Malicious Communication in VANETs — Lessons Learned in Microsimulation of Traffic

Andrew Koster[1], **Luiz H. Dias Souza**[1], **Ana L. C. Bazzan**[1]

[1]Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brazil

Abstract. *In this paper, we analyse the effect of malicious communication in a Vehicular Ad hoc Network (VANET) on the throughput of a traffic network. VANETs have been proposed as a communication solution for car to car communication. Nevertheless, there are no guarantees that vehicles will communicate truthfully, particularly if an individual car could gain from lying. We analyse such malicious communication by simulating a road network in SUMO, a microscopic traffic simulator, and comparing the flow through the network under varying amounts of truthful and malicious agents. Furthermore, we analyse the use of a computational trust model, to potentially counteract any adverse effects of the malicious agents. We found that malicious communication actually has a beneficial effect on the flow through the system, by causing traffic to be better distributed among the possible routes. Using trust further increased performance.*

1. Introduction

Vehicular Ad hoc Networks (VANETs) allow for unprecedented amounts of information to be sent between vehicles. Current applications of VANETs focus on safety mechanisms, such as early warning systems for rear-end collisions, e.g. the Cooperative Forward Collision Warnings (C2C-CC, 2007). However, potential for the use of car to car (C2C) communication goes far beyond safety mechanisms, and VANETs provide a potential infrastructure for this.

Nevertheless, malicious agents may be able to exploit the C2C communication to advance their own goals, at the cost of diminishing performance of other traffic participants. For instance, when using C2C communication to predict upcoming congestion, it may be worth an agent's while to lie about the state of the road, in order to deter more cars from choosing the same route, and thus optimizing his own travel time; potentially at the cost of other road users' travel time, who choose a longer, slower, suboptimal route due to this malicious communication.

Kraus et al. (2008) analysed a traffic scenario in which deceitful agents can exploit C2C communication in order to optimize their own travel time. Their research indicates that the effect of deception on travel times is low, however the simulation is for a downtown environment with all cars moving in different directions. Koster et al. (2013) analyse a different scenario: they use a macroscopic simulation of a highway scenario, and their results indicate that malicious communication can significantly impact travel times. In the same paper they propose the use of trust to mitigate the impact of this malicious communication.

For this paper we reimplemented Koster et al.'s work, using the SUMO microscopic simulator (Behrisch et al., 2011) in order to discover what impact malicious communication has, when using a more realistic simulation. We further analyse the behaviour of individual cars, and the network as a whole, in the presence of malicious communication, as well as the possible use of the proposed trust model to counteract the negative effects of malicious agents.

The experimental scenario is similar to the one used by Koster et al. The scenario represents a highway with a single alternative route. The highway, or main road, is both shorter, and faster than the alternative, but if it becomes congested then the alternative route is a good choice. Koster et al. perform experiments using a macroscopic traffic simulation. However, simulation at this level has a number of problems. A macroscopic traffic model is a mathematical model that employs traffic flow rate variables (Transportation Research Council, 2010); it uses numerical methods to calculate properties of the network under given circumstances. Cars are not represented individually, and thus the VANET cannot be simulated. Instead, the communication is assumed to happen at specific locations in the network and all cars can communicate with each other.

To more accurately simulate the way communication over the VANET influences the traffic, we use a microscopic traffic model. In a microscopic traffic model, each car is modelled individually, with its own route, and individual movement through the network. We use the SUMO simulator (Behrisch et al., 2011), an open source microscopic road traffic simulation package. SUMO explicitly includes the ability for online interaction with the model through TraCI, the Traffic Control Interface. This allows the simulation of traffic, and the simulation of communication (and vehicles' changes in behaviour based upon that communication) to be interleaved.

In Section 2 we briefly summarize related work; in particular Koster et al.'s work. In Section 3 we describe the experimental setup and in Section 4 we discuss our experiments and the results, before concluding the paper in Section 5.

