# Characteristics of a vehicular network

Antony Rowstron[1] and Giovanni Pau[2]
[1]Microsoft Research, Cambridge, UK
[2]University of California at Los Angeles, California, USA.

## ABSTRACT
If the full range of applications and protocols that exploit vehicular networks are going to be adopted then it is key to understand the performance characteristics of these networks: both the spatial and temporal properties. Many of the applications, services and protocols that could drive deployment of vehicular networking can only be developed if we understand the properties of the network that will be created. Yet, to really understand these properties we need studies of large-scale vehicle traffic traces, which are difficult to obtain. To date, most studies have used traces of small street or freeway sections, or the traces have been generated using unrealistic mobility models.

To address this issue we present an extensive analysis of a mobility trace generated by Los Alamos National Laboratory for Portland, Oregon. It is generated using the TRANSIMS micro-traffic simulator, combined with macro-level traffic information gathered in the city from censuses and studies over a year. Hence, this trace is far more accurate than previously studied synthetic traces. We analyze it to identify and characterize the key properties starting from the characteristics of the vehicle movement. We then explore the properties of the network that would be formed, considering different radio ranges and penetration rates. Finally, we then evaluate how these impact the design space for algorithms that use either one-to-many or one-to-one communication patterns.

## 1. INTRODUCTION
It is quite feasible that in the next few years vehicular networks will become a reality. This could either be through a car manufacturer incorporating 802.11p, or an equivalent, or through the widespread deployment of vehicle mounted consumer devices, such as Satellite Navigation/GPS devices also known as a personal navigation device (PND). Indeed, in Western European countries, already one in four vehicle owners also have a vehicle mounted PND. In the USA, the penetration rate is lower at around 10%, but the popularity of PNDs, in both Europe and the USA, is growing rapidly. A natural progression in the evolution of these devices would be to equip them with commodity 802.11 WiFi. This would then create an ad hoc vehicular wireless network consisting of these PND devices. Indeed, some manufacturers have already included commodity WiFi in their devices [20].

To help vehicular networks emerge it is important to demonstrate protocols and applications that work on the networks that would be formed. Understanding the characteristics of the created network is fundamental to creating these applications and protocols. Characterizing the network allows us to really understand the constraints and feasibility of protocols and services that can be run on, and even exploit, the networks formed. For example, are these networks frequently and dynamically partitioning? Is the density sufficient, in general, to support multi-hop routing? If the network is partitioning and we develop disruption tolerant protocols what is the end-to-end delay that would be expected? How will the network change as the device penetration rate increases? In this paper we attempt to address these questions.

We also consider how the base characteristics of the network impact the type of applications that can be supported by the network. Much of the pragmatic applied research for vehicular-based networks has focused on one-hop vehicle-to-infrastructure scenarios [15, 3, 28], and the opportunistic use of WiFi access points. Are we limited to simple one-hop vehicle-to-infrastructure scenarios, or single local one-hop information propagation? If we can do multi-hop, what sort of properties do the routing protocols need to handle? Much wireless networking research makes the assumption that the network is not partitioned, even if it is mobile. This means that routing protocols assume that there exists a set of nodes that can directly route a packet from a source to destination, for example AODV [22] and DSR [8]. Alternative, geographic routing protocols [9, 14, 29, 19] do not attempt to maintain a path between the source and destination, they create the path dynamically when packets are routed. However, they do assume that there exists a set of nodes that can forward the packet from the source to the destination, and if progress cannot be made they initiate a mechanism to route around the void. If the network is partitioned, then this makes little sense. Many protocols assume that each node has a static fixed identifier and a dynamic location, and that a source is able to determine the current location of the destination. This means that they require

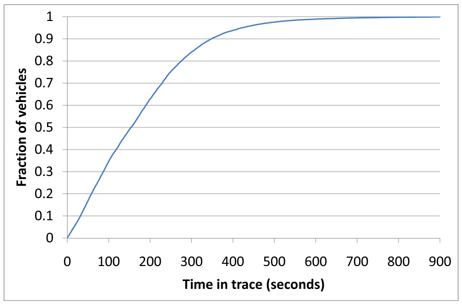


**Figure 1: A CDF of vehicles versus time in trace.**

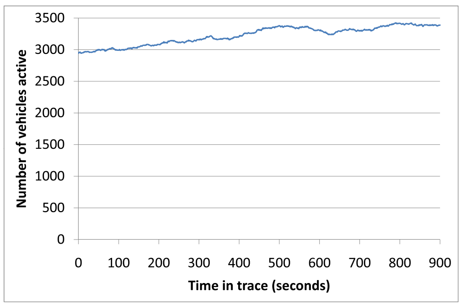


**Figure 2: Number of active vehicles per second during the trace.**

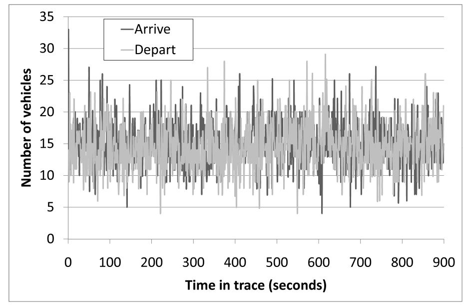


**Figure 3: Number of vehicles per second arriving and departing the trace from seconds 1 to 899.**

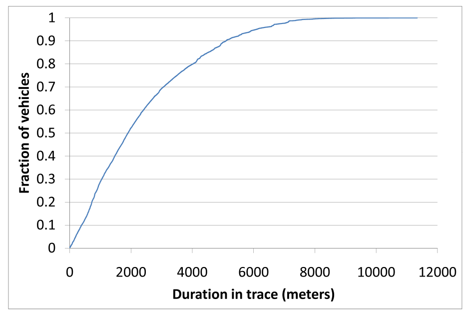


**Figure 4: A CDF of vehicles versus distance covered during trace.**

a mechanism to translate between a fixed identifier and the associated dynamic location. This can be implemented as a distributed hash table (DHT), or an equivalent service like GLS [16, 27], but is this feasible?

