



ELSEVIER

Contents lists available at ScienceDirect



Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Optimization models to characterize the broadcast capacity of vehicular *ad hoc* networks

Lili Du^{a,*}, Satish Ukkusuri^b, Wilfredo F. Yushimoto Del Valle^b, Shivkumar Kalyanaraman^c^a Department of Decision Sciences and Engineering Systems, Rensselaer Polytechnic Institute, Troy, New York, USA^b Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute, Troy, New York, USA^c Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, New York, USA

ARTICLE INFO

Article history:

Received 22 March 2008

Received in revised form 10 July 2008

Accepted 16 July 2008

Available online xxxx

Keywords:

Vehicular *ad hoc* networks

ATIS

Broadcast capacity

Information flow

Optimization

Integer program

ABSTRACT

Broadcast capacity of the entire network is one of the fundamental properties of vehicular *ad hoc* networks (VANETs). It measures how efficiently the information can be transmitted in the network and usually it is limited by the interference between the concurrent transmissions in the physical layer of the network. This study defines the broadcast capacity of vehicular *ad hoc* network as the maximum successful concurrent transmissions. In other words, we measure the maximum number of packets which can be transmitted in VANET simultaneously, which characterizes how fast a new message such as a traffic incident can be transmitted in VANET. Integer programming (IP) models are first developed to explore the maximum number of successful receiving nodes as well as the maximum number of transmitting nodes in VANET. The models embed an traffic flow model in the optimization problem. Since IP model cannot be efficiently solved as the network size increases, this study develops a statistical model to predict the network capacity based on the significant parameters in the transportation and communication networks. MITSIMLab is used to generate the necessary traffic flow data. Response surface method and linear regression technologies are applied to build the statistical models. Thus, this paper brings together an array of tools to solve the broadcast capacity problem in VANET. The proposed methodology provides an efficient approach to estimate the performance of VANET in real-time, which will impact the efficacy of travel decision making.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Vehicular *ad hoc* networks (VANETs) are infrastructure light vehicular networks, in which instrumented vehicles communicate with one another across wireless links directly or through possible intermediate vehicles. Different paradigms for information communication have been proposed under the umbrella of vehicle infrastructure integration (VII, 2007) by the USDOT: this includes vehicle to vehicle (V2V) communication systems such as VANET; vehicle to infrastructure communication (V2I) and a hybrid mix of both V2V and V2I. The recent deployment projects have demonstrated the significant benefits of V2V and V2I systems on alleviating congestion, improving safety and in providing other non-traffic applications. Similarly, VANETs have a great potential to be applied to traffic control and traffic safety management so that the transportation system performance can be improved.

* Corresponding author. Tel.: +1 5182766033.

E-mail addresses: dul@rpi.edu (L. Du), ukkuss@rpi.edu (S. Ukkusuri), yushiw@rpi.edu (W.F. Yushimoto Del Valle), shivkuma@ecse.rpi.edu (S. Kalyanaraman).

There is a growing belief that the success of **VANET-based** inter-vehicle communications will facilitate a plethora of potential applications. In general, VANET applications will deliver functions and services in **five** key areas: (i) **traffic safety: provides** drivers with timely information to enable cooperative vehicle interactions with roadway environment to mitigate collision/injuries/deaths; (ii) **vehicular mobility and traffic operation: increases** road capacity by dynamic cooperative traffic operation and management; (iii) **consumer convenience: provides** information share for people sharing common interests; (iv) **vehicle probe data acquisitions: aggregates** real-time vehicular data for use in mobility and environmental applications; (v) **opportunistic data sharing: provides** opportunistic extension of the wireless mesh networks in improving the capacity and coverage, or information service for emergency situations. The first real **5.9 GHz dedicated short range communication technology** (DSRC) was approved by FCC in December 2003 for **vehicle infrastructure integration initiative** (VII) applications in USA. Similarly, vehicle to vehicle communication for safety applications was tested in the **advanced safety vehicle** (ASV) program sponsored by the Japanese Ministry of Land, Infrastructure and Transport (MLIT) (ITS, 2008). Other projects in Europe such as **NOW** (2008) and **FleetNet** (2008) demonstrate the significant traffic management benefits that can be obtained by a meaningful integration of wireless communication and vehicular networks. VANET technologies, once successfully developed, will have significant impacts to the society in general.

