Simulation of Malware Propagation and Effects in Connected and Autonomous Vehicles

Jonathan Francis Roscoe
Emerging Security Paradigms
BT Applied Research
jonathan.roscoe@bt.com

Oliver Baxandall
Emerging Security Paradigms
BT Applied Research
oliver.baxandall@bt.com

Robert Hercock
Emerging Security Paradigms
BT Applied Research
robert.hercock@bt.com

***Abstract*—Connected and autonomous vehicles (CAVs) are an emerging technology that will introduce new threats to the general public. Impending standards (such as ISO21434) demonstrate that there is a real cyber security risk and a need for supporting infrastructure in the form of vehicle security operations centre.**

In this concept paper we discuss some of the issues facing vehicle security as the technology matures over the next few years and look at how epidemiological models for malware might be developed to address concerns over vehicle cyber threats.

We detail our development of Mobius, a bespoke tool for simulating and analysing malware events in CAVs and explore how the technology might be applied to support real-world decision making.

As a part of the need for cyber resilience, we suggest there is a key role for vehicle simulation software capable of modelling cyber threats to assist with threat analysis and decision making for highway authorities, OEMs and fleet operators, amongst others. We present a summary of compartmental epidemiological models and the role they can play in understanding malware propagation for CAVs.

***Index Terms*—autonomous, geospatial, simulation, transport, vehicle, security, malware, propagation**

I. Background

The automotive market today is changing rapidly with both vehicles and infrastructure becoming more connected. New models arriving on the market are exchanging an increasing volume of data which is being shared with other vehicles, infrastructure and increasingly to the cloud and with technologies at its ‘edge’ which are providing insight, analytics and control mechanisms which did not exist before. Beyond these connected capabilities there are the less understood autonomous technologies which will introduce even greater data sharing needs.

The de facto standards for describing vehicle automation were proposed in a recommendation document by SAE. In the model, levels 1 to 3 describe driver assistance whilst levels 4 and 5 describe high or full operational autonomy . As these levels are realised, the requirements for cyber security become more significant. Global uptake of L3-L5 vehicles is expected to reach 40% by 2030, and 80% by 2035 [1] and the Zenzic roadmap anticipates cyber security becoming a core part of vehicle approval in 2030 [2].

A number of cyber attacks against modern vehicles with varying levels of autonomy have been identified in the literature, against features including tyre-pressure sensors [3], LIDAR ranging [4], wireless entry [5] and infotainment systems [6]. These attacks can have superficial consequences such as the ability to display messages, may compromise security of the vehicle and bypass locking mechanisms or compromise safety by endangering occupants and other road users. These attacks may come from nation states, criminal organisations, hacktivists/artists and fraudulent operators [7].

A significant quantity of data is generated during routine CAV operation including sensor data, bus messaging, third-party communications and observations. There are many parties with an interest in this data ranging from manufacturers, local authorities, law enforcement and commercial entities for their own applications. Vehicle security operations centres (VSOCs) will be vital in the future for aggregating and analysing the vast amount of data.

The use of rule-based and other more intelligent mechanisms for automatically classifying data is a necessary tool but can also lead to alert fatigue. Existing automatic techniques may also struggle to recognise threats not previously identified. Consequently, the continued development of artificial intelligence and machine learning techniques to analyse data is a key requirement of SOCs and an approach to anomaly-based threat detection is required.

A. Malware Simulation for CAVs

Infectious disease modelling in the literature can be traced back to Bernoulli in 1760 [10] leading to a modern significant understanding with models by Kermack and McKendrick in the 1930s [11], [12] and a significant body of established research by the 1970s [11].

There are a number of epidemiological models for modelling the propagation of malware, one of the most significant being the 1991 Kephart-White models [13] which simplify individuals in a network to a system composing of a number of discrete states such as *susceptible*, *infected* and *recovered*. These are normally referred to as compartmental models, one of the simplest and well-known is the SIR model. The SIR model assumes an immunity once an individual has recovered, which is why modifications such as the SIS and SEIR models

This work was initially carried out as part of ResiCAV, an InnovateUK project funded by the Centre for Connected & Autonomous Vehicles.

may be favoured. An overview of various models is shown in Figure 1.

