Social Welfare-based Optimization for Data/Service Delivery to Connected Vehicles via Edges

Deepak Gangadharan, Oleg Sokolsky, Insup Lee
University of Pennsylvania
and
BaekGyu Kim
Toyota InfoTechnology Center

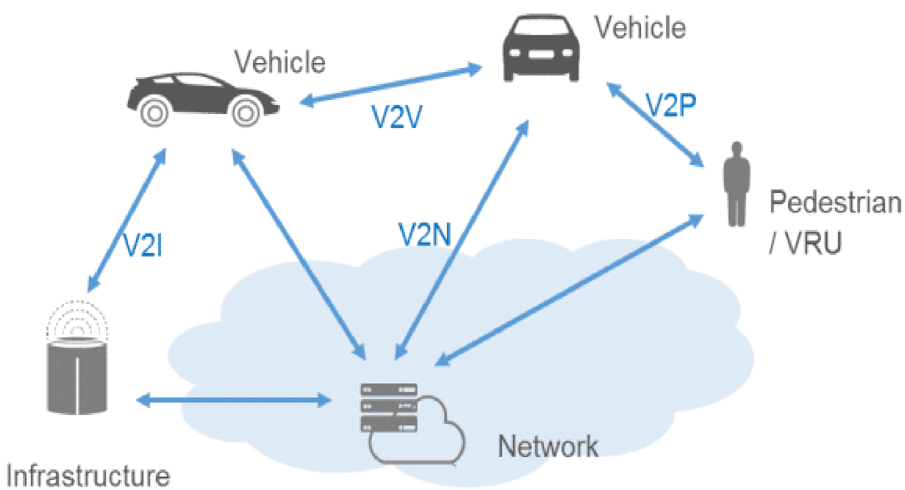
Penn Engineering PRECISE

TOYOTA InfoTech Envisioning Mobility

Motivation: Increased Connectivity For Automotives

- Vehicle connectivity with various entities
  - to other vehicles → V2V
  - to infrastructure → V2I
  - to cloud → V2N/V2C
  - to pedestrian → V2P

- Potential for efficient delivery of data/services!



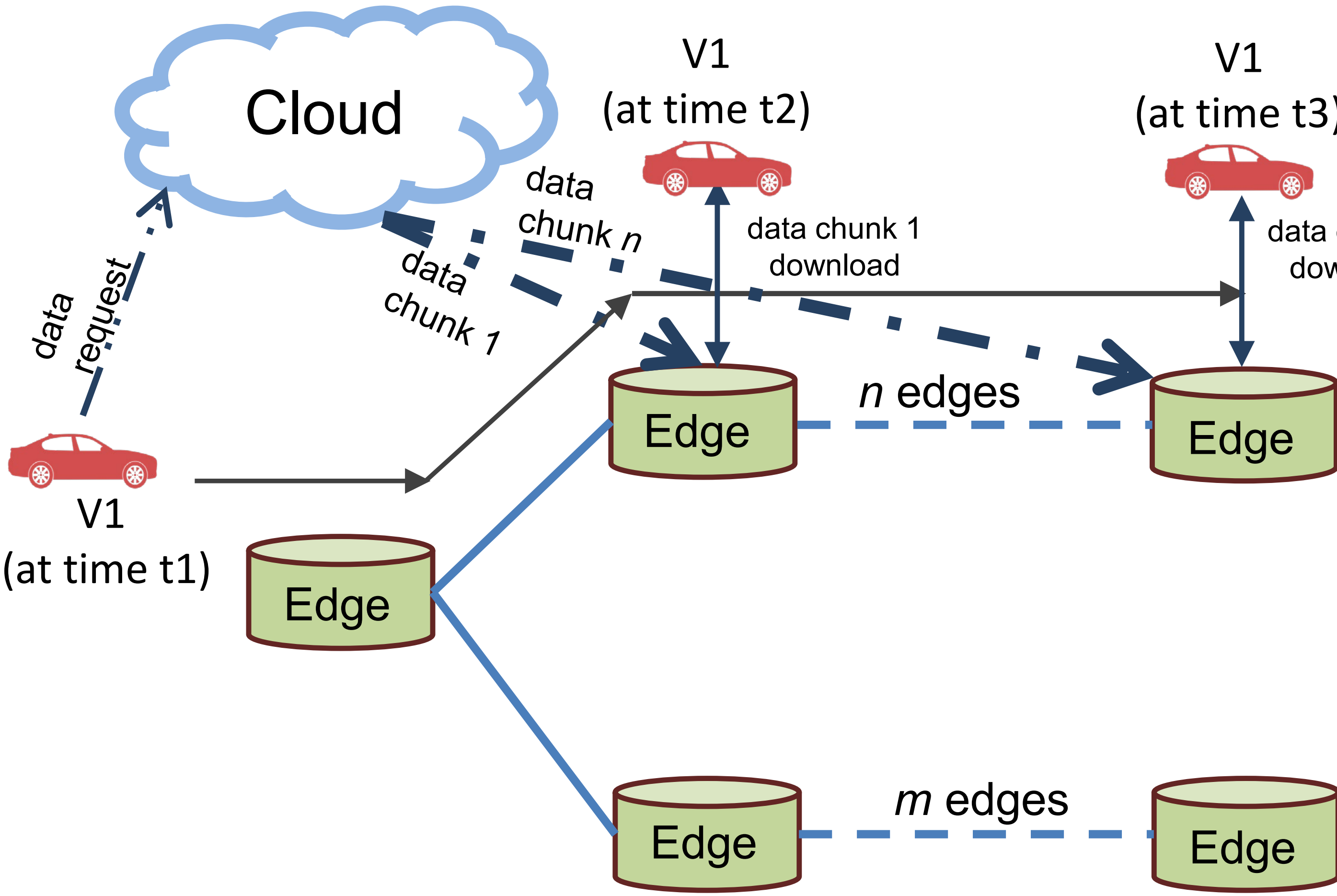
Motivation: U.S. Government Policy Trend

- USDOT *Connected* Vehicle Pilot Projects (~$42 million)
  - New York City DOT Pilot
  - Tampa-Hillsborough Expressway Authority Pilot
  - Wyoming DOT Pilot
- New York City DOT Pilot
  - (**Goal**) Eliminate traffic related deaths and reduce crash
  - (**Approach**) Exchange information about hazardous situations using Vehicle-to-Vehicle (V2V), Vehicle-to-Infrastructure (V2I), Infrastructure-to-Pedestrian (IVP) communications