In this paper we use a realistic mobility trace for a 3km by 7km region of the city of Portland, Oregon covering the main downtown area and two different freeways. It was generated by Los Alamos National Laboratory using vehicle flow data from censuses and other sources, and uses a micro-level simulator (TRANSIMS) for detailed per-vehicle movement information. We believe that this trace is significantly more accurate than previous mobility traces.

The rest of the paper is organized as follows: Section 2 provides an overview of the trace used. Section 3 describes the base properties of the trace, and Section 4 then looks at the base properties of a vehicular network running across the vehicles, or subset of the vehicles, in the trace. Section 5 examines how the network topology impacts both one-to-many and one-to-one communication patterns. Section 6 draws some high-level conclusions about the impact of the results on applications, protocols and services that could run on this network. Finally, Section 7 describes related work, and Section 8 concludes.

## 2. THE TRACE

We ran the experiments using a mobility trace provided by Los Alamos National Laboratories (LANL) for the city of Portland, Oregon. The trace was originally created for homeland security reasons, and considerable effort went into ensuring the trace was realistic. It was generated using both macro- and micro-simulations using a vehicular traffic simulator, TRANSIMS [24, 17, 2]. TRANSIMS is a micro-traffic simulator that models the interaction between vehicles traveling along streets and at junctions etc. To increase the accuracy of the simulations, LANL incorporate macro-level per-vehicle activity flows, generated from census and other survey information gathered explicitly for the trace, over a period of a year. From this information it is possible to infer schedules for vehicles, such as a car leaves a house at a particular time and travels to an office. Generating the mobility traces requires significant computation resources as well as the detailed city activity surveys. It is believed to be realistic of the traffic flow in downtown Portland, and covers 15 minutes from 8:00AM on a weekday, for an area 3km by 7km area of downtown Portland straddling the Willamette River, covering both urban streets and two freeways (I5 and I405).

## 3. BASE PROPERTIES OF THE TRACE

The trace is 900 seconds long, and provides position information for 16,528 unique vehicles, e.g. cars, trucks and buses. When a vehicle is active a new location is provided per second. In order to understand the core properties of the trace, we present a number of base statistics extracted from it.

The trace covers a 3km by 7km region, and vehicles will enter and leave this region. Figure 1 shows a CDF of vehicles versus the duration of the vehicle in the trace in seconds. We define the duration as the period between the first and last reference to the vehicle in the trace. Across all vehicles, the median vehicle duration is 154 seconds, or approximately

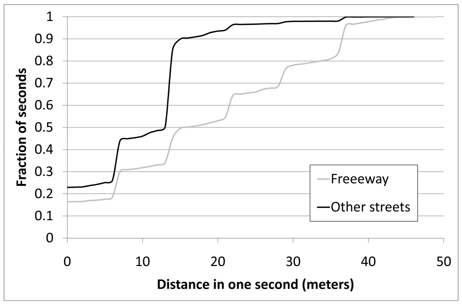


**Figure 5: A CDF of vehicle seconds versus distance covered per second.**

17% of the 900 second trace. The 95th percentile is 424 seconds, while the 99th percentile is 616 seconds. During the trace a small number of vehicles (381) leave and then re-enter the trace region. For these vehicles we consider the duration to be from when the vehicle first appears in the trace, to the last reference to the vehicle in the trace. Also, if the destination of a vehicle lies within the region then once the vehicle reaches that destination, it's location is not updated and we consider the vehicle to have left the trace.

In order to understand more about the density of vehicles in the trace, Figure 2 shows the number of vehicles active in the trace per second over the 900 seconds of the trace. A vehicle is considered active if there is a location update associated with the vehicle in that second. At the start of the trace there are 2,942 vehicles active (or, 17.8% of all vehicles in the trace) while at the end of the trace this has increased to 3,388 (or, 20.5% of all vehicles in the trace). Given that the trace represents 15 minutes from 8AM it is unsurprising that the density of vehicles increases during the trace.

Figure 3 shows the number of vehicles that arrive and depart from the trace per second over the period of the trace, ignoring the first and last second of the trace. The arrival rate of vehicles is slightly higher than the departure rate, as would be expected, as the total number of vehicles active per second increases slightly during the trace.

Figure 1 showed the vehicle duration in terms of seconds, and in contrast Figure 4 shows a CDF of vehicles versus the distance, in meters, they cover while active in the trace. The distance covered is calculated by summing the direct linear distance between each location coordinate for a vehicle while it is active in the trace. For the small number of vehicles (381) that leave the trace region and then re-enter, we sum the distance covered while present in the trace and ignore the distance covered while outside the trace. The median distance covered by a vehicle in the trace is 1,907 meters, the 99th percentile is 7,472 meters, and the maximum is 11,337 meters. Given that the trace covers a region of 3km by 7km these figures are consistent with what would be expected.

To understand the vehicle speeds in the trace we calculated for every vehicle active in the trace at second $t$ and $t+1$, the distance covered in the second. Figure 5 shows a CDF of fraction of the seconds measured versus the distance covered in meters. The total number of seconds calculated is the sum of duration of each vehicle in the trace. In order to help understand the distribution of speed across freeways and other streets, for each measurement we classified the vehicle as either being on the freeway or on a street. The trace includes two major freeways, the I405 and I5, and when classifying a measurement we included all entry and exit ramps as belonging to the freeway. Approximately 35% of the measurements are attributed to freeways. The average velocity for all vehicles in the trace is 29.6 mph.