In all of these compartmental models, the transition rate is a key aspect, that determines the rate at which individuals move between compartments. These transition rates allow tuning the model for different virus characteristics. The distinction between compartments and their transition states is important depending on the nature of the disease (in either the biological or cyber domain), for example, an exposed individual not yet infectious may be an undetected incubator.

For the SIR model, the transition rate between susceptible and infected (βSI) is defined as:

$$\frac{d(\frac{S}{N})}{dt} = -\beta \frac{SI}{N^2} \tag{1}$$

where S is the susceptible population, I is the infected population and N is the sum of all three. β is average contacts per individual multiplied by probability of transmission and $\frac{SI}{N^2}$ is the fraction of contacts resulting in a susceptible individual becoming infected.

Between infected and recovered, the transition rate is γI, simply a proportion of the infectious population (I). This maybe adapted for a time period t with:

$$\gamma = \frac{1}{t}. \tag{2}$$

For malware with a human interaction element, the Maki-Thompson rumour model has also been used [14]. These and related models have been successfully applied to large computer networks [15]–[17].

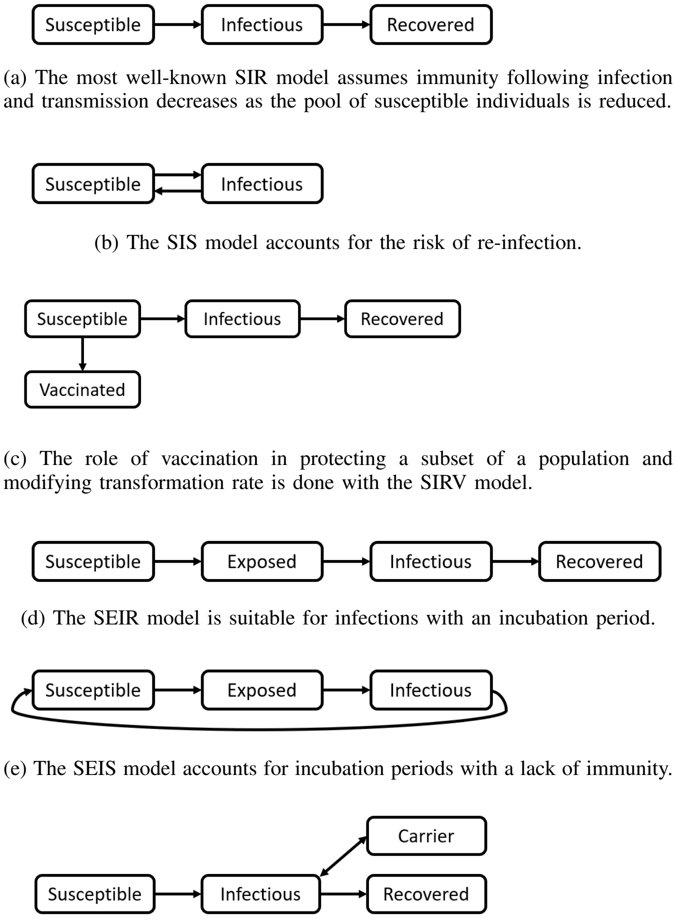
Mobile systems present new challenges to modelling malware propagation due to their transient nature and diversity of communication channels. Mickens and Noble [18] highlighted that the Kephart-White approach is insufficient for modelling propagation in mobile environments. This is because propagation relies on a statistic of average connectivity for each node. Mickens and Noble propose a new approach known as probabilistic queuing to account for node velocities and non-homogeneous connectivity.

II. Mobius: Traffic Simulation Dashboard

Mobius (shown in Figure II) is a combination of traffic simulation and web-based geospatial analytics. It was developed with the aim of enabling anomaly detection in a population of mixed manual, semi-autonomous and autonomous vehicles. The aim is to assist in identifying vehicles that may have become compromised as well as supporting planning and decision making processes.

A. Simulation of Urban MObility (SUMO)

SUMO [20] is an EPL[1] licensed, open-source, microscopic road traffic simulation package. It is the foundation upon which the Mobius back-end is built. Within SUMO, is a traffic control interface (TraCI), which provides a client/server architecture for accessing SUMO. This allows predefined alterations to the simulation, retrieval of values or the running of custom code during the process. After installation of SUMO, a scenario is generated for a predefined area, which converts a geographical topology of available traffic routes into graph's stored as XML, that vehicles are simulated moving in. The regions analysed were Central London, Manchester, Oxford and Martlesham. Supported vehicle types varied by region as appropriate, modelling ships, cars, buses, trucks, motorcycles, passenger vehicles, bicycles and pedestrians.



