Edge supported real time road accident prediction

Cailean Oikawa

April 25, 2022

INTRODUCTION

VANET technology can enable the fast dissemination of data about road conditions and traffic parameters such as vehicle position, velocity, acceleration. We want to define a framework for the fast collection and sharing of data that can be consumed by machine learning algorithms in order to predict road accidents before they occur so that actions can be taken to avoid them.

We aim to characterize the necessary network and computing resources used by nodes in vehicle-to-vehicle as well as vehicle-to- infrastructure connections to support a sufficiently low average age of information which enables accurate predictions to be made by a machine learning model.

We assume that a heterogeneous network of vehicles and road side units exists to support the protocol. Vehicles are capable of collecting information about road conditions such as slippery or loose surface, low visibility and locally storing the information indexed by location and time of collection. The dissemination of information among vehicles occurs according to non-scheduled gossip events between vehicles and roadside units which aims to maintain a threshold of information freshness. Vehicles can then communicate their position,velocity,acceleration with all nearby vehicles in real time to support fast prediction of road accidents given the shared data context about specific road conditions.

We aim to provide vehicles with data at all times about roadways within a radius of 300m and freshness no older than 5 minutes.

PROTOCOL

A simple gossip protocol is implemented within the ns-3 simulator on top of a simulated 802.11p network layer. Position in the network is discretized into 15m x 15m tiles of road surface which all vehicles measure passively for local storage. All transmissions are broadcasts on a shared single channel and consist of either tile-requests or tile-replies. According to a fixed period vehicles compute an interest vector of tiles within a given radius of their position and broadcast a request for any tiles for which local age of information is beyond a threshold limit. Any vehicle receiving a request will consult local datastore to determine whether or not fresher information can be transmitted. If a vehicle can reply it will schedule a transmission with a certain delay linearly proportional to how much fresher the local data is to the threshold requested. This ensures that for multiple vehicles receiving a request if multiple can reply then the reply with the best data will be transmitted first allowing duplicate reply transmissions to be automatically aborted and reduce redundant network traffic.

$$Delay = AOI_{local}/AOI_{request}$$

All vehicles picking up a reply will compare the age of information of new data with that of local and keep the freshest data. A supporting infrastructure participates in the protocol as a single node with multiple stationary access points.

EXPERIMENT

Two settings are considered, an urban center with low traffic speed and many intersections, and a section of two lane highway with high speed and no intersections. The discretized tile map of both the highway and urban scenario are shown bellow - the scaled width of the page is about 16km. Traffic traces are generated in SUMO using sections of road taken from OpenStreetMaps. All traces are generated specifying a through traffic factor of 5 cars per kilometer per hour. The highway scenario runs for 2300 seconds and simulates a total of 113 vehicles travelling in both directions. The urban center scenario runs for 1150 seconds and simulates a total of 154 vehicles travelling along random short routes. Infrastructure access points are placed at random tiles such that no two access points are closer than a specified distance **d**. This ensures that maximum distance between any two access points is **2d** and minimum distance is **d**. The parameters altered over subsequent simulations are 1. frequency of request for tiles, 2. radius over which tiles are requested, 3. infrastructure density. Performance data are collected from the simulation in order to determine how well the network succeeds or fails in different scenarios - each nodes local datastore is examined every 2.5 seconds in order to determine if the network successfully disseminated required data to it. A failure to disseminate to a node is defined as a node not having data for tiles within 300m fresher than 5 minutes while data fresher than 5 minutes does exist elsewhere in the network for these tiles. We measure the overall performance of the network as the ratio between failures and successes across all nodes each node examined every 2.5 seconds.

A network load is characterized as total bytes transmitted over the channel divided by duration of simulation in seconds divided by number of nodes participating including vehicles and infrastructure access points.