The potential applications of **VANET** have sparked a significant amount of research in both wireless communication and transportation areas. Although, **VANETs** have certain similarities to other mobile ad hoc networks, they are characterized by certain fundamental properties which makes their study more challenging (Blum et al., 2004; Moreno et al., 2005; Nekovee, 2007). For instance, **VANETs** are characterized by dynamic vehicular network in which mobile nodes are varying both spatially and temporally. In addition, driver behavior, mobility constraints due to space and existing infrastructure and high speed/acceleration of vehicles cause unique challenges to solve these problems. Furthermore, the number of vehicles in **VANET** cannot be scheduled or controlled in real-time because drivers will enter or exit the traffic network at different time instances. Several efforts have been invested on the performance of **VANET** in the recent years. For example, Artimy et al. (2004, 2005a,b) studied the information propagation in **VANET** by simulation of the transportation networks with an embedded communication network. Their simulation results show that the traffic factors, such as traffic density and traffic speed, affect the connectivity of **VANET** significantly. Wu et al. (2004) studied the spatial information propagation in **VANET**. Their analytical expression calculates the time delay of information transmission. Jin and Recker (2006) employed stochastic models to discuss the reliability of inter-vehicle communication under both uniform and general traffic streams. Assuming the presence of equipped vehicles follows an independent homogeneous Poisson process, Wang (2007) proposed a closed form expression for the expected propagation distance and its variance in the case of no transmission delay. Du et al. (2007) and Ukkusuri and Du (in press) explored the connectivity of **VANET** at every time instant by integrating traffic flow features.

Recently, there is a great deal of interest in understanding the fundamental transmission capacity of wireless networks in which transmission is permitted only between single **source-destination** pairs. This is because this allows us to characterize the throughput of the wireless network which determines the efficiency and amount of information flow in the ad hoc wireless network. In a VANET, a related problem of characterizing is very important since the throughput of the information is dependent on the vehicular mobility and the opportunities available for communication. As in a self-organized system, each individual vehicle broadcasts its own messages to the entire network nodes without cooperation. Therefore, one of the primary problems for **VANET** is data loss due to packet collision among concurrent transmissions (Jamieson and Balakrishnan, 2008). When the network loses data, retransmissions have to be conducted if the application requires a certain level of data fidelity, and doing so delay information propagation and degrade the information propagation efficiency. Therefore, the broadcast capacity acts as a key performance limit for information propagation efficiency in **VANET**. Our goal in this paper is to understand this issue comprehensively using optimization approaches to characterize the efficiency of information flow in **VANET**.

1.1. Related work

In the context of wireless ad hoc networks, there has been some related work which addresses the issue of network capacity. A seminal work is the research by Gupta and Kumar (2000) which defines per node throughput in a wireless network by a bit-distance product. This work supposes the network transports one bit-meter when one bit has been transported a distance of **1 m**. The throughput of the network is measured by the number of bit-meters that are transported per second. Their results show that as a wireless network is formed by n identical randomly located nodes, each node capable of transmitting at **W bits/s** and with a fixed broadcast radius, the local throughput per node is $\Theta\left(\frac{W}{\sqrt{n} \log n}\right)$ bits/s under a non-interference protocol. It is also shown that if the distribution of nodes, the traffic patterns and the transmission range of each node are optimally chosen, the bit-meter product that can be transported by the network per second is $\Theta(\sqrt{An})$ bit-meters/s. Extending the case of stationary nodes in Gupta and Kumar (2000), Grossglauser and Tse (2002) study a mobile ad hoc network, in which n mobile nodes communicate in random **source-destination** pairs. Allowing the unbounded delay and using only the one-hop relaying, their results show the mobility of nodes increases per node throughput asymptotically as the number of nodes becomes arbitrary large. Studying a wireless ad hoc network with finite immobile nodes, Jain et al. (2005) propose methods to compute the upper and lower bounds of the network capacity with no power control. Behzad and Rubin (2006) propose that independent of nodal distribution and traffic pattern, the throughput capacity of an ad hoc wireless network is maximized by properly increasing the nodal transmission power. Gastpar and Vetterli (2002) extend the work of Gupta and Kumar (2000) by assuming a network coding model where nodes could cooperate in arbitrary ways.