Figure 5 shows a number of things: the median distance covered per second for the freeways is 16 meters (36.0 mph) while for the other streets it is 13 meters (29.3 mph). The 90th percentile for the other streets is 15 meters (33.8 mph) and for the freeways it is 36 meters (81.4 mph), and for both types of street this seems reasonable given the speed limits in the USA. The maximum speed observed on the freeway, is 49 meters per second (110 mph), which seems high, but only represents a very small fraction of the total vehicle seconds in the trace. This could be attributable to emergency vehicles, but we have been unable to confirm this. In Figure 5 a distinct step is visible at approximately 7.5 meters, and multiples thereof, which would not be necessarily expected. We believe that these are artifacts of how the traces are generated. Unfortunately, we do not have access to the original configuration files used to generate the traces, so are unable to confirm this.

| Penetration rate | Number of vehicles |
|---|---|
| 0.00390625 | 64 |
| 0.0078125 | 129 |
| 0.015625 | 258 |
| 0.03125 | 516 |
| 0.0625 | 1033 |
| 0.125 | 2066 |
| 0.25 | 4132 |
| 0.5 | 8264 |
| 1.0 | 16528 |

**Table 1: Penetration rates used with the number of vehicles in the trace that this rate represents.**

# 4. BASE PROPERTIES OF THE NETWORK

Having explored the basic properties of the vehicles in the trace, we now look at the base characteristics of the vehicular network that would be formed if WiFi-enabled devices were mounted in the vehicles in this trace. For all these results, we vary the fraction of vehicles carrying a device, which we refer to as the *penetration rate*, as shown in Table 1. Unless otherwise stated, at penetration rates below 1.0 we randomly select the vehicles equipped with wireless from the original trace. The penetration rates cover a wide range and to provide some context, recently, TomTom stated that 25% of all vehicle owners in Western Europe also had a PND, and for the USA this rate is 10%. Of course, not all these PNDs will be used concurrently. However, we also observe that there are a small number of manufacturers of PND devices that share most of the market. Therefore, while penetration rates of 12.5% and higher may seem optimistic, we believe that the lower penetration rates considered do cover the range which is feasible in practice.

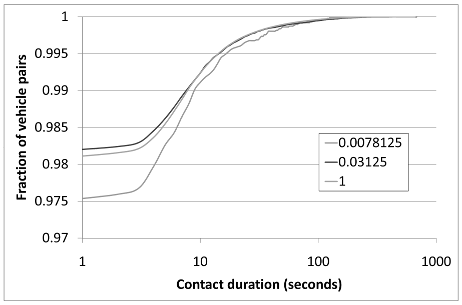


**Figure 6: A CDF of vehicle pairs versus contact time in seconds with connectivity range of 150m.**

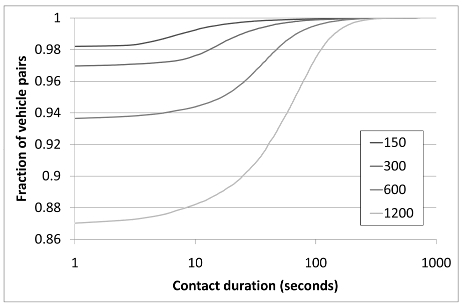


**Figure 7: A CDF of vehicle pairs versus contact time in seconds with varying connectivity range with a penetration rate of 0.03125.**

In the following sections we use a default wireless connectivity range of 150 meters, which seems realistic based on current measurements of commodity WiFi with internally mounted antennas. We also compare the performance with longer ranges that could potentially be achieved if the antenna was mounted on the exterior of the vehicle.

In all our measurements we are looking to determine the basic properties, and are not modeling channel contention or other such phenomenon that would be encountered in reality. In this paper we are attempting to understand the lower or upper bounds of what could be achieved.

## 4.1 Inter-vehicle contact duration

We first explore the inter-vehicle contact time. In order to do this we analyzed the trace to measure the number of seconds during the trace that each unique pair of vehicles was in connectivity range. So, for the penetration rate 0.00390625 we considered 4,032 unique pairs of vehicles, while at penetration rate 1.0 we considered 273,158,256 pairs. Figure 6 shows a CDF of the vehicle pairs, normalized to the total number of pairs, against the contact duration in seconds. We show the results for a subset of penetration, the penetration rates that achieved the best and worst performance, and for a penetration rate of 1.0.

As would be expected, Figure 6 shows that, independent of the penetration rate, the vast majority of vehicle pairs never encounter each other. Figure 6 uses a log plot, and the intersection point of each line with the y-axis indicates

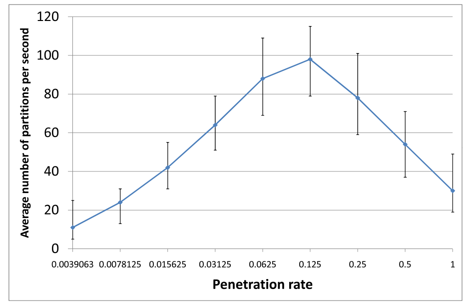


**Figure 8: Number of partitions versus penetration rate, mean with max min error bars, for 150 meter communication range.**

the fraction of pairs that did not come into contact. In all cases *less than 2.5% of the pairs* are able to connect at all. Due to the temporal diversity that we observed in the base trace this is not so surprising. Figure 6 shows that in all the cases the majority of inter-vehicle connection times are in the range of 10 to 100 seconds, but there are some connections that last considerably longer. The longer connectivity durations are due to vehicles travelling in the same direction. The shorter connection periods are predominantly due to the other vehicle travelling in the opposite direction.

Intuitively, a nice property is that the penetration rate should have little impact on connection duration, and this can be seen in Figure 6. However, varying the communication range would be expected to change the connection duration. In order to understand the impact it has, we analyzed the trace using four communication ranges, 150, 300, 600 and 1200 meters. Figure 7 shows the fraction of vehicle pairs versus contact duration for the different communication ranges, with a penetration rate of 0.03125. The results show that roughly each time the communication range doubles, the fraction of vehicle pairs that can communicate doubles. This is counter intuitive, as doubling the communication range will increase the area covered by a factor of four. The roads force the traffic not to be uniformly spread over the region, and this result clearly demonstrates the impact of the more linear road structure, where doubling the distance will only double the number of vehicles.