(a) The most well-known SIR model assumes immunity following infection and transmission decreases as the pool of susceptible individuals is reduced.

(b) The SIS model accounts for the risk of re-infection.

(c) The role of vaccination in protecting a subset of a population and modifying transformation rate is done with the SIRV model.

(d) The SEIR model is suitable for infections with an incubation period.

(e) The SEIS model accounts for incubation periods with a lack of immunity.

(f) The carrier model considers disease that may lie dormant.

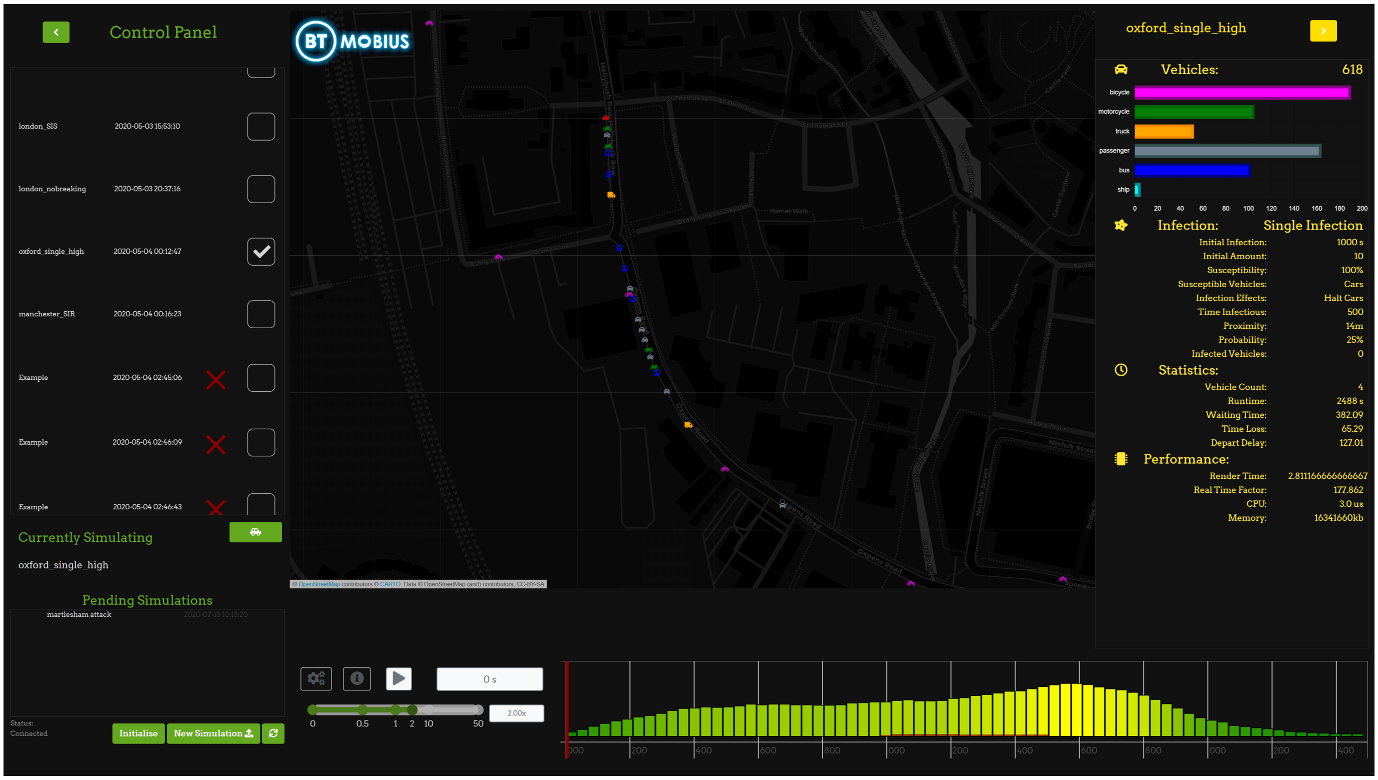
Fig. 1. Individuals are placed in one of a number of compartments, transition rates define how individuals move between these and model propagation of disease. Different compartmental models are used depending on the nature of disease. Not all of these models may be directly applied to the cyber domain

B. Visualisation

The front-end of Mobius is entirely web-based and uses common technology rendering libraries using Node.js and Express such as leaflet (for maps), bootstrap (a framework for HTML, CSS and JavaScript elements) and Vis.js (for timeline dynamics).