*Approx. 300 road side units (RSU) will be installed in NY city

Source: www.cvp.nyc

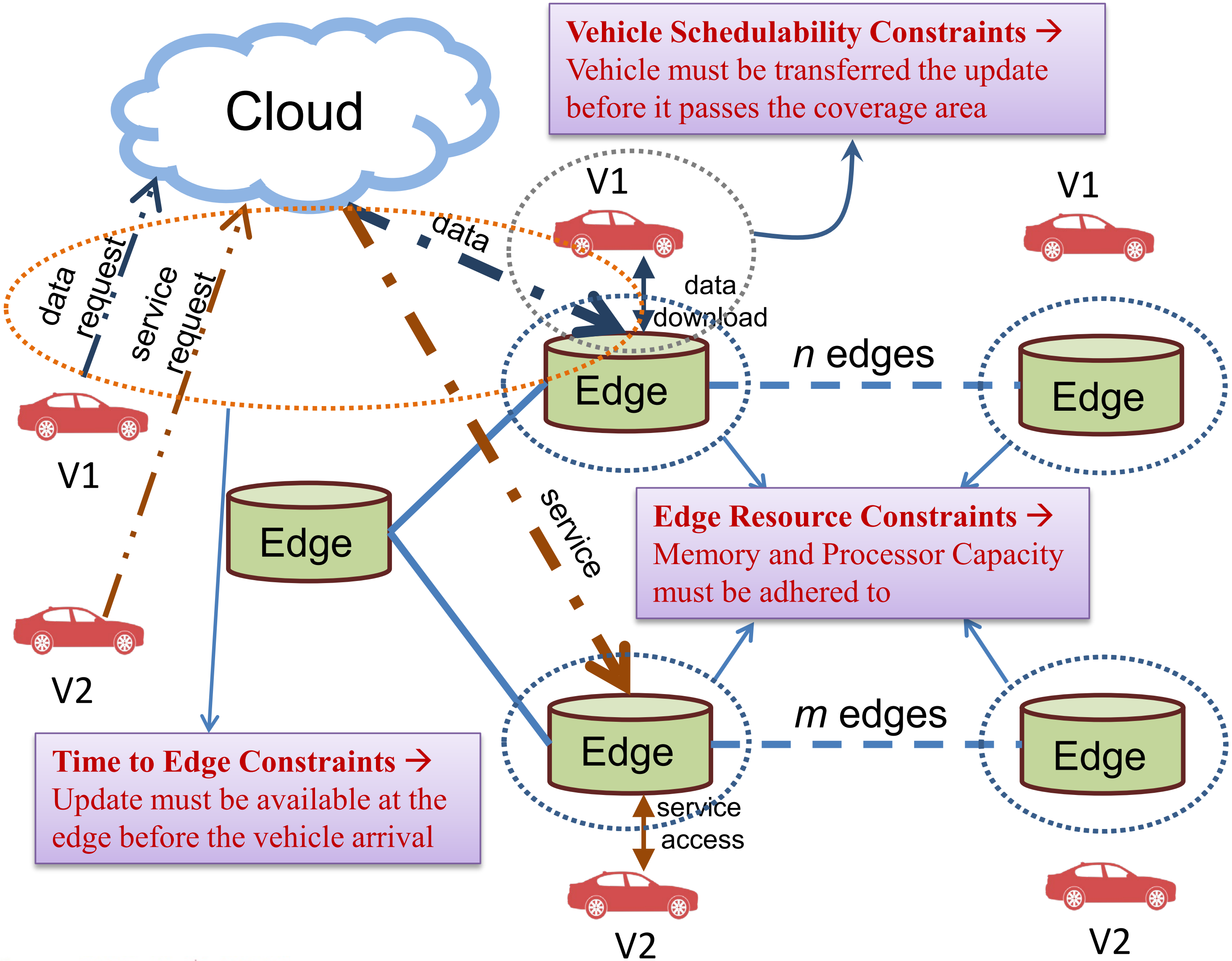
Motivation: Data Delivery Scenario



Data Delivery

- Vehicle V1 requests the cloud for some data such as map data
- Cloud determines the appropriate edge to send the data
- V1 downloads the data chunks from the respective edges as it passes by them
- **Problem: Necessary to determine the appropriate edges to push the update chunks in order to optimize a user/system objective**

Problem Scenario – Potential Constraints



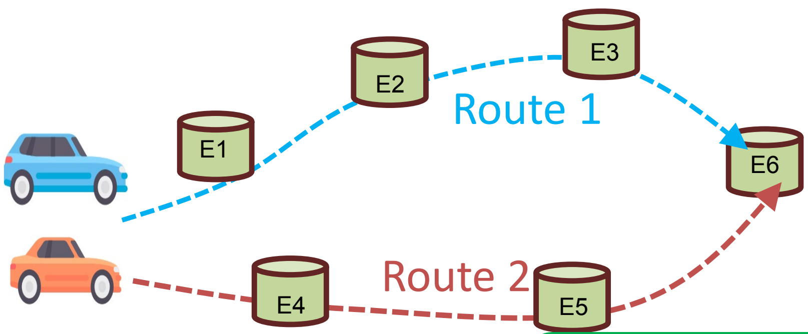
Vehicle Schedulability Constraints → Vehicle must be transferred the update before it passes the coverage area

Edge Resource Constraints → Memory and Processor Capacity must be adhered to

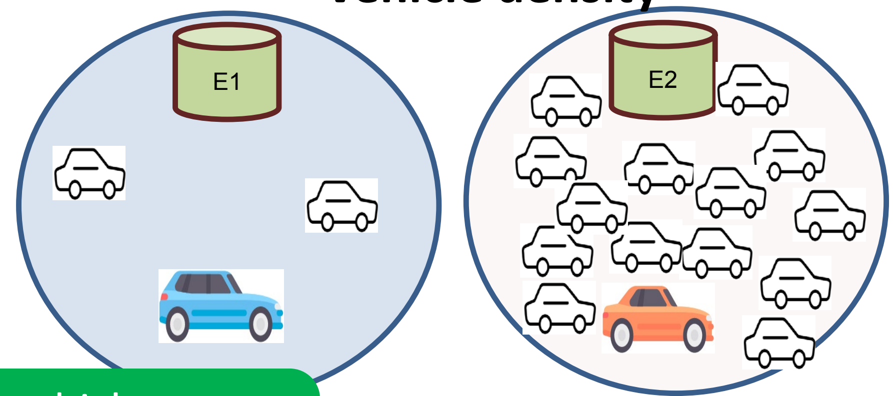
Time to Edge Constraints → Update must be available at the edge before the vehicle arrival

