has a dual role to play. Indeed, it needs to efficiently disseminate the summaries generated by the MDP. Then, on demand, it is also in charge of harvesting the summaries to an actuator node.

1) Summary Diffusion: Any regular node periodically advertises a packet with newly generated summaries to its current neighbors. Each packet is uniquely identified (generator ID + locally unique sequence number). This advertisement to neighbors provides more opportunities to the agents to harvest the summaries, and the duration of periodic advertisements should be chosen properly to fulfill the desired latency requirements because harvesting latency depends on it. Neighbors receiving a packet store it in their local summary databases. Therefore, depending on node mobility and encounters, packets are opportunistically diffused into the network.

Figure 4 depicts the case of a vehicular node $C1$ encountering with other vehicular nodes while moving (for the sake of readability, only C2 is explicitly represented). Encounters occur when two nodes exchange summaries, i.e., when they are within their radio ranges and have a new summary packet to advertise. In the figure dotted circles and time stamped triangles represent respectively radio ranges and C1 encounters. In particular, the figure shows that C1 (while advertising $S_{C1,1}$) encounters C2 (advertising $S_{C2,1}$) at time $T - t_4$. As a result, after $T - t_4$ C1 includes $S_{C2,1}$ in its storage, and C2 includes $S_{C1,1}$.

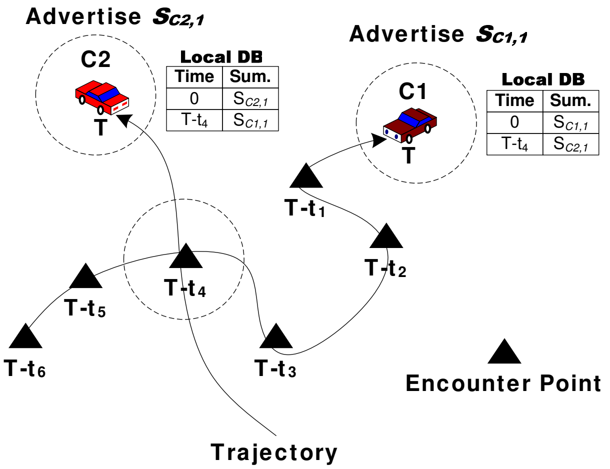


Fig. 4. Mobeyes single hop passive diffusion.

2) Summary Harvesting: In parallel with diffusion, summary harvesting can take place. A MobEyes actuator node can request the collection of diffused summaries by proactively querying its neighbor regular nodes. The ultimate goal is to collect all the summaries generated in a given area. Obviously, the actuator is interested in harvesting summaries it has not collected so far: to focus only on missing packets, a MobEyes actuator compares its already collected packets with the packet list at each neighbor, by exploiting a space-efficient data structure for membership checking, i.e., a Bloom filter[2].

Therefore, the MobEyes harvesting procedure consists of the following steps:

1) The actuator node broadcasts a “harvest” request with its Bloom filter.
2) Each neighbor prepares a list of “missing” packets from the received Bloom filter.
3) One of the neighbors returns missing packets to the actuator.
4) The actuator sends back an acknowledgment with a piggybacked list of just received packets. Upon listening or overhearing this, neighbors update their missing packet lists for the actuator.
5) Steps 3 and 4 are repeated until there is no remaining packet.

V. SIMULATION RESULTS

We have performed a range of simulations on Mobeyes in order to illustrate the effect of NCR. In these simulations, we are interested in the harvesting latency as a function of speed and the harvesting rate as a function of time. From [3], we know that the latency drops as speed increases, but we do not know how the NCR will affect the latency. For mobility modeling, we used the steady-state Random Waypoint on graphs [5] and the Track Model [6] on the topologies illustrated in Fig. 5.

To capture the most representative features of motion in MANET, we have developed a track group mobility model based on a Markov Chain approach during the past one year. The tracks are represented by freeways and local streets. The grouped nodes must move following the tracks and share the same group movement. Each group member has an internal random mobility within the scope of a group. At each intersection (switch station) a group can be split into multiple smaller groups; or may be merged into a bigger group. The track model allows also individually moving nodes as well as static nodes. Such non-grouped nodes are not restricted by the switch stations and tracks and are able to move around the whole field. The track model has been tested with real freeway/street maps from the US census bureau.