The core visual component is a geographical map, with markers for each vehicle in the simulation. Vehicles size, shape and colour can be altered depending on the type of simulation

[1] https://www.eclipse.org/legal/epl-v20.html



london_SIS 2020-05-03 15:53:10 □

london_nobreaking 2020-05-03 20:37:16 □

oxford_single_high 2020-05-04 00:12:47 ☑

manchester_SIR 2020-05-04 00:16:23 □

Example 2020-05-04 02:45:06 ☒ □

Example 2020-05-04 02:46:09 ☒ □

Example 2020-05-04 02:46:43 ☒ □

Currently Simulating

oxford_single_high

Pending Simulations

martlesham attack 2020-07-13 10:13:20

Status:
Connected

© OpenStreetMap contributors © CARTO, Data © OpenStreetMap (and) contributors, CC-BY-SA

oxford_single_high

Vehicles: 618

Infection: Single Infection

Initial Infection:	1000 s
Initial Amount:	10
Susceptibility:	100%
Susceptible Vehicles:	Cars
Infection Effects:	Halt Cars
Time Infectious:	500
Proximity:	14m
Probability:	25%
Infected Vehicles:	0

Statistics:

Vehicle Count:	4
Runtime:	2488 s
Waiting Time:	382.09
Time Loss:	65.29
Depart Delay:	127.01

Performance:

Render Time:	2.8111666666666667
Real Time Factor:	177.862
CPU:	3.0 us
Memory:	16341660kb

being displayed, for example, allow for a distinction to be made between infected vehicles.

C. Simulation Management

Simulations are managed through and SQLite database, which stores the simulations initial parameters and basic information such as the running length when simulated. The simulation itself is stored as a JSON file, with an id that links itself to the simulation in the database. The file is organised as GeoJson by vehicle/simulated entity, with details on the type of vehicle, its speed and position at fixed points of time. The creation of new simulations is controlled using an HTML form, containing parameters for the simulation. Node.js spawns a child process, which from the a virtual command line runs a python file for creation of the simulation given the specified parameters. The simulation is stored in the database and marked as pending in the front-end, until the child process finishes, at which point the database is updated with the new simulation for visualisation purposes.

D. Application Areas

Mobius has a number of potential application areas that we have investigated.

1) V2X Malware Propagation: Our primary use case is in ascertaining the risk that malware propagating over the broad number of communication channels present in modern and feature vehicles may pose. Simulation of malware propagation amongst a transient population of vehicles through V2X (vehicle-to-vehicle, infrastructure, person, etc.) communications involves the use of biological epidemiological models to understand how malware may behave across a population.

2) Vehicle Impacting Malware: Our simulator has a number of pre-programmed malware behaviours such as forced braking and constrained acceleration. There is however a broader variety of malware that may impact the behaviour of a manually driven or autonomous vehicle.

3) Autonomous vs. Manually Controlled Vehicles: Vehicles are rarely designed to operate in isolation, understanding the behaviour and interactions of a population of vehicles is invaluable to ensuring operational safety.

Driver assistance mechanisms are already commonplace, and these have an impact on driver behaviour. Understanding the consequence of increased awareness and propensity to recover due to such systems is important to understanding the broader societal impact such levels of automation have.

As we work towards levels of complete autonomy for vehicles, the percentage of autonomous vehicles on the road will increase. It's still unclear how exactly a mixed population of manual and autonomous vehicles will interact and there is a potential for autonomous vehicles, operating for maximum safety, may be subverted by manually driven ones.

The increase of automation will change the behaviour of vehicles, due to a reaction to driver assistance [22], fully autonomous characteristics [23] and human response to autonomous vehicles [24]. Monitoring the behaviour of vehicles as witnessed through smart infrastructure is one mechanism through which cyber attacks may be identified.

4) Non-Security Applications: As a tool for analysing the movement of vehicles Mobius has many potential application areas that we intend to explore in future research.