2. Background

The problem of detecting and preventing malicious behaviour in VANETs has been studied from a number of perspectives, with various solutions being proposed. The most common attacks that are studied are those at the network level, where encryption and specific protocols can prevent malicious behaviour (Raya et al., 2006). However, these solutions cannot prevent legitimate participants from sending false information. One of the first attempts to detect such false information was made by Golle et al. (2004). Their system relies on sensor information and knowledge about the network topology to decide whether or not the information can be true, allowing them to distinguish truthful nodes from those spreading false information, and by using a cryptographic solution, identities are preserved. However, it remains to be seen whether such an approach can be used in practice: VANETs are unreliable networks with a strict limitation on the available bandwidth (Scheuermann et al., 2009), and using a significant portion of that to exchange information about the network topology may be impractical. Additionally, with thousands, if not millions, of cars in a network, storing and tracking identities itself becomes problematic.

Other approaches deal with malicious information in C2C communication by using computational trust models. Many computational trust models have been proposed for a variety of domains (Pinyol and Sabater-Mir, In Press). As Zhang (2011) points out, however, C2C communication faces problems that are not addressed in such models. Most conventional trust models rely on repeat interactions with an individual agent to build up a trust relationship. In the absence of such repeated interactions, they turn to their peers, who may have repeated interactions, or a centralized authority that holds reputation information. Unfortunately, in communication in a VANET, none of these methods are available. Koster et al. (2013) resolve this by considering the VANET in its entirety as trustworthy, or not.

The main building blocks of Koster et al.'s system are (1) a computational trust model that is able to assess the trustworthiness of various information sources, in particular C2C communication, and (2) a possibilistic BDI agent, to reason about uncertain information and make decisions. We will not go into details of the possibilistic reasoning engine, but the main purpose of the system is to incorporate information from various different information sources and make an informed decision in order to obtain some goal; usually arriving at the destination as fast as possible. Some background knowledge about the functioning of their trust model, however, is necessary to understand the experiments of Section 4.

Koster et al. use the truthfulness of past messages to calculate the trustworthiness of the VANET. This trust evaluation is a numerical evaluation in the range $[0, 1]$, which can be interpreted as the likelihood that a message received over the VANET will be truthful. However, the communication received over the VANET will consist of multiple, potentially conflicting messages. If we assume a message communicates a Boolean statement, then different vehicles can communicate its assertion and its negation. We thus assume the communication itself gives a likelihood over the truthfulness of the message content. The frequency of any literal φ, as given by C2C communication, is:

$$F_{C2C}(\varphi) = \frac{|M(\varphi)|}{|M(\varphi)| + |M(\neg\varphi)|} \tag{1}$$

where $M(\varphi)$ is the set of messages received that contain φ.

After evaluating the truthfulness of a statement φ, by evaluating sensor data, such as the odometry of the car, or by obtaining information afterwards; for instance, from a trusted, centralized, database when the car is in its home garage, the same messages can be used to evaluate the trustworthiness of the network. Koster et al. do not differentiate between reasons for a message being wrong, but simply evaluate all erroneous communication as untrustworthy. The trust model used is based on BRS Jøsang and Ismail (2002), and it allows for the discounting of older information, thereby accounting for the dynamicity of a VANET. Let M^+ be the set of truthful messages and M^- be the set of false messages. Furthermore, let $time(m)$ be the time a message was received and now be the current time. The trustworthiness of the VANET is computed as:

$$Trust(C2C) = \frac{D_1 + \sum_{m \in M^+} \lambda^{now - time(m)}}{D_2 + \sum_{m \in M^+ \cup M^-} \lambda^{now - time(m)}} \tag{2}$$

Where D_1 and D_2 together allow for the specification of a default value for trustworthiness and $0 < \lambda \leq 1$ is the discount factor over time.