114 In addition, there is only a random picked source and destination, and the rest of the nodes act as relays. Their results show
 115 that under **these** conditions the throughput is $O(\log(n))$.

116 Based on different network features, there is other research, for instance, [Yi et al. \(2003\)](#) and [Huang et al. \(2007\)](#) propose
 117 the throughput of the wireless network increases with directional antennas. [Li et al. \(2001\)](#) present simulation-based results
 118 on the throughput of the small ad hoc network. [Negi and Rajeswaran \(2004\)](#) study the throughput of the wireless network, in
 119 which communication links apply ultra-wide band links.

120 Only a few papers deal with the transmission limits in wireless ad hoc network, in which transmissions are permitted
 121 from a source to a group of destinations. [Koskinen and Virtamo \(2005\)](#) evaluated the probability of successful transmission
 122 in a random time slot with the assumptions that all nodes employ the same transmission range and nodes are randomly
 123 located. [Son et al. \(2006\)](#) used Mica2 platform to study the relationship between concurrent transmissions and **signal-to-**
 124 **interference-plus-noise ratio** in static wireless networks. Their research emphasizes the effects that the concurrent transmis-
 125 sions have on packet error rate.

126 In relation to **VANET**, there are two key limitations which should be considered to model the capacity of the network.
 127 First, most of the previous research **focus** on general wireless ad hoc networks, and either assume the network is static or
 128 assume nodes move randomly in the study area. Obviously, those assumptions are not applicable to **VANET**. For example,
 129 in **VANET**, vehicles move on the fixed road network under a set of traffic control policies. Transportation science studies have
 130 shown that under the given traffic demand, road capacity, and speed limit, average traffic density and speed are not fully
 131 predictable. Consequently, the movements of vehicles in **VANET** can be simulated using data about the traffic conditions.
 132 Second, even though nodal capacity of wireless network has been extensively studied in terms of network size, transmission
 133 power, bandwidth constraints, and antenna directionality, very few studies discuss about the concurrent transmissions
 134 which influence the capacity of wireless network. Moreover, it is clear that the node density, node distribution and nodal
 135 mobility in road networks will characterize the message transmission limits in **VANET**. Therefore, it is valuable to investigate
 136 the interaction between the transmission limits in **VANET** and the traffic flow features. To the best of our knowledge, there is
 137 no previous work which addresses the above issues in **VANET**.

138 Motivated by the above points, assuming the given **VANET** employs a broadcast transmission protocol, we consider that
 139 the number of successful concurrent transmissions in **VANET** reflects the information broadcast capacity of the entire net-
 140 work, and also may cause the performance bottlenecks. In addition, we account for the fact that number of successful con-
 141 current transmissions is affected by various factors, such as traffic density, vehicular distribution, traffic pattern, network
 142 size, transmission power, bandwidth constraints, and antenna directionality. This research, therefore, defines the broadcast
 143 capacity of **VANET** as the *maximum successful concurrent transmissions*.¹ With this definition, we further investigate the effect
 144 of traffic flow network features and communication network configurations on the capacity of **VANET** on freeway segments. Our
 145 research objective is to explore the explicit relationship between the broadcast capacity of **VANET** and the significant factors in
 146 both transportation and wireless communication networks. Specifically, two **integer programming** (IP) models are developed
 147 which incorporate traffic flow features to optimize the capacity of **VANET**. Since the IP models cannot be solved efficiently in
 148 **real-time**, we develop an advanced statistical model to predict the network capacity by the significant parameters in both trans-
 149 portation networks and communication networks. In the description that follows, the performance of **VANET** refers to the capac-
 150 ity of **VANET** and nodes are also referred as vehicles.