Problem Scenario – Influencing Factors

Vehicle routes

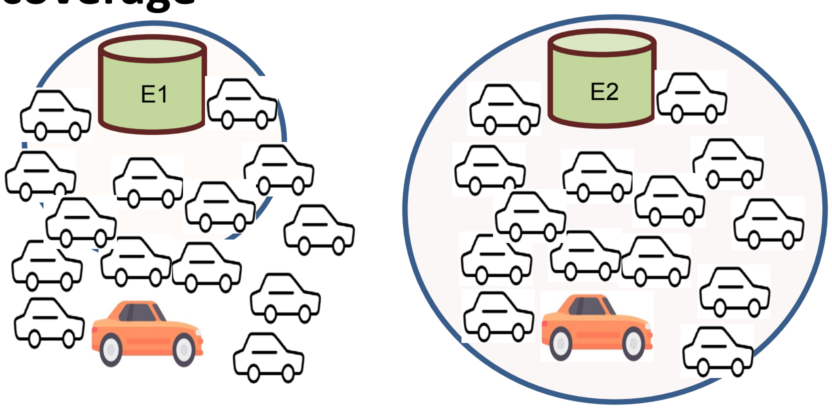


Vehicle density

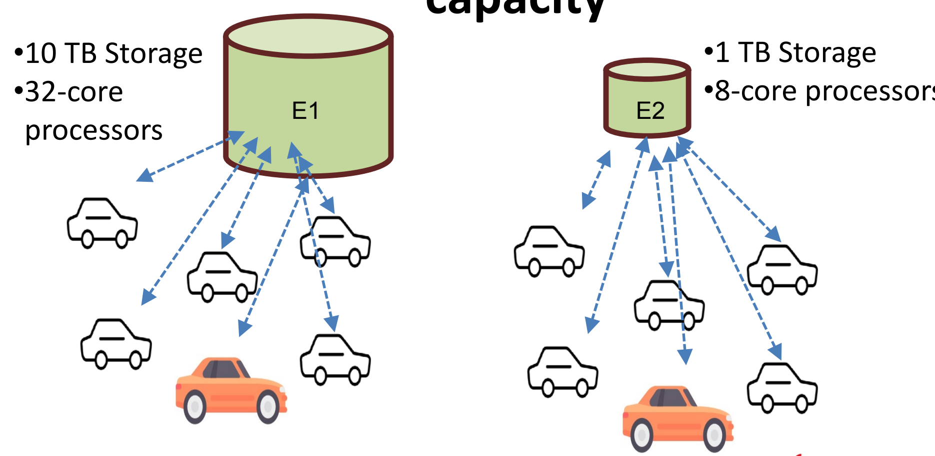


Can a group of vehicles successfully get data/services on their routes?

Field communication coverage



Memory/computing capacity



Optimization Problem Formulation – Constraints for Data/Service Delivery

- Problem size parameters
  N – Number of vehicles, *M* – Number of edges

- Route indicator
  $x_{i,j}$ – Binary value indicating whether vehicle *i* passes through edge *j*
  (1 indicates that it passes and 0 indicates that it doesn't)

- Decision Variable
  $m_{i,j}$ – Memory chunk of vehicle *i* allocated to edge *j*
  $serv_{i,j}$ – Binary decision variable denoting vehicle *i* receives its service from edge *j*

Penn Engineering PRECISE

TOYOTA InfoTech Envisioning Mobility

Optimization Problem Formulation – Constraints for Data/Service Delivery

- <u>Range Constraint</u>
  m_app_i – Memory requirement for vehicle *i* update

$$\left.\begin{array}{l} m_{i,j} \geq 0 \\ m_{i,j} \leq m\_app_i \end{array}\right\} \forall i,j: x_{i,j} = 1$$

Memory chunk allocated must take a value between 0 and m_app_i both values included

- <u>Edge Memory Constraint (Update Delivery)</u>
  m_occ_j - Size of memory currently occupied on edge *j*
  m_e_j - Total memory size of edge *j*

$$\sum_{i=1}^{N} x_{i,j} * m_{i,j} + m\_occ_j \leq m\_e_j \ , \forall j = 1 \ to \ M$$

Sum of all memory chunks for all vehicles on edge j cannot exceed the available memory

Penn Engineering PRECISE

TOYOTA InfoTech Envisioning Mobility

Optimization Problem Formulation – Constraints for Data/Service Delivery

- <u>Edge Memory Constraint (Service Delivery)</u>
  d_i- Size of input data from vehicle to edge
  r_i- Size of output data from edge to vehicle

$$\sum_{i=1}^{N} serv_{i,j} * (d_i + r_i) + m\_occ_j \leq m\_e_j \;, \forall j = 1 \; to \; M$$

  Sum of memory required for input and output data for all vehicles on edge j cannot exceed the available memory

- <u>Edge Computation Constraint (Service Delivery)</u>
  P_occ_j - Processing resources already in use on edge *j* (can be number of processing units in time or number of VMs available)
  P_j - Total processing capacity of edge *j*
  p_i - Processing resource requirement of vehicle *i*

$$\sum_{i=1}^{N} serv_{i,j} * p_i + P\_occ_j \leq P_j \;, \forall j = 1 \; to \; M$$

Penn Engineering PRECISE

TOYOTA InfoTech Envisioning Mobility

Optimization Problem Formulation – Constraints for Data/Service Delivery

- Accumulation Constraint

$$\sum_{j=1}^{M} x_{i,j} * m_{i,j} = m\_app_i, \forall i = 1\ to\ \mathrm{N}$$

Sum of memory chunks for a vehicle over all edges should be equal to the update size