In Figures 6 and 7 it is also interesting that the contact durations are dominated by short-lived connections, in most cases between 10 and 100 seconds. This implies that we need to ensure that protocols are efficient at using the connection when it exists.

When drawing conclusions from this trace, a number of things need to be considered, first the trace is only 900 seconds long, and, for example, the contact times will be biased by the vehicles leaving and joining the trace region. In reality vehicles will not just leave the region. Secondly, it should be noted that just because two vehicles are in communication range, it does not mean that they can connect, for example, a building could be between them which will block the signal.

## 4.2 Network connectivity

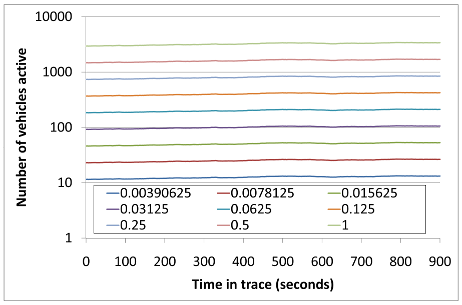


**Figure 9: Number of active vehicles over the 900 second, for different penetration rates.**

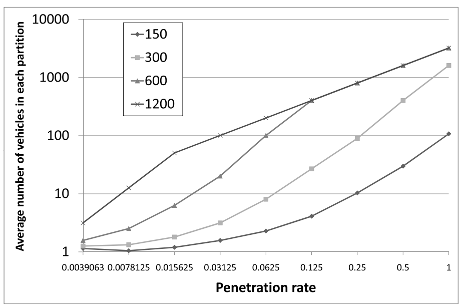


**Figure 10: Average number of vehicles per partition versus penetration rate for varies communication ranges.**

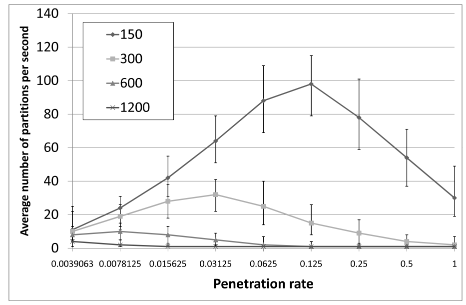


**Figure 11: Number of partitions versus penetration rate, mean with max min error bars for varies communication ranges.**

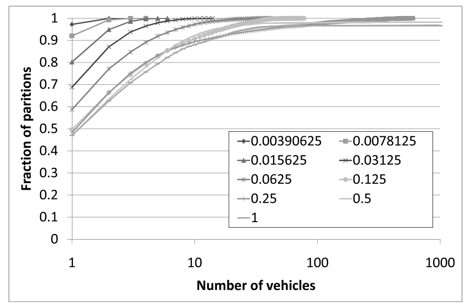


**Figure 12: CDF of fraction of partitions versus numbers of vehicles in each partition, for 150 meter communication range.**

Having examined the inter-vehicle connection durations, we now look at the connectivity of the network, and in particular whether the underlying network is partitioning. To do this we calculated, per second of the trace, the number of partitions formed by the network. Figure 8 shows the average number of partitions per second for the 900 second trace, plus the maximum and minimum number of partitions observed across the entire 900 seconds, versus the penetration rate with the default 150 meter communication range.

At first glance the results in Figure 8 appear counter intuitive, with a peak at penetration rate 0.125. However, to explain the results we need to consider the number of vehicles active per second for each penetration rate, and this is shown in Figure 9. Figure 9 shows the number of vehicles active per second versus the 900 seconds of trace, for each of the penetration rates. For the lowest penetration rate, on average across the 900 seconds there are only 12.6 vehicles active per second. Therefore, the number of active vehicles per second acts as a lower bound on the number of partitions, hence the value of 11 for the lowest penetration rate in Figure 8.

To understand this further, Figure 10 shows for each penetration rate and communication range the expected number of vehicles per partition per second. Interestingly, that for 150 meters communication range, even at penetration rate of 0.25 there are still only, on average, 10 vehicles per partition. At full density, and 150 meters communication range there are, on average, 30 partitions (Figure 8), but each only has 107 vehicles on average (Figure 10).

Figure 11 shows for all communication ranges the average number of partitions per second versus the penetration rate. As would be expected, increasing the communication range decreases the number of partitions and, as a natural consequence, increases the average number of vehicles per partition. A single partition is formed only with communication ranges of 600 and 1200, from penetration rate 0.125 and 0.015625, respectively. As a single partition is formed, all vehicles in the trace are part of the partition, which is why the 600 and 1200 lines in Figure 10 are identical from penetration rate 0.125.

Figure 10 showed the average number of vehicles expected in each partition. In order to understand this further, and the difference between freeways and other streets, we analyzed the partition size distribution in detail. Figure 12 shows a CDF of all partitions observed during the 900 seconds, with a communication range of 150 meters, for the different penetration rates. For clarity, we have cut the x-axis at 1000 vehicles. For all penetration rates the median size of a partition is approximately 1. The large difference between the mean (Figure 10) and the median is an indicator of the skew in partition sizes.

Figure 12 shows the results for all vehicles in the trace, and next, we examine the size of partitions considering only vehicles on a freeway or freeway ramp. Figures 13 and 14 shows

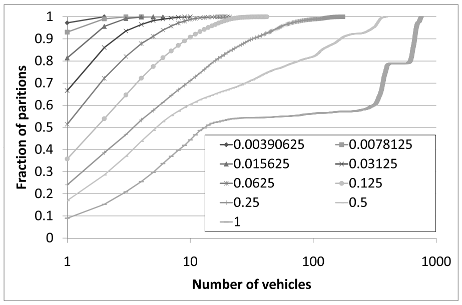


**Figure 13: CDF of fraction of partitions versus numbers of vehicles in each partition, for 150 meter communication range, for freeway.**