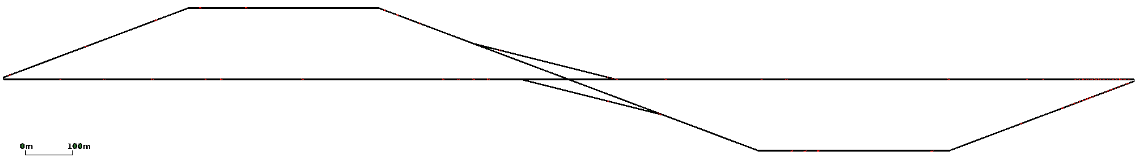

Figure 1. The road network in SUMO

Finally, to evaluate the likelihood of a Boolean statement being true, based on communication over the VANET, the frequency in the communication of Eq. (1) and the trustworthiness of the VANET, as in Eq. (2) are combined:

$$T_{C2C}(\varphi) = Trust(C2C) \cdot F_{C2C}(\varphi) + (1 - Trust(C2C)) \cdot F_{C2C}(\neg\varphi) \qquad (3)$$

Fundamentally, this relies on the assumption that trust can be interpreted as a likelihood: if a vehicle communicates φ, over the VANET with trustworthiness t, it can also be seen as communicating $\neg\varphi$ with trustworthiness $1 - t$.

In Section 4 we will analyse the functioning of this trust model in a scenario with communication over a VANET, but first we describe the experimental setup.

3. Experimental Setup

As mentioned in the introduction, we use the SUMO microsimulator to model the traffic scenario. In this section we describe the specifics of the network that is modelled, and how the cars can communicate and reason about this communication.

3.1. The network

We modelled the network to have similar properties to the network in Koster et al.'s original work. The alternative roads are slightly longer than the main road (2.5 km, compared to the main road being 2.3 km long), as well as having a lower maximum speed (60 km/h or 16.7 m/s, while on the main road it is 100 km/h, or 27.8 m/s). This causes the minimum travel time along the side road (taking both alternative routes) to be 149 seconds, while along the main road the minimum travel time is 86 seconds. The network, as modelled in SUMO, is depicted in Figure 1.

If there is no congestion, vehicles will preferably travel along the main road. In order for this to generate congestion, we must limit the outflow of the network: in SUMO the destination nodes, where cars exit the network, have infinite capacity. The maximum speed along the road is constant, and SUMO does not allow more cars to enter the system than physically fit on the road. Therefore by simply adding more cars we can never create congestion. Rather, we must artificially create a bottleneck at the exit points, where we lower the speed in order to cap the flow of the network to be a realistic simulation of a road network. We cap the flow at 0.48 cars per second, if cars travel along both roads (the main road and the alternative), by setting a speed limit on the final stretch of road to 1.2 m/s. This means that, if 1 car enters the system every 3 seconds, and the cars are distributed optimally, the cars will be able to flow through the system. However, using just the main road, the inflow will be greater than the outflow, causing the simulator to "drop cars", which cannot be inserted due to the road being full.

In practice, due to SUMO leaving more space between cars than the absolute minimum required by the configuration, the flow is capped at 0.3: there are at least eight meters between two cars, rather than the minimum of five of the ideal network. The way we evaluate network performance is by calculating the flow of the entire network.

If all cars try to drive along the main road, the flow of the network converges on 0.15, because the side road is not used, effectively making the system a single road with a bottleneck at 1.2 m/s. The performance increase of the network by sending cars along the side road can thus be seen as the percentile increase over this baseline, with 100% being the maximal flow of 0.3 cars/second.

3.2. Communication and reasoning

Movement within the network is interleaved with communication and reasoning. Each vehicle has two choices to make: at the start, whether to take the main road or the side road, and in the middle, when it can switch between roads. A configurable percentage of the cars has accurate knowledge about the state of the roads: it knows the average speed on each part, and chooses its route according to this. The rest of the cars must rely on C2C communication. The cars communicate about a single Boolean statement: "the main route is congested". In a truthful scenario, a car travelling along the main route communicates that it is congested, if it is moving at less than half than the maximum speed of the road, otherwise it communicates that the road is not congested. A car travelling along the side road does not have that information available, but it does have a belief about the state of the main road: it chose to drive along the side road because it believed the main road was congested when it had to make the choice. We allow the confidence in such beliefs to degrade over time, due to the dynamicity of the environment. If the confidence is greater than some threshold, a car on the side road will communicate that the main road is congested. Otherwise it will not communicate anything.