- Time to Edge Constraint

$t\_comm_{i,j}$ – Time required to send update data to the edge *j* after vehicle *i* initiates update download

$t\_trv_{i,j}$ – Time required to travel to edge *j* by vehicle *i* after initiating the download

$$m_{i,j} * t_{comm_{i,j}} \leq m_{i,j} * t_{trv_{i,j}}, \forall i, j$$

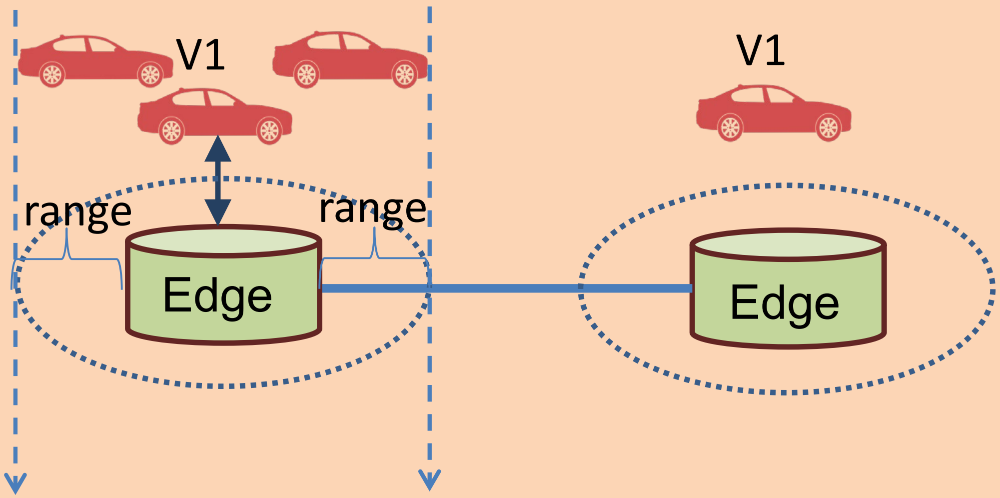
Update chunks must be available at the edges before the vehicle arrives at the edges

Penn Engineering PRECISE

TOYOTA InfoTech Envisioning Mobility

Optimization Problem Formulation – Constraints for Data/Service Delivery

- Bandwidth Schedulability Constraint



Macroscopic Flow Model

k – vehicle density, i.e., number of vehicles per unit distance
q – vehicle flow, i.e., number of vehicles passing a fixed point per unit time
v – Vehicle speed

$$q = k \times v$$

Optimization Problem Formulation – Constraints for Data/Service Delivery

- Bandwidth Schedulability Constraint
  - Speed Density Relationship
    $v = v_f \times (1 - \frac{k}{k_{jam}})$,
    where v_f is the free flow speed (or speed limit) and k_{jam} is the vehicle density during a traffic jam
  - Vehicles in the coverage area can be modelled as a queuing network, specifically in this scenario as a M/D/C queue (poisson arrival, deterministic service and C servers)
  - As per M/D/C queue modelling, the minimum number of bytes received by a vehicle D_{min}= $B/(k_{jam} \times v)$ [1], where B is the bandwidth of the edge device
  - **The bandwidth schedulability constraint for vehicle requiring update is**
    $$\boldsymbol{m_{i,j} \leq D_{min}}$$

[1] Analytical Models and Performance Evaluation of Drive-thru Internet Systems, IEEE JSAC 2011

Optimization Problem Formulation – Constraints for Data/Service Delivery

- Bandwidth Schedulability Constraint
  - For service delivery, as per M/D/C modelling, the minimum number of bytes received by a vehicle is

$$D_{min}^{serv} = \frac{\mathrm{B} \times \left(\frac{L}{v} - t_{p_i}\right)}{k_{jam} \times L}$$

    where L is the coverage distance, t_{p_i} is the time taken to execute a service for vehicle *i* with p_i number of VMs

  - **The bandwidth schedulability constraint for vehicle requiring a service is given by**

$$\boldsymbol{serv_{i,j} * (d_i + r_i) \leq D_{min}^{serv}}$$

Penn Engineering PRECISE

TOYOTA InfoTech Envisioning Mobility

Edge Bandwidth Utilization

- Total bandwidth utilization can then be calculated as

$$bw^{util} = bw^{util,data} + bw^{util,serv} + bw^{util,oth}$$

where

$bw^{util,data}$ - Bandwidth utilization of the vehicles requiring data delivery

$bw^{util,serv}$ - Bandwidth utilization of the vehicles requiring service delivery

$bw^{util,oth}$ - Bandwidth utilization of the vehicles that do not require service or data delivery

Optimization Objective

- <u>Congestion Cost Objective</u>

  Minimization of total bandwidth cost considering all edges

  - Motivation : In order to avoid congestion of vehicles at an edge, a cost must be determined in utilizing the bandwidth resource given by $bw_j^{cost} = \beta(1 + bw_j^{util})^2$ [2], where bw_j^{util} is the bandwidth utilization on edge *j* and β is the bandwidth cost factor

  - **The optimization objective is**

$$\boldsymbol{minimize \sum_{j=1}^{M} bw_j^{cost}}$$

[2] On-Demand Bandwidth Pricing for Congestion Control in Core Switches in Cloud Networks, IEEE CLOUD 2016

Social Welfare Function for Data/Service Delivery in Connected Vehicles

- Two functions used to compute social welfare – **Utility function** and **Congestion Cost function**

- **Utility Function**
  - A function of the data allocated for a vehicle on the edge given by $U_{i,j}(m_{i,j})$
  - A concave function is used such that the utility increases as the allocated data increases
  - $\log(1 + m_{i,j})$ is the preferred choice based on earlier works

- In order to capture the effect of faster delivery, we must use different utility functions at different edges for a vehicle

  Reason – The utility value must decrease across edges on a route for vehicles which need faster delivery → So $U_{i,j}(m_{i,j})$ scales down as we go from one edge to the next

Social Welfare Function

- **Congestion Cost Function** – Same as the bandwidth congestion cost function used earlier denoted by $C_j(bw_{util,j})$ → Is a convex function

- Social Welfare Function

$$S(B) = \sum_{i \epsilon V} \sum_{j \epsilon E} U_{i,j}(m_{i,j}) - \sum_{j \epsilon E} C_j(bw_{util,j})$$

where $U_{i,j}(m_{i,j})$= $\log(1 + m_{i,j}) \times 1/(s \times t_{trv_{i,j}})$ and s is the scaling factor determining the utility of the edge