Electric Fleet Organisation: Electric vehicle infrastructure is generally limited in comparison to traditional fuels. This leads to an increased need for forward planning and resource management.

Mesh Network Organisation: The use of drones and portable towers to establish ad hoc wireless networks is an established solution in emergency scenarios [25]. One of the key challenges in this application is maintaining sufficient coverage, particularly as nodes in the network may need to replaced over time.

Civic Planning: Traffic management is a major concern for city planners and Mobius can enable simulation of a variety of scenarios, with different vehicle types and behaviours. New challenges will be introduced due to the change in behaviour that automation will bring and the ability to simulate a population transitioning from manually operated to autonomous vehicles will be essential to optimising infrastructure.

Blue Light Planning: Data on the geographical distribution of medical and criminal incidents, zoning and demographics can be valuable in strategically placing blue light resources.

III. Simulating Malware for CAVs

Mobius combines traffic simulation with abnormal vehicle behaviour controlled by epidemiological models to simulate malware events in CAVs.

A. Simulating Malware Events

The first stage of simulating malware events was to identify a number of potential scenarios that are likely to affect CAVs. These include attempts of theft, deliberate operator misuse or malicious attacks. Each of these scenarios can have a different impact on vehicle behaviour, to model these we create new vehicle classes for SUMO, each representing a different form of malware, with an associated change in behaviour. This included:

- Uncontrolled acceleration
- Speed limiting
- Forced breaking
- Erratic control / lane discipline

TraCI enables control of the simulation and at each step, an analysis of vehicle positions is carried out and a probability of infection is calculated using the model that has been defined. In line with that calculation, vehicle classes are changed to an infected form, in line with the parameters of the simulation.

The initial pool of susceptible vehicles is configurable, and typically based on local statistics of vehicle distribution, so that malware impacting a specific component (such as an ECU present in certain vehicles from a certian manufacturing date) can be simulated. The simulation also allows the user to specify both the type of vehicle and the percentage of those vehicles that are susceptible.

B. Malware Propagation

To simulate epidemiological spread of malware between vehicles, a base compartmental epidemiological model was specified. The model accepts a parameter specifying a subset derivative that the simulation can follow. After an initial infection at a specified time, at each simulation step for each infected vehicle, the non-infected vehicles within a stated proximity are collected and infected with a probability for a length of time, after which the vehicle either enters a cured state or a normal state (depending on the type of epidemiological model). Its also important to note that both the type and proportion of vehicles that are susceptible to infection can both be altered.

We implemented the following models:

- SIR malware propagation model
- SIS malware propagation model
- a single infection instance

In the single infection instance we target a single vehicle for infection, with no ability for malware to spread to other vehicles. SIR allows us to model a scenario where vehicles can be permanently patched and SIS allows us to investigate scenarios where vehicles can be recovered, but not patched.

As pointed out by Mickens and Noble [18], the standard compartmental models are not suited to mobile networks due to reliance on average connectivity. Trullols-Cruces et al. [19] explored an SIR model for large-scale vehicular networks. In our implementation on the compartmental models, we evaluate the transition rate (βSI) at each step in a simulation cycle for only a localised set of vehicles, inline with some form of inter-vehicular (V2X) communication.

These models require further development to account for a greater range of scenarios. For example, we posit that smart infrastructure (traffic lights, signage, etc.), test and diagnostic equipment (in vehicle workshops) and consumer technologies may be vectors for infection. Fuel stations and parking facilities could become “super spreaders” due to the high density of vehicles and prolonged proximity.

IV. Results

We present here initial observations from our experiments with SIS modelling. It should be noted that a number of parameters such as range, infectious period and behaviour can be configured. These will vary based on the precise nature of the malware and the region of the simulation.

In these experiments we assume a single source of infection has been introduced to a locality and a roadside process (such as factory reset of infotainment system) can be used to recover the vehicle and eradicate the malware after a period of 100 seconds. During this 100 seconds, the vehicle will be forced to brake. Any susceptible vehicle that comes within a predefined distance is at risk of becoming infected also. There is no latency from infection to effect.