A car, having to make a choice thus receives messages and aggregates them using the formulas of Section 2. In the absence of prior communication (at the start), the trust evaluation is simply the default trust, which we set at 1, and the trustworthiness of the statement "the main route is congested" as given by C2C communication is simply the maximum likelihood estimation given by the individual messages. However, the car has a preference to travel along the main road if it is not congested. The incorporation of new information into the agent's reasoning system is not straightforward. Koster et al. (2013) propose to use a possibilistic BDI agent, and use belief revision to resolve this issue. However, in our simple system, this is not necessary and we simply assume that a car has an inherent strength, determined a priori, to its desire to travel along the main road. If the new information that the main route is congested is more trustworthy than this inherent strength, then it will travel along the side road instead. The strength of the desire thus functions as a threshold for when the information is accepted.

At the second choice point, the car makes a similar evaluation, but this time it can evaluate the trustworthiness of the VANET. If it was travelling along the main road, and its average travel time was under half of the maximum speed, it believes the main road was congested. As a consequence, any message stating the reverse is marked as being false. When travelling along the side road, the actual state of the main road cannot be evaluated, but if the side road is congested, then we assume the car would have been better off travelling along the main road instead. Therefore, if the average speed of the car along the side road was less than half of the maximum speed (along the side road), it believes that the main road was not congested (or at least less congested), and therefore any message stating that the main road was congested is marked as being false. This gives the agent the sets M^+ and M^- of Equation (2).

Malicious agents, rather than communicating truthfully, as we described above, always say what suits them best, which may be the truth, or it may be a lie. If they are travelling along the main road, they are best suited by keeping traffic away from it, and thus always communicate that the main road is congested. If they are travelling along the side road, the reverse is true, and thus they always communicate that the main road is not congested.

As some preliminary experiments, we test the setup with cars that all have accurate information about the road network (they know the average speed on each road). Unless mentioned otherwise, we run the simulation for 12,000 timesteps (seconds), because this is generally enough to reach an equilibrium. The flow of the simulated network is simply the total number of cars passing through the network, divided by the time it takes for them to pass through (for which we disregard the time it takes to reach a stable equilibrium). If congestion reaches the start of the simulator, less cars are allowed to enter the network, which decreases the flow. The score of a simulation is thus:

$$\left(\frac{cars(end\_time) - cars(equilibrium\_start\_time)}{end\_time - equilibrium\_start\_time} - 0.15\right)/0.15 \tag{4}$$

Which is the percentage of flow increase over the basic scenario where only the main road is used, as compared to the ideal scenario where both roads are used optimally.

If all cars have accurate information about the road network, this results in a performance of 83%. When analysing the network usage we see that there are large oscillations, because when a road has been congested for a long time, all the cars are moving at 1.2 m/s, whereas if a road is recently congested, the average speed will be higher. This means that a half-empty road with cars travelling at this minimum speed may still not be prioritized over an almost completely congested road, because cars moving onto that road will still be travelling fast, thus increasing the average speed, and causing cars behind them to also prioritize the almost entirely congested road over the almost empty road. This seems to indicate that pure average travel speed on a road is not good enough for cars to choose between two roads. However, this is not the main part of our research, it just gives us a good benchmark for results with communication against which to measure performance.

4. Experimental Results

Using the model as described in the previous section, we run a number of simulation experiments, testing the effect of communication over a VANET, the effect of malicious agents, and the trust model of Section 2.

4.1. Truthful communication

First, we establish a baseline. In order to evaluate the effect of malicious communication, we need to know how truthful communication affects the network. We assume that with