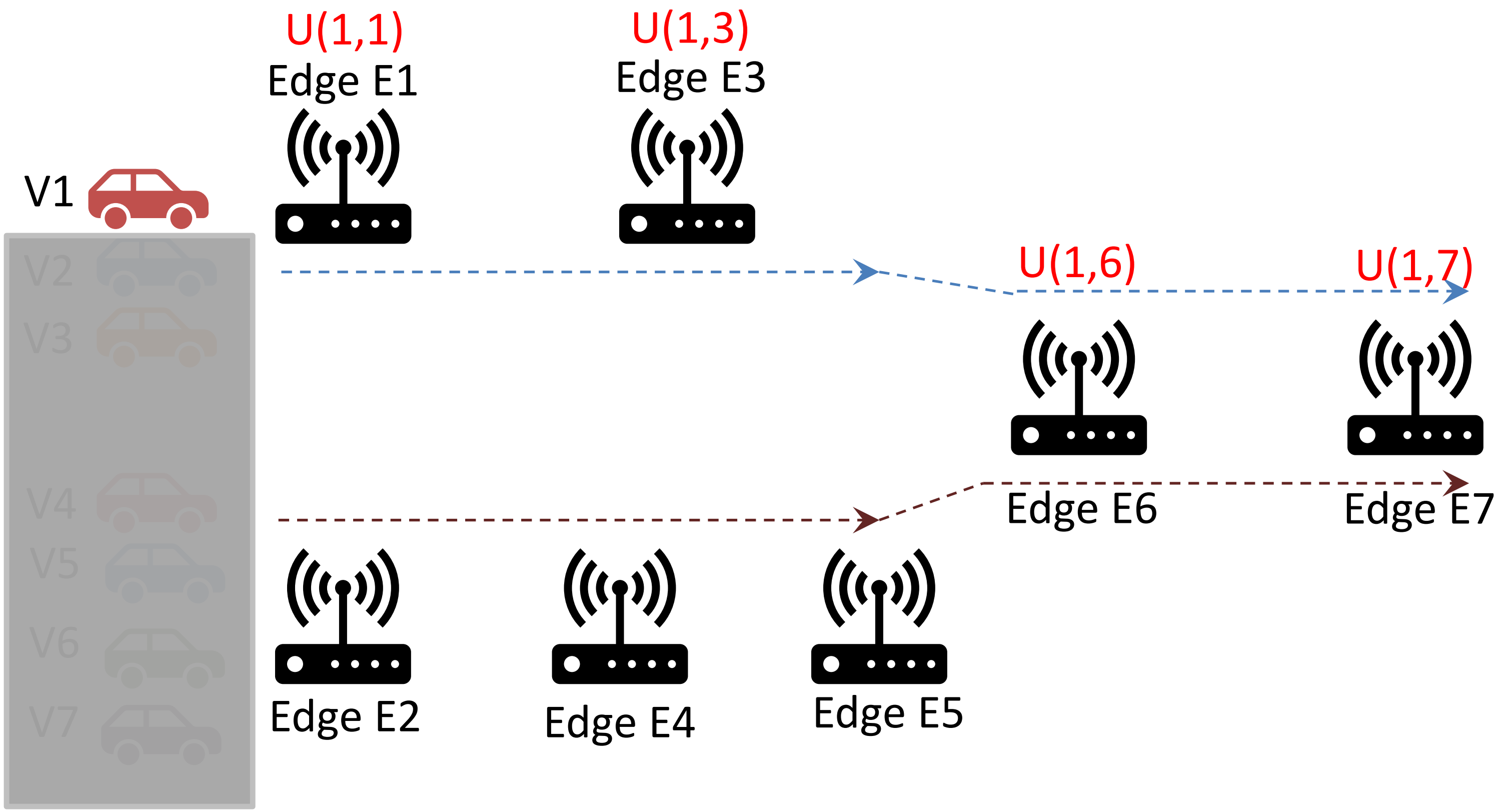
- It is a concave function and the objective function will be to **maximize $\boldsymbol{S(B)}$**

Penn Engineering PRECISE

TOYOTA InfoTech Envisioning Mobility

What does it mean?