434 vehicles with randomised journeys were introduced to a transport network derived from an area of Central London over a two hour period, with the majority of travel occurring

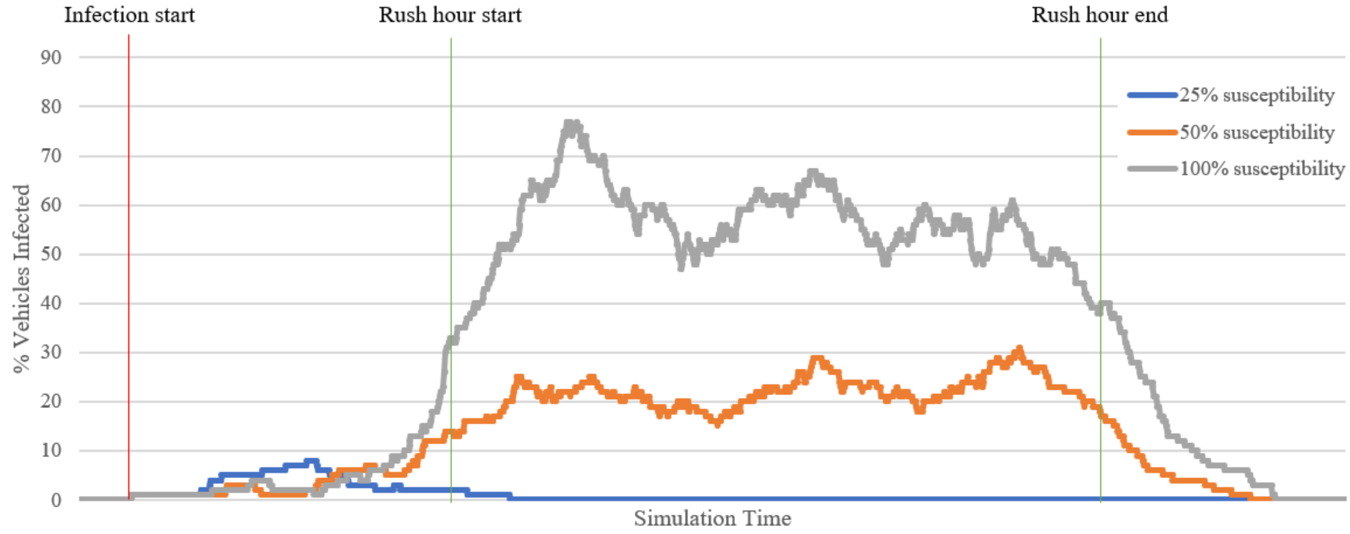


Fig. 2. Infection spread through V2V communication in Central London at varying rates of vehicle susceptibility in an SIS model. The data is for a two hour period. With a susceptibility of 25% the malware does not persist throughout the population. With 50% and 100% susceptibility, the infection remains until traffic reduces.

in a 1 hour slot. Cars, motorcycles, HGVs and buses were all potentially vulnerable infection with a distance of 20-metres for infection, but we varied the percentage of those that were susceptible in each test, starting with 25% through to 50% and lastly with 100%.

Figure 2 shows a simulation of infection in a busy city area with varying levels of susceptibility over a two hour period with a period of high traffic. This graph demonstrates the value of diversity in computer networks, with a higher diversity, there is less likely to be a common vulnerability and subsequently lower susceptibility. We can see that with a low susceptibility of 25% the transient nature of vehicles means the likelihood of transmission remains low and all vehicles eventually recover. With higher susceptibility, the infection persists in a population until the traffic is reduced.

In our second round of experimentation we investigated the role that traffic density has on the severity of infection. In Figure 3, there was an initial peak of infection, but due to the low-likelihood of vehicles becoming infected and the low number of vehicles on the roads, the infection rapidly disappeared.

The results from this experiment show that with a larger population, malware can persist for longer and potentially cause reinfections and have a larger impact, even with a relatively low rate of infection.

It is worth noting, that as vehicles were forced to brake when infected, there is limited opportunity to travel the network and spread the malware - a susceptible person must come into proximity. Not all malware will follow this pattern.

V. Conclusions

We have presented an overview of the compartmental models that play an important role in understanding the nature of disease and how it is best managed. These models have been successfully applied to the cyber domain, but mobile networks present new challenges. The application of compartmental models to mixed autonomous vehicle networks is a novel contribution. There already exist many popular variants of the compartmental models and we have demonstrated the applicability and suitable modifications for the purpose of cyber resilience of CAVs through Mobius, novel software for simulating malware propagation and impact in vehicle networks. Our simulations have demonstrated that diversity in vehicle technology can be an effective barrier to transmission of malware through a population.