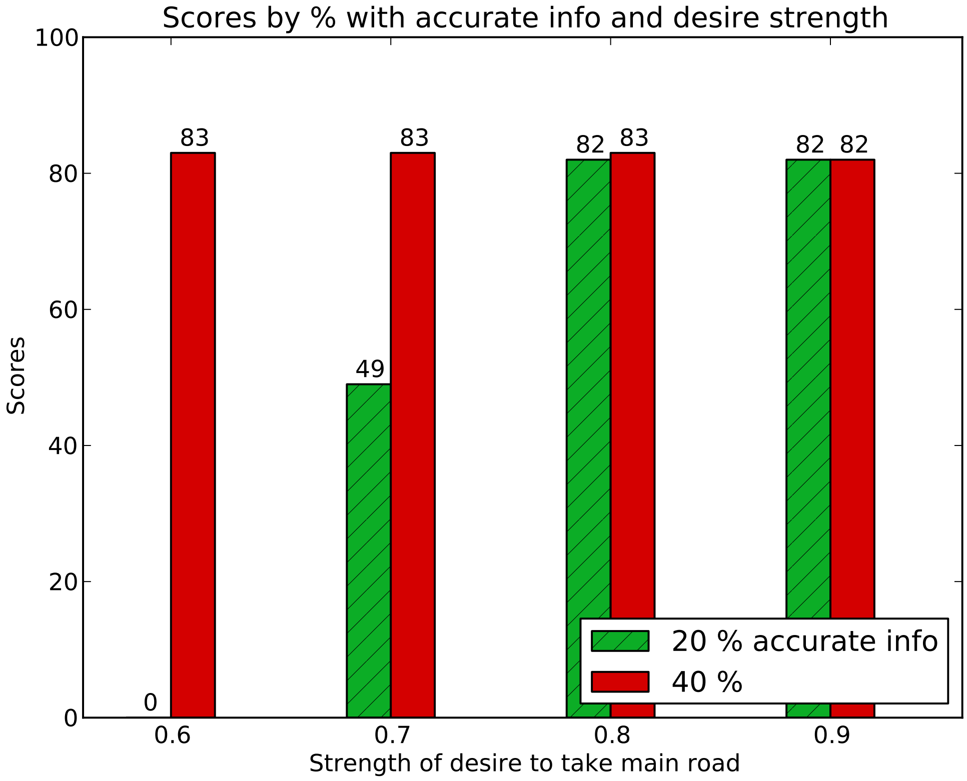

Figure 2. Network performance without liars

no communication and no other information, cars will simply choose to drive along the main road, resulting in the baseline flow of 0.15 cars/second. In the previous section we showed that if all cars know the average speed on all roads, performance increases to a flow of 0.275, or 83%. We run a number of experiments with varying numbers of vehicles with this accurate information, and assume all other cars rely on communication over the VANET, as described in the previous section. Simultaneously, we vary the a priori strength of the desire to drive along the main road. Because we assume all cars communicate truthfully, we do not use the trust model.

The results can be found in Figure 2. We have omitted from the graph the results for when no cars have prior information, because this results in no improvement over the baseline at all, just as the situation with a low threshold and 20% of the cars having accurate information about the network. This is because if nobody is travelling along the main road, the only communication over the VANET comes from cars that are travelling along the side road. These, falsely, believe the main road is congested, and propagate this false belief. Eck and Soh (2013) call this the *institutional memory phenomenon*, which occurs in large team information sharing. Sharing information in traffic can be seen as a special instance of this. Eck and Soh point out that using trust does not solve the issue, but in our scenario we observe differently, as discussed below. Without trust we see a number of principal approaches that can resolve the issue.

- By setting the strength of the desire very high, we create a high innate resistance to travelling along the side road, creating a balance of cars that take the main road and obtaining the same result as the benchmark score of 83%. However, as can be seen in the situation with 20% of the cars having accurate information (and thus choosing the main road), we still need a fairly high threshold to prevent vehicles from taking the wrong route. This is because the cars on the side road keep propagating the outdated belief through the network.
- By using meta-information about the communication: at the moment we only take the communicated information into account, not the number of agents that are communicating that information. For instance, if there are a very low number of cars reporting congestion, it is possible that this is a phantom belief, whereas if a large number of cars report congestion, it is more likely that those cars are actually stuck in congestion. Similarly, in some VANET communication protocols (Altayeb and Mahgoub, 2013), information such as location, direction and speed are communicated in the network layer: this information can be extracted and used in the decision process.
- Have beliefs degrade at a faster pace: as described in Section 3.2, agents deal with the dynamic environment by having the strength of beliefs decay over time. We define a threshold, under which a belief is no longer communicated. By making this threshold very high, cars on the side road stop communicating about the state of the main road very quickly.

Which approach is best is highly dependent on the real road conditions. Regardless, these results seems to indicate that even if there is some method for keeping malicious agents out of the system some further mechanism is needed to prevent such "phantom" beliefs from propagating. This has been studied in other domains, such as sensor networks, or other collaborative sensing scenarios (Glinton et al., 2010), and solutions from these domains might be applicable to C2C communication in VANETs as well.