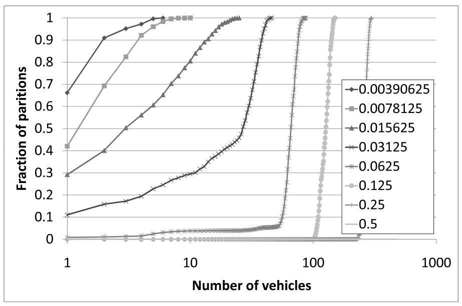


**Figure 14: CDF of fraction of partitions versus numbers of vehicles in each partition, for 1200 meter communication range, for freeway.**

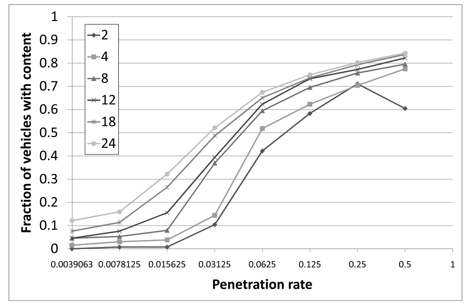


**Figure 15: Fraction of vehicles receiving content versus penetration rate with various numbers of seeds, with range 150 meters.**

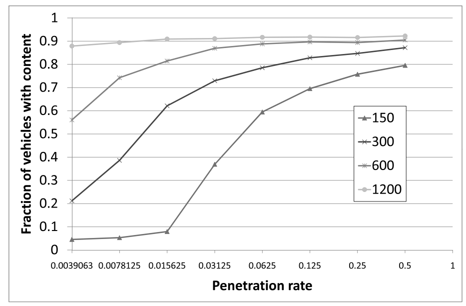


**Figure 16: Fraction of vehicles with content versus penetration rate for 8 seeds with varying communication ranges.**

a CDF of all partitions observed during the 900 seconds, for a communication range of 150 and 1200 meters respectively, for the different penetration rates for freeway vehicles. The freeway has a much higher density of vehicles in general, and this is therefore reflected in higher numbers of vehicles in each partition. Indeed, for a penetration rate of 1.0 we observed 27,057 partitions in the full trace, with 150 meters, with only 4,228 partitions in total for the freeway vehicles. Figure 14 shows the impact of increasing the communication range to 1200 meters. Interestingly, at 0.25 and higher, we can observe that all partitions have 100 or more vehicles. Indeed, from 0.125 and higher there is only *one* partition per second, containing all the vehicles.

These results show the importance of building protocols and services that are tolerant to dynamic partitions. As we will see later, despite the fact that the network is partitioned, the rate of change of partition membership is quick, allowing disruption tolerant protocols to still perform well. If a protocol expects all communicating vehicles to be in the same partition, then the number of vehicles each can reach will be small, unless both the communication range and penetration rate are high.

### 4.3 Summary of network topology

To summarize, these results indicate that the network, at reasonable connectivity ranges and penetration rates, is highly partitioned and dynamic. The inter-vehicle contact time is usually short, in the region of tens of seconds. Expanding the communication range will not necessarily yield as large an increase in the number of vehicles reachable as could be expected.

## 5. HIGHER-LEVEL PROTOCOLS

We have so far examined the base properties of the trace and the network topology that would be formed. In this section we evaluate the higher-level properties that impact one-to-one as well as one-to-many communication patterns.

### 5.1 One-to-many

We first evaluate the performance of protocols that rely on one-to-many communication patterns, for example, epidemic dissemination. In the experiment we assume that there are a number of vehicles that act as seeds to the network, that have information that we would like epidemically distributed in the network. In the experiments, we assume that the seeds are additional to the base set of vehicles selected for a particular penetration rate. Hence, after randomly selecting the vehicles for a penetration rate, we randomly select a further set of vehicles to act as mobile seeds. It is assumed that each of the seeds is pre-loaded with some content required by all devices.

To transfer the content between two vehicles, the two vehicles must be within range for $q$ seconds, where we fix $q = 20$. We assume a per-second transfer rate, meaning if a vehicle is in contact with another vehicle for $k$ seconds then $k/20$th of the content is transferred. Vehicles store partially trans-

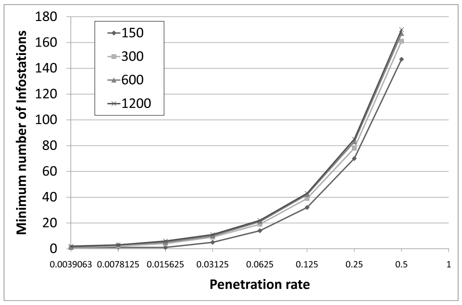


**Figure 17: Lower bound on the number of infostations required to achieve the same dissemination as V2V.**

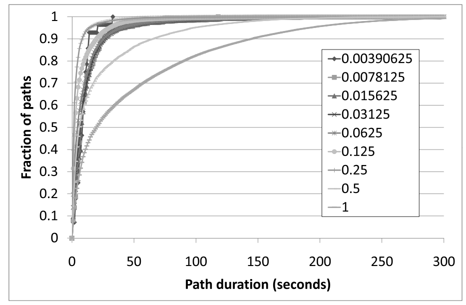


**Figure 18: CDF of fraction of paths versus path duration in seconds for all penetration rates with a communication range of 150.**

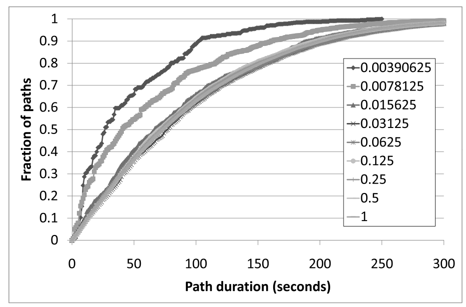


**Figure 19: CDF of fraction of paths versus path duration in seconds for all penetration rates with a communication range of 1200.**

ferred content, and can resume transfer when another vehicle with the content is encountered. Conservatively, the vehicles *do not* become content sources until they have the full content. From experimental results we believe that a contact time of 20 seconds at 150 meters should enable us to transfer approximately 20MB between devices mounted within a vehicle.