Further work is needed to define models that account for the diverse potential vectors of disease transmission that are not present in traditional epidemiological models due to biological constraints. Computer systems can often be "inoculated" against specific attacks once they are known, where a susceptible or infected machine can be updated (over the air, or at a garage), the SIRV model for example may be useful in this regard [21].

Information on smart infrastructure and the various network nodes that may be responsible for vehicles other than the vehicles themselves needs to be accounted for in future models. For example, signalling, fuel stations and parking facilities are potentially risky "super spreaders" of malware. In addition there are numerous existing and emerging technologies for communications between CAVs and infrastructure, modelling of these and their specific characteristics would be desirable for ascertaining the risk of new technologies.

Geospatial analytics will play a crucial role in analysing the behaviour of vehicles and detecting anomalous behaviour when intelligence from intra-vehicle IDS may not be sufficient.

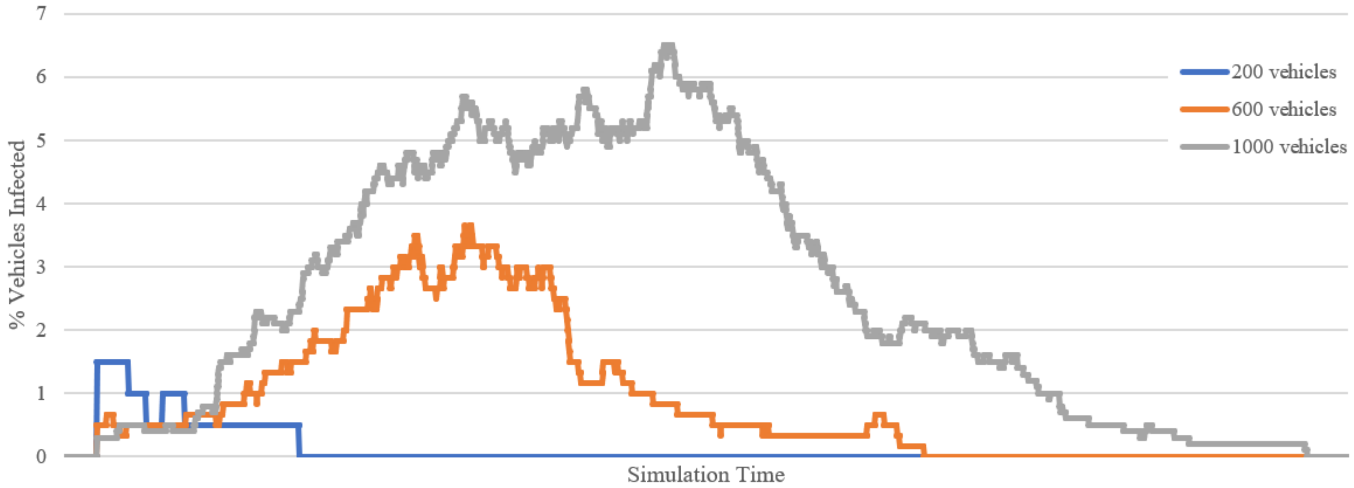


Fig. 3. Results for a trial of low susceptibility rates with varying traffic density. The infection persists for longer, affecting a greater percentage of vehicles for a longer period of time as population increases.

As new standards that necessitate establish security operation centres[2], there is an increasing need for robust technology to understand vehicle behaviour.

References

[1] Catapult Transport Systems, "Market Forgecast for connected and autonomous vehicles," 2017.

[2] Zenzic, "UK Connected and Automated Mobility Roadmap to 2030," 2019.

[3] I. Rouf, R. D. Miller, H. A. Mustafa, T. Taylor, S. Oh, W. Xu, M. Gruteser, W. Trappe, and I. Seskar, "Security and privacy vulnerabilities of in-car wireless networks: A tire pressure monitoring system case study." in *USENIX Security Symposium*, vol. 10, 2010.

[4] M. Harris, "Researcher hacks self-driving car sensors," *IEEE Spectrum*, vol. 9, 2015.