RESULTS

Results are good for the highway scenario, despite fairly sparse traffic data is disseminated well with relatively low infrastructure density. Results are fairly bad in the urban scenario even with very dense infrastructure - the network here fails to reach anything below a 30 percent failure rate even with very high infrastructure density and very frequent tile requests. Examining the highway simulations there are clear diminishing returns on all parameters - many settings of the simulation provide failure rates well below 5 percent but no scenario manages to achieve better than 1 percent which is not surprising given the simplicity of the protocol.

We can calculate a notion of efficiency by calculating the success rate divided by the network load in bytes / duration / nodes. We can see that the top efficiency parameter setting that give failure rates lower than 5 percent is request radius 2000m requesting every 15 seconds with infrastructure gaps of 1500-3000m. We could argue that it is not worth the investment in physical infrastructure to build more access points in order to shave off a few percent from the error rate.

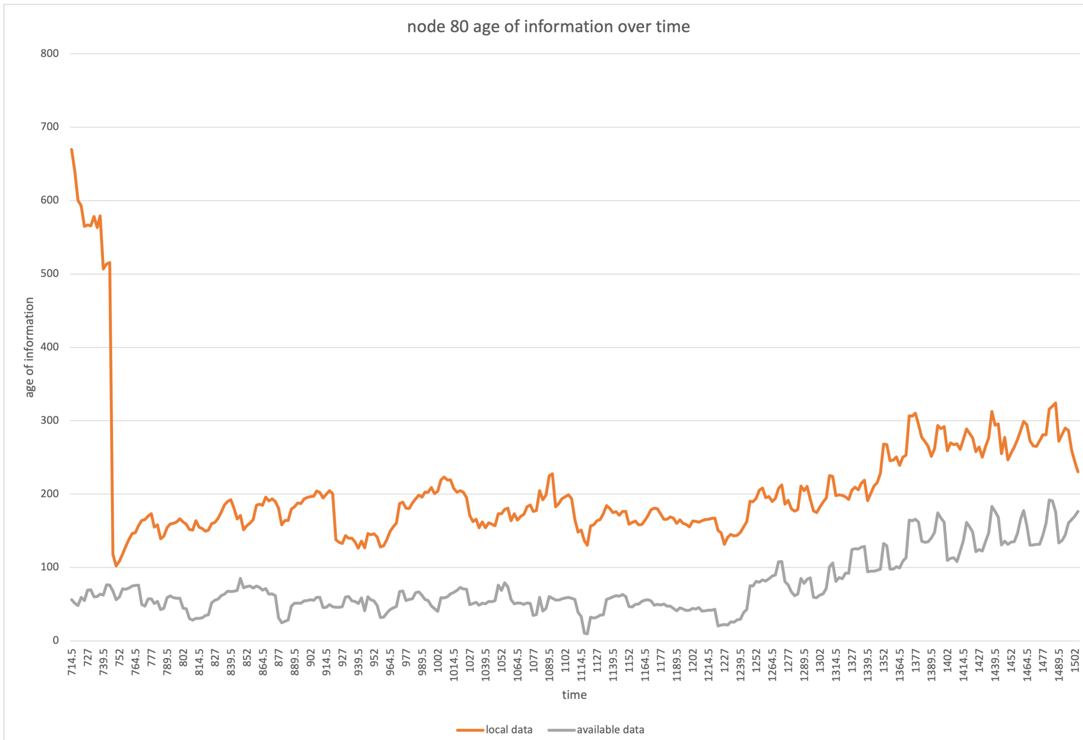
Examining charts of age of information over time for one node in each scenarios which are representative of network behaviour overall it can be seen that in the highway setting the high frequency of data collection and gossip interactions leads to very low failure rate but in the urban setting data is not disseminated quickly enough. Shown below are node 70 in the highway scenario with radius of request 2000 requesting every 10 seconds infrastructure gap between 1500-3000m, node 80 in the urban scenario with radius of request 4000 requesting every 5 seconds infrastructure gap between 200-400m. Both nodes examined have local data steadily within a few multiples of the best available data at all times and this pattern is not particular to these two nodes. This suggests that the performance of the network can be improved by increasing infrastructure density / rate of request / radius of request but the peak performance of the network disseminating fresh information is limited by the rate of physical world rendezvous between nodes and the rate at which each tile in the network is measured.