We simulate the full 900 seconds of the trace, and then measure the fraction of vehicles with the full content, *excluding* the original seed vehicles. It should be noted that as the seed vehicles are selected randomly, they are unlikely to be present in the trace for the full 900 seconds. At lower penetration rates and with small numbers of seeds, the length of time each seed was active had a big impact on the results. To compensate for this we run each experiment 20 times and present the median result.

Figure 15 shows the fraction of vehicles at the end of the trace that have the content versus the penetration rate with a communication range of 150 meters with varying number of seeds. These results show that even at very low penetration rates, with a small number of seeds, the content can spread rapidly.

Figure 16 shows the impact of varying the communication range. For the 1200 and 600 ranges you would expect, at penetration rates 0.125 and higher when the network is a single partition (see Figure 11), that all vehicles would have the content, rather than only 90% of them. This is an artifact of selecting the seed nodes randomly, and not ensuring that they are in the trace from the start. Therefore, some vehicles have left the trace before the content is spreading. Further, before one vehicle can spread the content to another vehicle it must have the entire content, which requires 20 seconds of contact time with other vehicles already containing the content. This slows the initial rate at which content spreads.

These results demonstrate the feasibility of spreading large content using an epidemic one-to-many approach. The other approach to getting content into the system is to use infostations. Figure 17 shows the lower bound on the number of infostations required to deliver the content to the same number of vehicles as in the epidemic case, at the different penetration rates. This assumes a unicast model, where the infostation transfers content to a single vehicle at any point in time, and at the same rate as the content is transferred between vehicles. Hence a vehicle needs to be in contact for 20 seconds with the infostation to receive the content. Therefore, each infostation can deliver the content to 45 vehicles. We also assume that the infostation can be placed in a location where there are always vehicles available to transfer the content. Hence, the results in Figure 17 should be considered conservative and a lower bound on the number of infostations that would be required. Figure 17 shows the results for each of the communication ranges used in the epidemic version. As would be expected, given the relatively little impact that the communication has on the epidemic version, the number of infostations required to match the performance is relatively similar.

It should be noted that at the end of the 900 seconds, in the epidemic version, the number of vehicles that are carrying the content is higher than the original number of seeds. This means that during the next 900 seconds, the content will spread faster, as there are effectively many more seeds. This means that over a longer period the number of infostations required effectively will increase over time to match the same distribution rate.

## 5.2 One-to-one

Having considered epidemic dissemination, we next consider how routing protocols would perform on this network topology.

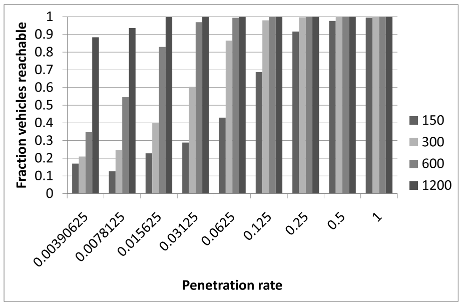


**Figure 20: Fraction of active vehicles reachable for all penetration rates and communication ranges.**

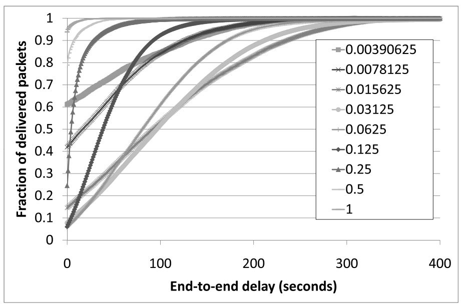


**Figure 21: CDF of packets delivered versus the end-to-end delay in seconds at 150 meters for all penetration rates.**

We first consider the fragility of routing paths in this network topology. A path between a source and destination is defined as existing if there are a set of active vehicles that could forward a packet from the source to a destination. To understand the duration of each path, we measured for every pair of vehicles when a path exists, and the duration of the path. We assume that the set of vehicles that forward the packet can change during the duration of a path, but once the pair of vehicles is partitioned the path is considered failed.

We have already examined the partition sizes and characteristics, and we can conclude from this that for lower penetration rates and for most communication ranges the fraction of active vehicles that can communicate directly with each other is limited; if two vehicles are not in the same partition they cannot communicate! For vehicles that can communicate, Figure 18 shows a CDF of number of paths versus path duration, across all penetration rates, for a communication range of 150 meters. For clarity, we have cut the x-axis at 300 seconds, and for all penetration rates the 99th percentile is below 300 seconds. The results show that the duration of the paths is short, for all penetration rates below 1.0, the median path duration is less than 6 seconds. Figure 19 shows the same results for a communication range of 1200 meters. The impact of extending the communication range is to increase the path duration, but the median is less than 75 seconds in all cases. This implies that multi-hop routing could be feasible, but at lower densities or with low communication ranges, then the path duration could be short lived.

Due to the partitioning and short-lived paths, we next examine the end-to-end delay that we could expect if we had a perfect disruption tolerant routing protocol. To do this we run an experiment where every second of the 900 trace we randomly select four active vehicles and each of these generate a single packet. This packet is then broadcasted through the network, such that at each second, any vehicle that is connected either directly or through other vehicles to a vehicle carrying the packet is considered to have received the packet. We measure the time when the packet is received by each vehicle that *was active when the source generated the packet*. This provides us with two metrics: the fraction of vehicles reachable, over time, that were active when the packet was generated, and the end-to-end delay that the vehicles that receive the packet experience. These results assume a zero transmission latency which, while unrealistic, provides an upper bound on the performance any disruption tolerant protocol could achieve on this network.

Figure 20 shows, on average per packet generated, the fraction of active vehicles that over the duration of the remaining trace receive the packet, for various penetration rates and communication ranges. If we compare these results with the results in Figure 11, we can see that in general the number of packets delivered to active vehicles is far higher than the average partition size. Hence, the vehicles moving between partitions are effectively delivering the packets. Therefore disruption tolerant protocols appear advantageous in this environment.