Utility Values

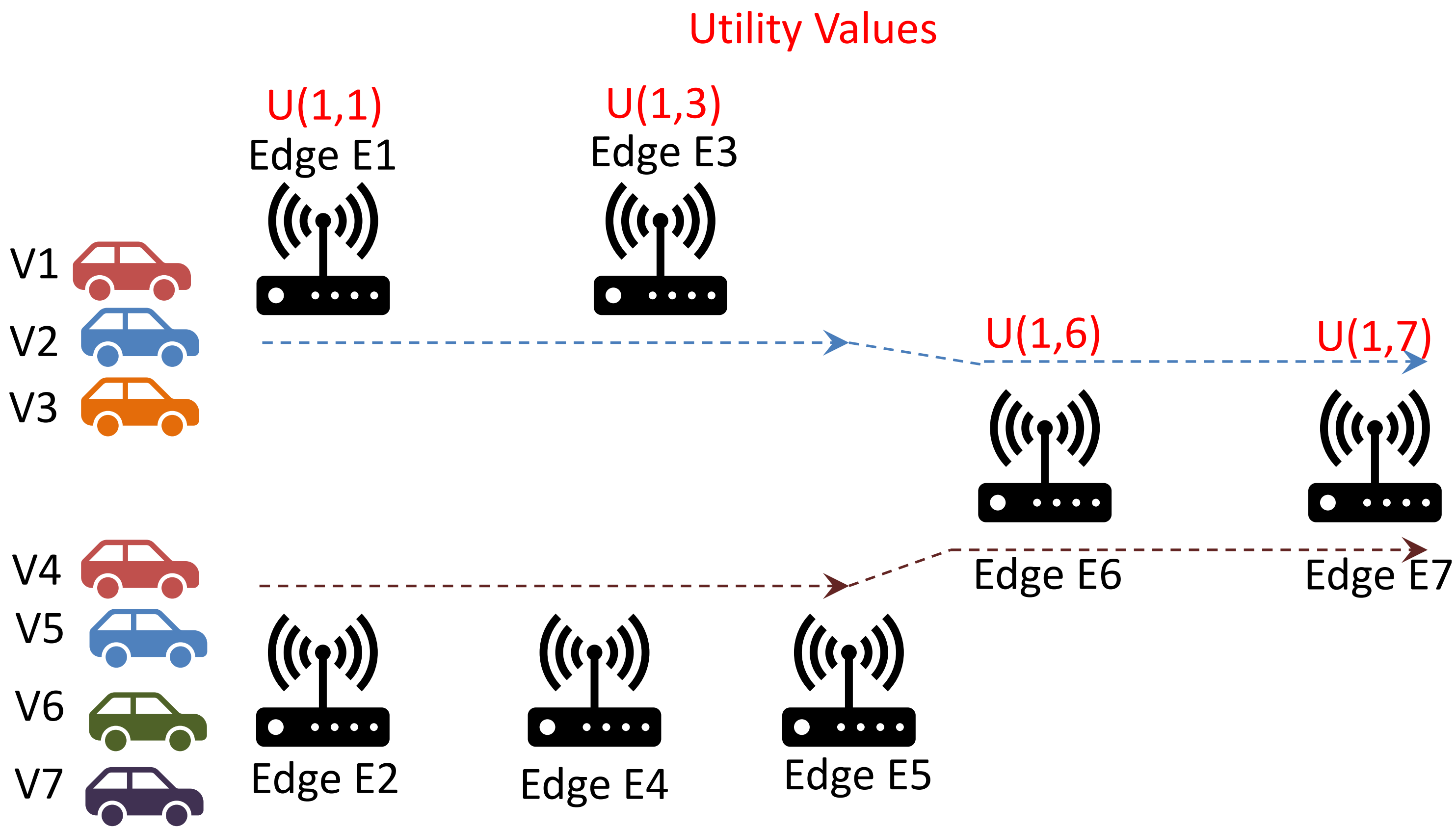


Utility value $U(1{,}1) = \frac{1}{t(1{,}1)} * \log\{1 + m(1{,}1)\}$, where

$t(1{,}1)$ - earliest time when V1 will reach E1, and

m$(1{,}1)$ – Data allocated to V1 on E1

What does it mean?

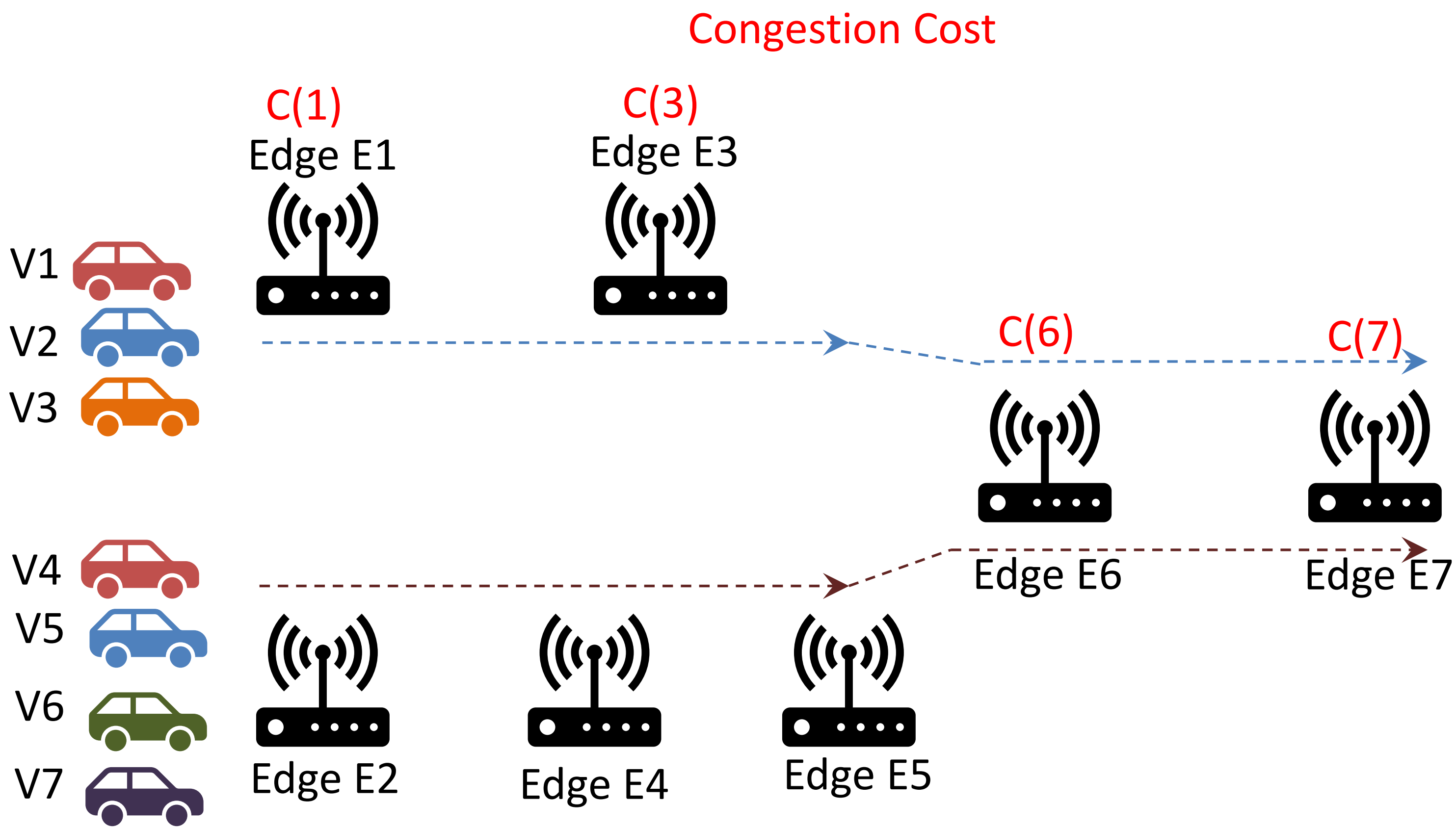


Total Utility for TU(V1) = U(1,1) + U(1,3) + U(1,6) + U(1,7)

Similarly total utility is calculated for each vehicle.