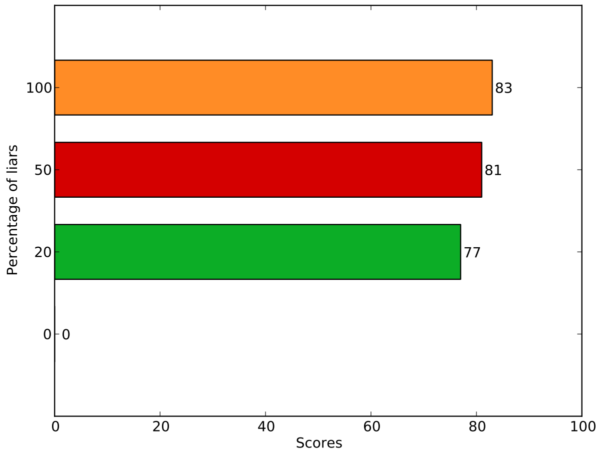

Figure 3. Network performance with liars

4.2. Malicious communication

To measure what effect malicious agents have upon the system, we add varying percentages of them to the model. The results are plotted in Figure 3. We use a strength of the desire to take the main road of 0.6, because we want the agents to use the communicated information. The lying agents will add noise to the system, and thus choosing an extremely high value for this belief will result in no agent taking the side road, regardless of the communication. How to calibrate the balance of a priori knowledge and untrustworthy communication over the VANET, however, remains an interesting avenue for future research.

As we can see, contrary to what Koster et al. (2013) found, malicious behaviour does not seem to have a detrimental effect on the network as a whole. The reason for this is that lies cause the cars to spread out more evenly over the network. Even at relatively low numbers of lying agents, this effect is clearly visible. The reason that having 20% malicious communicators in the system performs so much better than having 0% is because malicious agents always communicate, regardless of what their beliefs are. Thus malicious agents on the side road will be communicating that the main road is not congested. This ensures that cars, malicious and truthful will continue to choose the main road, keeping the “phantom” belief, that we saw appear with only truthful agents, from being propagated with too high a trustworthiness.

Malicious communicators also do not obtain, on average, a greater individual performance, because they are equally susceptible to other vehicles’ lies. In order to obtain a greater performance, it is necessary to influence cars ahead of the vehicle, in order to have a clear road upon arrival, but this effect is not achieved, in particular because the agent does not know which road it should take ahead of time.

4.3. Using trust

Although “malicious” communication is actually beneficial to overall network performance, and indifferent at individual level, we are still interested in what the trust model can do. We have plotted our results in Figure 4.

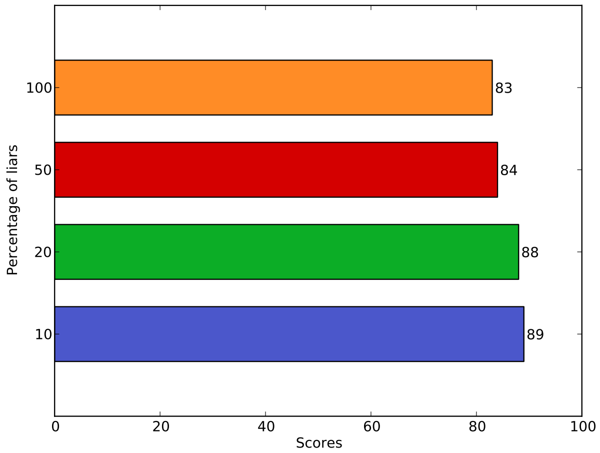

Figure 4. Network performance with liars and trust

As we can see, with a small number of lying vehicles and the use of trust, the network performance actually increases over that when cars can receive accurate information about the average speed on the roads. While trust does not come into play in the first choice, we hypothesize that at the second choice point the trust mechanism allows the cars to better choose the optimal road, thereby distributing themselves better over the second part of the network, which in turn frees up space on the second part. However, further analysis of the dynamics of the network are necessary to confirm this.