If we are to design disruption tolerant protocols, then it is also useful to understand the expected delivery delay. Clearly, designing protocols that handle a few seconds delay is very different to handling minutes of delay! Figure 21 shows a CDF of the fraction of packets delivered versus end-to-end delay for all penetration rates, using a communication range of 150 meters. What is interesting is that for the majority of penetration rates the median delay is less than 100 seconds. As would be expected, as the penetration rate increases, obviously the delay decreases.

# 6. SUMMARY: IMPACT ON DESIGN

We have presented a lot of details from the trace. At a high-level the key take away point of this study is that we need to build protocols that can efficiently adapt between highly disconnected and dynamically partitioning networks and non-partitioned networks. Many protocols perform well in a specific environment, but poorly across many environments. In particular, if protocols or services try to maintain state in the network, either at a particular location, for example at a junction [21], or try to maintain a distributed name to location service, like GLS [16], then these services will not perform well when the network is dynamically partitioning. We recognize that it is challenging to build applications and protocols that can efficiently work in low and high density scenarios and we believe that this is a major challenge for the community.

Infostation-based services should exploit the broadcast nature of the medium to ensure that as the vehicle density increases the number of infostations required does not increase dramatically. This also raises the issue about the popularity of the content transferred. For a content distribution system, it is possible to think of the penetration rate as capturing the impact of only a certain fraction of the vehicles wanting a particular piece of content, even if the number of equipped vehicles is higher. If the content is delivered by an infostation and is of interest to a small population then many infostations will be required. These issues suggest that building hybrid systems seems like a good way forwards, and these should use opportunistic caching of content as a way to increase the number of seeds even for unpopular content. If content is of broad interest, then the results show that a small number of well placed initial seeds performs well in such a system, even when the penetration rate is relatively low.

There has been interest in delay tolerant networking protocols for vehicle-based networks [12, 13]. Many delay tolerant protocols are designed to support *long-term* partitioning of the network, on the order of hours to days. This study shows that we need to build routing protocols that tolerate partitioning which results in end-to-end delays on the order of *seconds to minutes*, more *disruption* tolerant rather than delay tolerant.

We also observe that often vehicles become partitioned and then reconnect a few seconds later. This should be heavily impacting how we build end-to-end protocols for these environments. An end-to-end protocol like TCP/IP, with slow start and dependency on round-trip-time as well as its aggressive reaction packet loss, will not work well in these environments. This should drive the development of disruption tolerant end-to-end protocols.

# 7. RELATED WORK

Understanding mobility and it's impact on protocols design is critical [14, 18, 4], and we now provide an overview of related work. For a general taxonomy of synthetic mobility models see [5].

Generating realistic mobility traces is difficult. Real-world traces are always the most valuable, and several groups have collected real-world traces from 802.11 networks deployed on University Campuses [7, 10, 6, 25]. However, the mobility model pattern for humans walking around a campus are very different to those of vehicles travelling on roads. The DieselNet project at UMASS Amherst has gathered traces from an 802.11 testbed consisting of three campus Shuttle routes, where each route has a bus every 15 minutes in each direction. In [28] the inter-contact properties are studied and a synthetic model is derived for bus-based disruption tolerant networks. This is used to study epidemic dissemination in the network. The traces used in this study are for all vehicles in the city of Portland, rather than a specific subset, and includes thousands of vehicles. Buses, by their nature, have a periodic mobility pattern, compared to many vehicles. We believe that the results presented in this paper reflect more the properties of the larger-scale vehicular networks.

In general, the lack of real world traces has caused people to use synthetic mobility traces, but much of the prior work has used overly simplistic and therefore, unrealistic, general mobility models like random way point. However, an early attempt to use a more realistic model was by Lochert et al. [14] that used a traffic simulator developed by Daimler which uses maps. It is unclear how they parameterize the traffic and they consider experiments with a small number of source-destination pairs for only a few seconds. More recently map-based urban mobility models have become more popular. There has been the development of several micro-traffic simulators, including TRANSIMS, which is used to generate the trace used in this paper, as well as open source simulators like SUMO [11]. In [18] a microscopic traffic simulator is used to simulate 24-hours of traffic in the Zurich metropolitan area. The Zurich maps are provided by the Swiss planning authority and contain the information about the road capacity and maximum speed. In [18] they study the performance of several ad hoc routing protocols using this mobility trace and propose two mechanisms to improve the delivery ratio of AODV and GPSR. It is unclear how realistic the traffic trace is relative to the real traffic patterns in Zurich. The trace that we are using was generated by LANL for homeland security reasons. As well as using a micro-traffic simulator, they also used census and other information gathered from the city over a period of a year to ensure that the traffic trace generated was similar in characteristics to the traffic seen in Portland. This was a major exercise and we believe that this has generated one of the most accurate large-scale traffic mobility traces currently available.

Finally, a mobility evaluation framework called *IMPORTANT* was proposed in [1]. It includes a number of protocol independent metrics that capture key mobility characteristics. They used mobility traces generated using a variety of mobility models including Random Way Point, Manhattan, Group Mobility, and Freeway mobility models. It used *mobility metrics*, such as temporal dependency, spatial dependency, and geographic restrictions; and *connectivity graph metrics*, such as number of link changes, link duration, path duration, and path availability. Several other papers use these metrics to investigate the impact on particular routing protocols, for example on DSR and AODV [23, 26]. Many of the properties that we show are based on these metrics, however, in this paper we try and show the base properties of the network topology and then show how these impact the protocols that would run on these networks, including the impact of penetration rate and communication range.

# 8. CONCLUSIONS

In this paper we have used a detailed realistic mobility trace to try and explore both the base characteristics of the mobility and the network that would be formed, using different radio ranges and penetration rates. We believe that the trace analyzed is one of the best synthetic large-scale traces available, and that the insights we gain from this trace can be used to inform future protocol and application design.