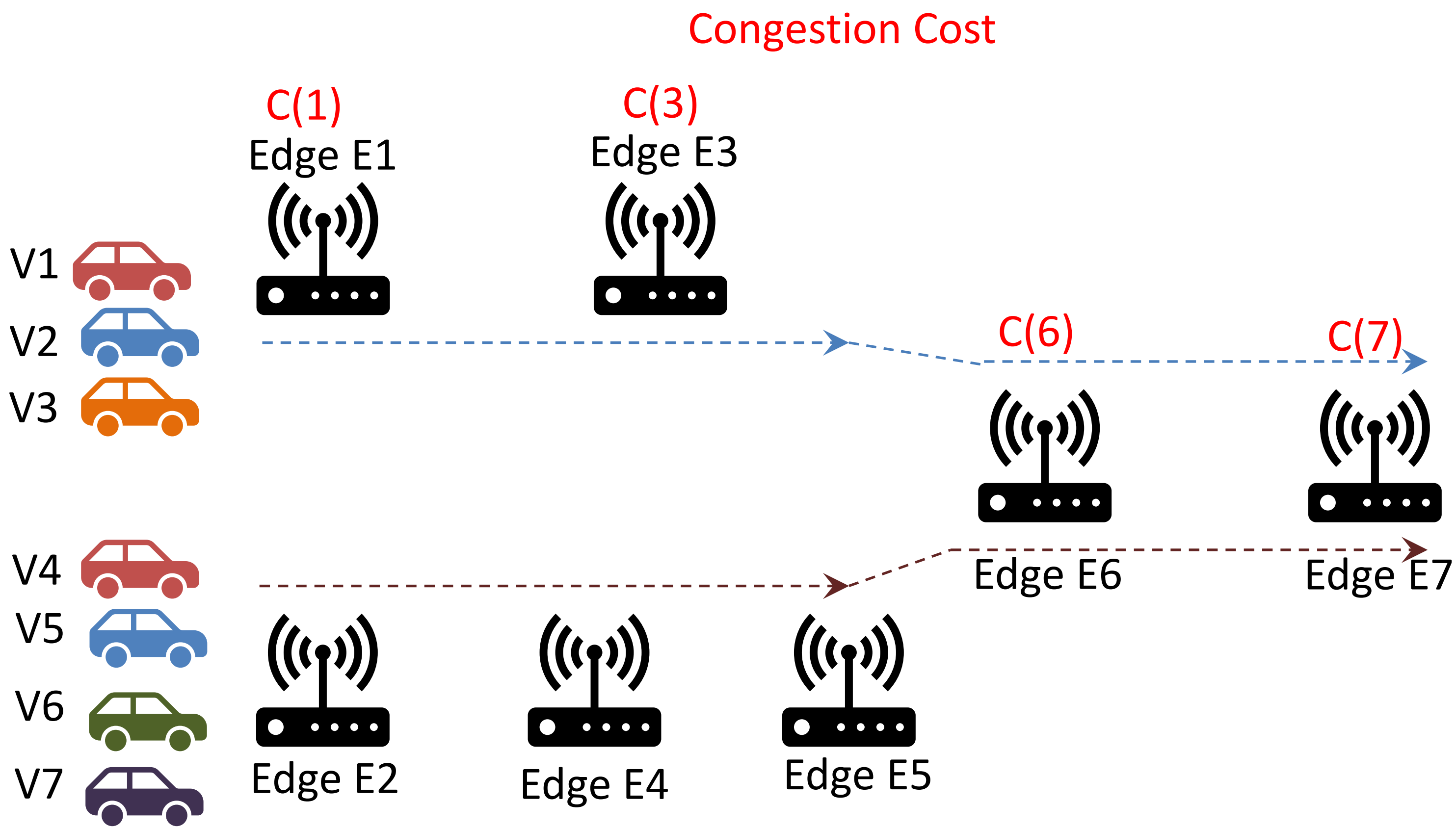
System Utility = $\sum_{\forall V_i \epsilon \, vehicles} TU(V_i)$

What does it mean?



Congestion Cost on Edge E1 $C(1) = \beta * (1 + BU(1))^2$, where
β – Congestion cost factor, and
$BU(1)$ – bandwidth utilization due to V1, V2 and V3 on E1

What does it mean?



Total System Congestion Cost $= \sum_{j \epsilon edges} C(j)$

Social Welfare Function = Total System Utility – Total System Congestion Cost

Experimental Results

- <u>Experimental Setting 1</u>

  - Number of edges (M) = {25,49,81,100}

  - Number of vehicles (N) requiring update was varied between 20 to 200

  - Vehicle jam density (k_{jam}) was randomly generated between 40 and 50 at all edges

  - The actual vehicle density in the coverage area was varied between 35 and 25 in steps of 5

  - Coverage distance (L_j) was assigned arbitrary values between 0.6 miles and 1.6 miles

  - Memory requirement of data requested by each vehicle was randomly generated between 60 and 80 Mbits

  - Memory capacity of the edge was randomly generated between 400 and 500 Mbits

  - Maximum bandwidth capacity of the edges was randomly generated between 8 and 15 Mbps

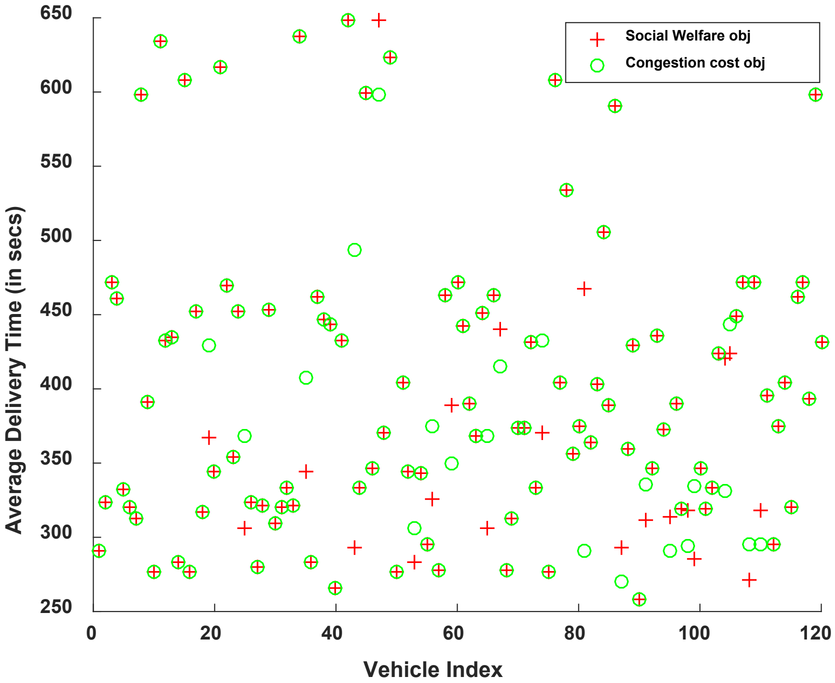
  - Free flowing velocity of the vehicles at the edges was randomly generated between 50 and 70 mph