In order to show the effect of an increased neighborhood changing rate on the performance, we use mobility patterns with three different NCR, a low NCR, medium NCR and high NCR on each set of simulation. And to remove the influence of the density from the results and perform cross-topology comparison, we normalized the results by the average density

---

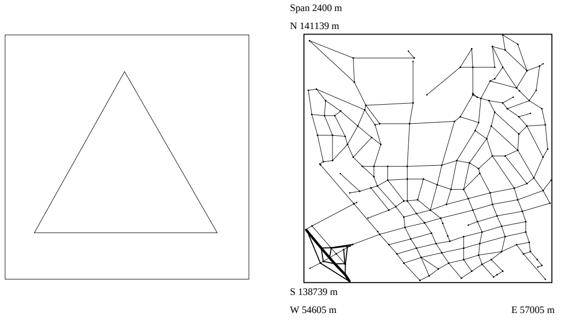
[2] For a definition of the Bloom filter, the reader is refereed to [3]

in Fig. 9. Finally, we evaluate the performance by randomly choosing 10 agents and run each of three scenarios 30 times. The simulation parameters can be found in Table I.

Network Simulator	ns-2 2.27
Mobility models	Random Trip [5], Track Model [6]
Pause time	0s
Data dissemination protocol	Mobeyes
$Hello^{aodv}$ Interval	3s
Event Generation Interval	10'000s
Number of harvesting Agents	10
Number of runs	30
Simulation time	2000s
Simulation Area	2400m x 2400m grid
Number of Nodes	100
Tx Range	250m
Speed	$5m/s \rightarrow 25m/s$
Density	$\#nodes \cdot \frac{\pi \cdot range^2}{X_{dim} \cdot Y_{dim}}$

TABLE I

SIMULATION PARAMETERS

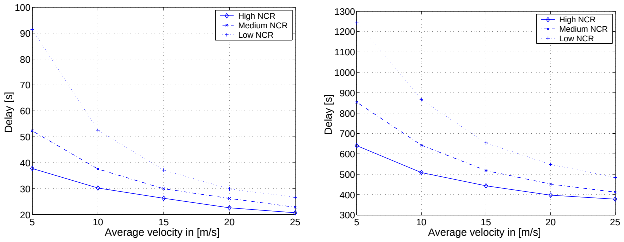


(a) Triangle topology (edge length=$760m$)

(b) Map Topology

Fig. 5. Two topologies used by the mobility models

We show in Fig. 6(a) and Fig. 6(b) the effect of the average velocity on the latency for respectively the triangle and the map topology. As we could expect, the latency is reduced when we increase the speed. However, what we can also see on those two figures is the influence of the NCR. The performance is improved as the NCR increases.

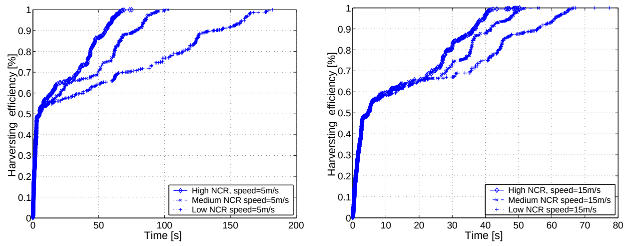


(a) Triangle topology

(b) Map Topology

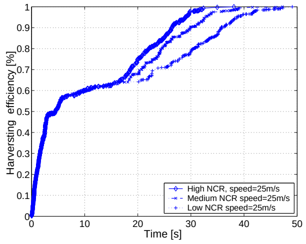
Fig. 6. Average diffusion latency, for different NCRs

In Fig. 7, we illustrate the harvesting rate as a function of time for average velocities $5, 15, 25m/s$. We can see the effect of increasing the NCR on the harvesting time. For example, on Fig. 7(a), after $75s$, Mobeyes has harvested 100% of the events for the high NCR, but only 90% and 70% for the medium and low NCR respectively. We see similar effects for all different average velocities. This shows that an increased neighborhood changing rate also increases the dissemination of data in the vehicular network.



(a) Average velocity= 5m/s

(b) Average velocity= 15m/s



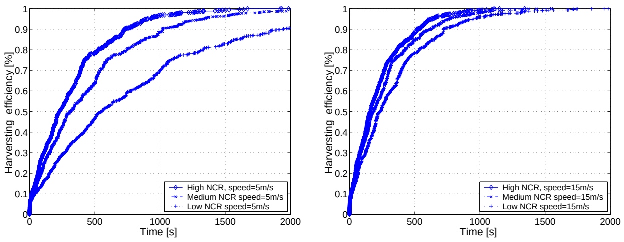
(c) Average velocity= 25m/s

Fig. 7. Faction of harvested summaries on the triangle topology

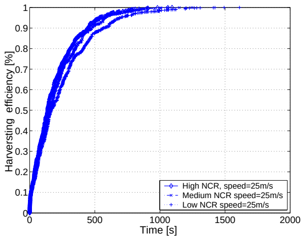
In Fig. 8, we made the same evaluation as in Fig. 7 but on the map topology. We can see a similar effect on the influence of the NCR on the harvesting rate. For example, on Fig. 8(a), after $100s$, Mobeyes has harvested 95% of the events for the high NCR, but only 90% and 70% for the medium and low NCR respectively. We also see similar effects for all different average speeds, although the effect of different velocities is more drastic on the map topology.