We show that the network is highly dynamic with frequent partitioning. We show that the inter-contact times are short. We then study how these properties impact the design of both one-to-one and one-to-many communication patterns, and conclude that disruption tolerant protocols are going to be very important. Furthermore, we need protocols that are efficient in high density non-partitioned networks as well as low density dynamically partitioning networks. We believe that developing such protocols is challenging.

# 9. REFERENCES


[1] F. Bai, N. Sadagopan, and A. Helmy. The IMPORTANT framework for analyzing the Impact of Mobility on Performance Of RouTing protocols for Adhoc NeTworks. *Ad Hoc Networks*, 1(4):383–403, 2003.

[2] C. Barrett, R. Beckman, K. Berkbigler, K. Bisset, B. Bush, K. Campbell, S. Eubank, K. Henson, J. Hurford, D. Kubicek, et al. Transims: Transportation analysis simulation system. *Los Alamos National Laboratory*, 2002.

[3] V. Bychkovsky, B. Hull, A. Miu, H. Balakrishnan, and S. Madden. A measurement study of vehicular internet access using unplanned 802.11 networks. In *ACM MOBICOM*, 2006.

[4] D. R. Choffnes and F. E. Bustamante. An integrated mobility and traffic model for vehicular wireless networks. In *VANET '05: Proceedings of the 2nd ACM international workshop on Vehicular ad hoc networks*, pages 69–78, New York, NY, USA, 2005. ACM.

[5] J. Haerri, F. Filali, and C. Bonnet. Mobility models for vehicular ad hoc networks: a survey and taxonomy. *EEE Communications Surveys and Tutorials*, 2008.

[6] W. Hsu and A. Helmy. Impact: Investigation of mobile-user patterns across university campuses using wlan trace analysis. *Arxiv preprint cs.NI/0508009*, 2005.

[7] R. Jain, D. Lelescu, and M. Balakrishnan. Model T: an empirical model for user registration patterns in a campus wireless LAN. In *Proceedings of the 11th annual international conference on Mobile computing and networking*, pages 170–184. ACM New York, NY, USA, 2005.

[8] D. Johnson and D. Maltz. Dynamic source routing in ad hoc wireless networks. In *Ad Hoc Networking*, 2001. Chapter 5, pg 139-172, Addison-Wesley.

[9] B. Karp and H. Kung. Greedy perimeter stateless routing for wireless networks. In *Mobicom'00*, Aug. 2000.

[10] M. Kim, D. Kotz, and S. Kim. Extracting a mobility model from real user traces. In *Proceedings of IEEE Infocom*, volume 6, 2006.

[11] D. Krajzewicz, M. Bonert, and P. Wagner. Sumo (simulation of urban mobility); an open-source traffic simulation. In *RoboCup2006*, 2006.

[12] J. LeBrun, C.-N. Chuah, D. Ghosal, and M. Zhang. Knowledge-based opportunistic forwarding in vehicular wireless ad hoc networks. In *Vehicular Technology Conference*, 2005.

[13] I. Leontiadis and C. Mascolo. Geopps: Opportunistic geographical routing for vehicular networks. In *Proceedings of the IEEE Workshop on Autonomic and Opportunistic Communications*, June 2007.

[14] C. Lochert, H. Hartenstein, J. Tian, H. Fussler, D. Hermann, and M. Mauve. A routing strategy for vehicular ad hoc networks in city environments. In *Intelligent Vehicles Symposium*, 2003.

[15] R. Mahajan, J. Zahorjan, and B. Zill. Understanding wifi-based connectivity from moving vehicle. In *IMC*, 2007.

[16] R. Morris, J. Jannotti, F. Kaashoek, J. Li, and D. Decouto. CarNet: A scalable ad hoc wireless network system. In *9th ACM SIGOPS European Workshop*, Sept. 2000.

[17] K. Nagel and M. Rickert. Parallel implementation of the TRANSIMS micro-simulation. *Parallel Computing*, 27(12):1611–1639, 2001.

[18] V. Naumov, R. Baumann, and T. Gross. An evaluation of inter-vehicle ad hoc networks based on realistic vehicular traces. In *Proceedings of the 7th ACM international symposium on Mobile ad hoc networking and computing*, pages 108–119. ACM New York, NY, USA, 2006.

[19] V. Naumov and T. R. Gross. Connectivity-aware routing (car) in vehicular ad hoc networks. In *IEEE Infocom*, 2007.

[20] D. Navigation. Dash express automotive navigation system, 2008.

[21] A. Oberhoffken. MIRP: Map Information Routing Protocol for mobile ad-hoc networks, 2004.

[22] C. Perkins and E. Royer. Ad hoc on-demand distance vector routing. In *Mobile Computing Systems and Applications*, Feb. 1999.

[23] N. Sadagopan, F. Bai, B. Krishnamachari, and A. Helmy. PATHS: analysis of PATH duration statistics and their impact on reactive MANET routing protocols. In *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, pages 245–256. ACM New York, NY, USA, 2003.

[24] L. Smith, R. Beckman, and K. Baggerly. Transims: Transportation analysis and simulation system. Technical report, LA-UR–95-1641, Los Alamos National Lab., 1995.

[25] C. Tuduce and T. Gross. A mobility model based on wlan traces and its validation. In *Proceedings IEEE INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, volume 1, 2005.

[26] N. Wisitpongphan, F. Bai, P. Mudalige, and O. Tonguz. On the routing problem in disconnected vehicular ad hoc networks. In *Proc. IEEE Infocom*, 2007.

[27] Y. Yu, G. Lu, and Z. Zhang. Enhancing location service scalability with HIGH-GRADE. In *2004 IEEE International Conference on Mobile Ad-hoc and Sensor Systems*, pages 164–173, 2004.

[28] X. Zhang, J. Kurose, B. N. Levine, D. Towsley, and H. Zhang. Study of a bus-based disruption tolerant network: Mobility modeling and impact on routing. In *ACM Mobicom*, Sept. 2007.

[29] J. Zhao and G. Cao. VADD: Vehicle assisted data delivery in vehicular ad hoc networks. In *IEEE Infocom*, 2006.