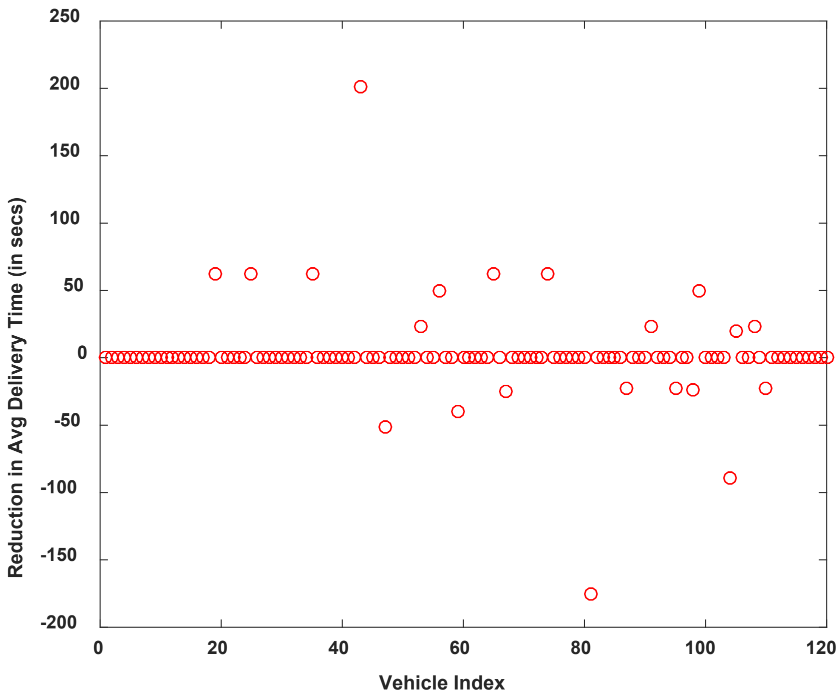
  - The route of the vehicles was randomly generated by picking connected edges in a grid

Penn Engineering PRECISE PENN RESEARCH IN EMBEDDED COMPUTING AND INTEGRATED SYSTEMS ENGINEERING

TOYOTA InfoTech Envisioning Mobility

Result 1: Data retrieval time using two objective functions

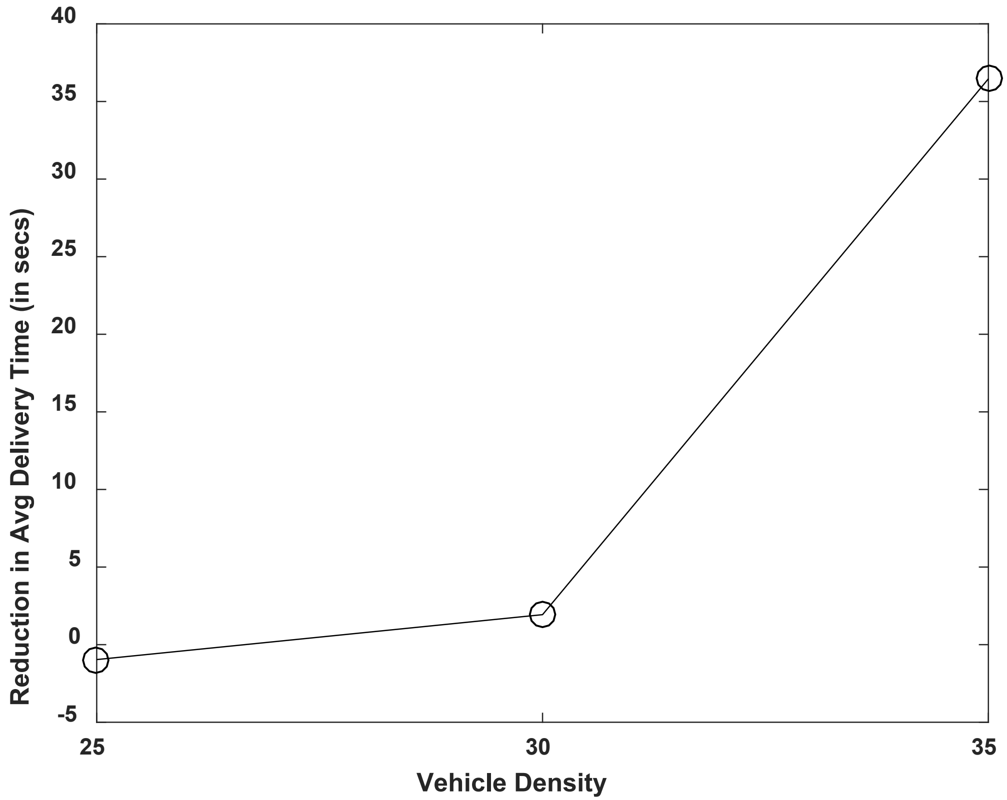




Penn Engineering PRECISE

TOYOTA InfoTech Envisioning Mobility

Result 2: Variation in Delivery Time Reduction with Vehicle Densities



- At lower vehicle densities, the average delivery time considering all vehicles is marginally higher
- As vehicle densities increase, the social welfare objective results in higher reduction in average delivery time
- For vehicle density = 35, the reduction is 36.46 secs

Conclusion

- Proposed an optimization framework for data/service delivery to connected vehicles considering vehicle flow densities
- Proposed a social welfare based delivery
- Demonstrated significant reductions in average delivery times with increasing vehicle densities

Thank You!

Questions