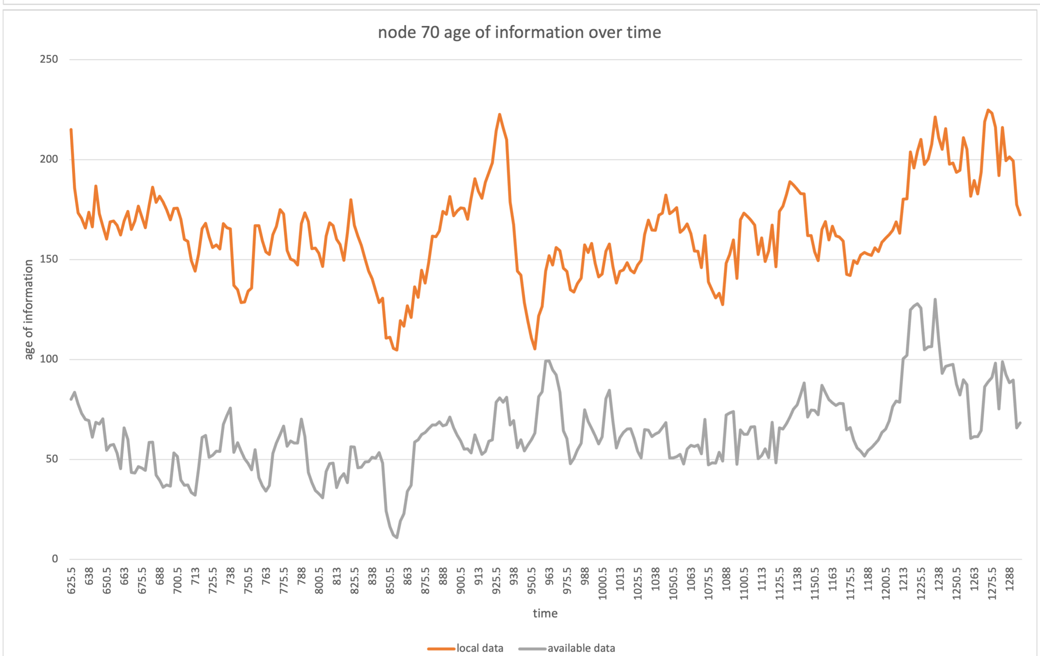
The selection of good parameters for a gossip protocol oriented data sharing between vehicles about road conditions depends most importantly on the characteristics of the real world traffic and for certain settings this approach does not provide the possibility of hitting certain age of information targets, here other approaches might be more suited. The setting of a highway style roadway is very conducive to a gossip protocol based approach as the traffic characteristic leads to high frequency of tile measurement and gossip interaction. Even with sparse infrastructure this approach seems feasible.

highway				
radius	period	id. density	failure rate (300m)	bytes / duration / nodes
1000	15	1500	0.1579	51.98876578
1000	15	1000	0.07612	60.60896213
1000	15	500	0.02251	71.94014294
1000	10	1500	0.1124	66.20267882
1000	10	1000	0.04836	75.17771237
1000	10	500	0.02622	96.42126266
1000	5	1500	0.08794	104.5581206
1000	5	1000	0.03469	121.4983813
1000	5	500	0.01674	162.0082311
2000	15	1500	0.03459	60.85043478
2000	15	1000	0.01842	65.14240803
2000	15	500	0.01165	77.03964264
2000	10	1500	0.03494	77.0757784
2000	10	1000	0.01688	83.21662876
2000	10	500	0.01407	101.8057534
2000	5	1500	0.02713	120.195582
2000	5	1000	0.01544	134.5840936
2000	5	500	0.01176	172.1769625
4000	15	1500	0.02127	81.4297195
4000	15	1000	0.01376	81.9553311
4000	15	500	0.009377	90.36776653
4000	10	1500	0.02435	203.9856381
4000	10	1000	0.01547	104.3081204
4000	10	500	0.01129	117.2373198
4000	5	1500	0.01969	156.5058205
4000	5	1000	0.01504	165.8231706
4000	5	500	0.01078	195.6643359