[5] A. Francillon, B. Danev, and S. Capkun, "Relay attacks on passive keyless entry and start systems in modern cars," in *Proceedings of the Network and Distributed System Security Symposium (NDSS)*. Eidgenössische Technische Hochschule Zürich, Department of Computer Science, 2011.

[6] C. Miller and C. Valasek, "Remote exploitation of an unaltered passenger vehicle," *Black Hat USA*, vol. 2015, p. 91, 2015.

[7] N. Huq, R. Vosseler, and M. Swimmer, "Cyberattacks against intelligent transportation systems," *TrendLabs Research Paper*, 2017.

[8] H. M. Song, H. R. Kim, and H. K. Kim, "Intrusion detection system based on the analysis of time intervals of can messages for in-vehicle network," in *2016 international conference on information networking (ICOIN)*. IEEE, 2016, pp. 63–68.

[9] O. Y. Al-Jarrah, C. Maple, M. Dianati, D. Oxtoby, and A. Mouzakitis, "Intrusion detection systems for intra-vehicle networks: a review," *IEEE Access*, vol. 7, pp. 21 266–21 289, 2019.

[10] D. Bernoulli, "Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir," *Histoire de l'Acad., Roy. Sci.(Paris) avec Mem*, pp. 1–45, 1760.

[11] N. T. Bailey *et al.*, *The mathematical theory of infectious diseases and its applications*. Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE., 1975.

[12] A. M'Kendrick, "Applications of mathematics to medical problems," *Proceedings of the Edinburgh Mathematical Society*, vol. 44, pp. 98–130, 1925.

[13] J. O. Kephart and S. R. White, "Directed-graph epidemiological models of computer viruses," in *Computation: the micro and the macro view*. World Scientific, 1992, pp. 71–102.

[14] J. Gani, "The maki–thompson rumour model: a detailed analysis," *Environmental Modelling & Software*, vol. 15, no. 8, pp. 721–725, 2000.

[15] E. Yom-Tov, N. Levy, and A. Rubin, "Modeling infection methods of computer malware in the presence of vaccinations using epidemiological models: An analysis of real-world data," *arXiv preprint arXiv:1908.09902*, 2019.

[16] S. Yu, G. Gu, A. Barnawi, S. Guo, and I. Stojmenovic, "Malware propagation in large-scale networks," *IEEE Transactions on Knowledge and data engineering*, vol. 27, no. 1, pp. 170–179, 2014.

[17] T.-F. Yen, V. Heorhiadi, A. Oprea, M. K. Reiter, and A. Juels, "An epidemiological study of malware encounters in a large enterprise," in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, 2014, pp. 1117–1130.

[18] J. W. Mickens and B. D. Noble, "Modeling epidemic spreading in mobile environments," in *Proceedings of the 4th ACM workshop on Wireless security*, 2005, pp. 77–86.

[19] O. Trullols-Cruces, M. Fiore, and J. M. Barcelo-Ordinas, "Understanding, modeling and taming mobile malware epidemics in a large-scale vehicular network," in *2013 IEEE 14th International Symposium on" A World of Wireless, Mobile and Multimedia Networks"(WoWMoM)*. IEEE, 2013, pp. 1–9.

[20] P. A. Lopez, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, and E. WieBner, "Microscopic traffic simulation using sumo," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2018, pp. 2575–2582.

[21] I. Masaaki, "Optimal strategies for vaccination using the stochastic sirv model," *Transactions of the Institute of Systems, Control and Information Engineers*, vol. 25, no. 12, pp. 343–348, 2012.

[22] M. Cunningham and M. A. Regan, "Autonomous vehicles: human factors issues and future research," in *Proceedings of the 2015 Australasian Road safety conference*, vol. 14, 2015.

[23] D. Silver, J. A. Bagnell, and A. Stentz, "Learning autonomous driving styles and maneuvers from expert demonstration," in *Experimental Robotics*. Springer, 2013, pp. 371–386.

[24] D. Moore, R. Currano, M. Shanks, and D. Sirkin, "Defense against the dark cars: Design principles for griefing of autonomous vehicles," in *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, 2020, pp. 201–209.

[25] M. Deruyck, J. Wyckmans, L. Martens, and W. Joseph, "Emergency ad-hoc networks by using drone mounted base stations for a disaster scenario," in *2016 IEEE 12th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. IEEE, 2016, pp. 1–7.

[2] See ISO/SAE DIS 21434