Then, an interesting feature may be observed by looking at Fig. 7 and Fig. 8. If we compare the two figures, we can see that the performance gap between different classes of NCRs is reduced as the speed increases. This is even more pronounced on the figures for the map topology. This may be intuitively explained by the fact that an increased speed reduces the average spatial and temporal dependencies created by the different topologies, as for example, faster nodes are less subject to the effect of shortest paths on a map topology, as they reach their targets quicker and then move again on a different path.

Finally, in Fig. 9, we perform a cross comparison between different topologies and different mobility models. For this set of results, we simulated Mobeyes with the Track model for the triangle topology and map topology. We added a simulation with the Random Waypoint model for the triangle topology. For the three scenarios, we have the same density and same NCR. What we can see on both figures is that with similar NCR, speed and density, the performance of data dissemination is almost identical. For example, in Fig. 9(b),



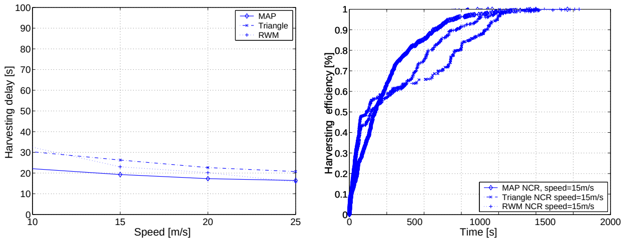
(a) Average velocity= 5m/s (b) Average velocity= 15m/s



(c) Average velocity= 25m/s

Fig. 8. Faction of harvested summaries on the map topology

although the harvesting slightly differs during the simulation, it is completed at roughly the same time. And in Fig. 9(a), the latency is also similar for all three cases. This shows the significance of NCR, as it is able to characterize the intrinsic properties of complex topologies or mobility patterns. Coupled with the average speed and the average density, the NCR describes all types of motion or topologies.



(a) Delay (b) Harvest Efficiency

Fig. 9. Comparison of Data Dissemination efficiency under different motion patterns and topologies for similar NCRs

VI. CONCLUSION

In this paper, we introduced the Neighborhood Changing Rate (NCR), a new parameter describing data dissemination in mobile ad hoc networks. We illustrated how the parameter was able to describe temporal and spatial dependencies that cannot be controlled solely by the average speed or the average density. Those dependencies are usually controlled by a multitude of other parameters that are difficult to tune for efficient performance evaluation. Instead, we proposed to group them into a single parameter, NCR. By its intrinsic property, the NCR measure can well characterize a dissemination/harvesting scenario and allow to predict the performance of such scenario. The NCR measure also allows us to evaluate cross-mobility and cross-topology data dissemination which can be controlled solely by three parameters: the speed, the density and NCR. Specific topologies, or mobility patterns become irrelevant.

In the future work, we intend to evaluate the effect of each parameter to see which one has the most significant effect on data dissemination.

REFERENCES

[1] M. Grossglauser, and M. Vetterli, "Locating Mobile Nodes With EASE: Learning Efficient Routes From Encouter Histories Alone", in *IEEE/ACM Transaction on Network*, pp. 457-469, Vol. 14, NO. 3, June 2006.

[2] W. Zhao, M. Ammar. and E. Zegura, "A Message Ferrying Approach for Data Delivery in Sparse Networks", in *Proc of ACM MobiHoc'04*, Tokyo, Japan, May 2004.

[3] Uichin Lee, Eugenio Magistretti, Biao Zhou, Mario Gerla, Paolo Bellavista, Antonio Corradi, "Efficient Data Harvesting in Mobile Sensor Platforms", in *Second IEEE International Workshop on Sensor Networks and Systems for Pervasive Computing (PerSeNS? 2006)*, Pisa, Italy, March 2006.

[4] J. Haerri, F. Filali, and C. Bonnet, "Performance comparison of AODV and OLSR in VANETs urban environments under realistic mobility patterns" in Proc. of the 5th IFIP Mediterranean Ad-Hoc Networking Workshop(of (Med-Hoc-Net'06), Lipari, Italy, June 2006.

[5] The Random Trip Framework, `http://lrcwww.epfl.ch/RandomTrip/`

[6] Biao Zhou, Kaixin Xu, Mario Gerla, "Group and Swarm Mobility Models for Ad Hoc Network Scenarios Using Virtual Tracks", in *Proc. of IEEE Military Communications Conference (MILCOM '04)*, Monterey, USA, October 2004.