urban				
radius	period	id. density	failure rate (300m)	bytes / duration / nodes
1000	15	1500	0.6259	52.52205128
1000	15	1000	0.5927	61.0530765
1000	15	500	0.5721	87.00106456
1000	10	1500	0.5535	71.08298774
1000	10	1000	0.5437	83.39174463
1000	10	500	0.4857	118.2531752
1000	5	1500	0.4991	122.5426533
1000	5	1000	0.4821	144.2207595
1000	5	500	0.4538	211.3893439
2000	15	1500	0.6183	76.12296544
2000	15	1000	0.5885	85.0955421
2000	15	500	0.5383	109.7436627
2000	10	1500	0.5341	103.8806466
2000	10	1000	0.5401	115.0380187
2000	10	500	0.4997	149.8579605
2000	5	1500	0.5021	181.4525753
2000	5	1000	0.5204	201.490809
2000	5	500	0.4544	261.8841634
4000	15	1500	0.6134	83.75204013
4000	15	1000	0.5848	92.21979086
4000	15	500	0.5529	116.3378972
4000	10	1500	0.5544	113.7513266
4000	10	1000	0.5477	124.8972812
4000	10	500	0.5115	158.469776
4000	5	1500	0.5091	201.2288199
4000	5	1000	0.5176	217.266131
4000	5	500	0.4662	277.3938972
4000	5	200	0.3641	563.8309835

highway					
radius	period	id. density	failure rate (300m)	bytes / duration / nodes	eff.
4000	10	1500	0.02435	203.9856381	0.004782935
4000	5	500	0.01078	195.6643359	0.005055699
2000	5	500	0.01176	172.1769625	0.005739676
4000	5	1000	0.01504	165.8231706	0.005939821
1000	5	500	0.01674	162.0082311	0.006069198
4000	5	1500	0.01969	156.5058205	0.006263729
2000	5	1000	0.01544	134.5840936	0.007315575
1000	5	1000	0.03469	121.4983813	0.007945044
2000	5	1500	0.02713	120.195582	0.008094058
4000	10	500	0.01129	117.2373198	0.008433407
1000	5	1500	0.08794	104.5581206	0.008722995
4000	10	1000	0.01547	104.3081204	0.009438671
2000	10	500	0.01407	101.8057534	0.009684423
1000	10	500	0.02622	96.42126266	0.010099225
4000	15	500	0.009377	90.36776653	0.010962128
2000	10	1000	0.01688	83.21662876	0.011813985
4000	15	1500	0.02127	81.4297195	0.012019322
4000	15	1000	0.01376	81.9553311	0.012033872
2000	10	1500	0.03494	77.0757784	0.012520924
1000	10	1000	0.04836	75.17771237	0.012658539
2000	15	500	0.01165	77.03964264	0.012829109
1000	10	1500	0.1124	66.20267882	0.013407312
1000	15	500	0.02251	71.94014294	0.013587546
2000	15	1000	0.01842	65.14240803	0.015068218
1000	15	1000	0.07612	60.60896213	0.01524329
2000	15	1500	0.03459	60.85043478	0.015865293
1000	15	1500	0.1579	51.98876578	0.01619773


node 80 age of information over time
age of information
800
700
600
500
400
300
200
100
0
time
local data
available data


node 70 age of information over time
age of information
250
200
150
100
50
0
time
local data
available data