5. Conclusion and Future Work

In general, the results we found in the previous section seem more similar to those found by Kraus et al. (2008) than to the results Koster et al. (2013) obtained, although our results do show that using trust can increase network performance quite significantly. This could be, because, just as in Kraus et al.'s simulation, the lies do not advance ahead of the malicious agent, whereas this was the case in Koster et al.'s macroscopic simulation: by communicating all at once before moving, the communication automatically preceded movement. One effect we found that is new is the significant performance increase with malicious vehicles. While we can explain this, it is a counter-intuitive result, and we will need to study this further in different network topologies to see whether this type of communication can be used in general to prevent congestion when there are uncongested side roads.

The results we obtained open many avenues for future research. Firstly we aim to validate our hypothesis for why using trust increases performance of the network, and whether this is a generally applicable principle, or whether this effect is limited to the kind of network topology we simulated here. We also aim to discover whether it is possible for malicious agents to have their misinformation advance ahead of them, and what effect this has on the network. In particular, our experiments show that the effect of communication, malicious or otherwise, is not straightforward. Phantom beliefs can be propagated, and apparent malicious communication can increase network performance. Further study of such effects is necessary if C2C communication is to be understood properly before it is deployed on a mass scale.

Acknowledgements

Andrew Koster is supported by CAPES (PNPD). Ana Bazzan is partially supported by CNPq. Luiz Souza receives a grant in the SIMTUR project from the CTIC/RNP.

References

Altayeb, M. and Mahgoub, I. (2013). A survey of vehicular ad hoc networks routing protocols. *International Journal of Innovation and Applied Studies*, 3(3):829–846.

Behrisch, M., Bieker, L., Erdmann, J., and Krajzewicz, D. (2011). SUMO - Simulation of Urban MObility: An overview. In *Proceedings of SIMUL 2011, The Third International Conference on Advances in System Simulation*, pages 63–68, Barcelona, Spain.

C2C-CC (2007). Car 2 car communication consortium manifesto.

Eck, A. and Soh, L.-K. (2013). Dynamic facts in large team information sharing (extended abstract). In *AAMAS'13*, pages 1217–1218, Saint Paul, MN, USA.

Glinton, R., Scerri, P., and Sycara, K. (2010). Exploiting scale invariant dynamics for efficient information propagation in large teams. In *Proceedings of AAMAS 2010*, pages 21 – 28, Toronto, Canada. IFAAMAS.

Golle, P., Greene, D., and Staddon, J. (2004). Detecting and correcting malicious data in vanets. In *Proceedings of VANET'04*, pages 29–37, Philadelphia, USA. ACM.

Jøsang, A. and Ismail, R. (2002). The beta reputation system. In *Proceedings of the Fifteenth Bled Electronic Commerce Conference e-Reality: Constructing the e-Economy*, Bled, Slovenia.

Koster, A., Tettamanzi, A., Bazzan, A. L. C., and da Costa Pereira, C. (Forthcoming 2013). Using trust and possibilistic reasoning to deal with untrustworthy communication in VANETs. In *Proceedings of the IEEE Conference on Intelligent Transport Systems*, The Hague, The Netherlands.

Kraus, S., Lin, R., and Shavitt, Y. (2008). On self-interested agents in vehicular networks with car-to-car gossiping. *IEEE Transactions on Vehicular Technology*, 57(6):3319–3332.

Pinyol, I. and Sabater-Mir, J. (In Press). Computational trust and reputation models for open multi-agent systems: a review. *Artificial Intelligence Review*.

Raya, M., Papadimitratos, P., and Hubaux, J.-P. (2006). Securing vehicular communications. *IEEE Wireless Communications*, 13(5):8 – 15.

Scheuermann, B., Lochert, C., Rybicki, J., and Mauve, M. (2009). A fundamental scalability criterion for data aggregation in VANETs. In *MobiCom 2009: Proceedings of the Fifteenth ACM SIGMOBILE International Conference on Mobile Computing and Networking*, pages 285–296, Beijing, China.

Transportation Research Council (2010). *Highway Capacity Manual 2010*.

Zhang, J. (2011). A survey on trust management for VANETs. In *Proceedings of IEEE AINA'11*, pages 105–112, Biopolis, Singapore.