151 The rest of this paper is organized as follows: Section 2 defines the network model used in this paper. Section 3 describes
 152 our IP models to investigate the capacity under our definition. Section 4 discuss our methodology to develop the statistical
 153 models, which are used to predict the capacity of **VANET** by significant parameters in both the road networks and the com-
 154 munication networks. Section 5 summarizes our main results, insights, and further identifies the future work.

155 2. Network model

156 We begin our study with some assumptions about the network model. We first briefly introduce our specifications about
 157 the network topology, the **medium access control**, and the antenna of the communication device. After those, we study in
 158 depth the application of signal-interference-plus-noise threshold in our study.

159 2.1. Network topology

160 The **VANETs** of interest in this study are comprised of a fleet of n vehicles moving on the **sample freeway segment** defined
 161 in Chapter 3 and with length L . Moreover, vehicles on the entire road segment are numbered from the entrance to exit.

162 2.2. Medium access control (MAC)

163 The time varying network topology and the lack of centralized control in **VANET** make the use of coordinated MAC pro-
 164 tocols difficult, but render a random access schema to be a preferred choice.

¹ A transmission includes a receiving node and a transmitting node.

165 Therefore, in this study, we assume the medium access control in VANET applies the spirit of a random access scheme.
 166 That is to say, vehicles independently broadcast messages to the nodes in the entire network without cooperation (Koskinen
 167 and Virtamo, 2005).

168 **2.3. Antennas**

169 We assume individual vehicles are equipped with directional antennas. In addition, if two or more messages are sent to
 170 the same vehicle simultaneously, they will result in message collisions and failed transmissions. But, one message can be
 171 received by multiple receiving nodes simultaneously.

172 **2.4. Signal-to-interference-and-noise ration (SINR)**

173 Given by Gupta and Kumar (2000), SINR is one of the conditions of a successful transmission: in any time slot, for a given
 174 receiver node k at y_k , and a transmitting node j at y_j , if node j can successfully transmit message to node k then the SINR at
 175 receiver is greater than some common threshold, TH , i.e.

$$177 \frac{p_j \|y_j - y_k\|^{-\alpha}}{N_0 + I} \geq \text{TH}, \quad (1)$$

178 where N_0 is the background noise on the frequency channel utilized by network, p_j is the transmission power of node j , and I
 179 denotes the sum of the interference power. Defining the transmission state of each node by e_i ,

$$181 e_i = \begin{cases} 1 & \text{node } i \text{ is transmitting packets,} \\ 0 & \text{o.w.,} \end{cases}$$

182 the sum of interference can be measured by

$$184 I = \sum_{1 \neq i, j} e_i p_i \|y_i - y_k\|^{-\alpha}. \quad (2)$$

185 In order to simplify the problem, we ignore the background noise and assume that all nodes transmit messages with the
 186 same power, $p_i \equiv p$, then the SINR condition for a successful transmission is reduced to

$$189 \sum_{i \neq j} e_i \|y_i - y_k\|^{-\alpha} \leq \frac{1}{\text{TH}} \|y_j - y_k\|^{-\alpha}. \quad (3)$$

190 Eq. (3) demonstrates that two key factors impact the success of a transmission: (1) the distances between the transmitters
 191 and receivers and (2) the interference around the receivers. In other words, the distribution of vehicles, the transmission
 192 range, and the SINR threshold of the wireless communication device will influence the success of the transmission in VANET
 193 significantly. This observation serves as a starting point for our study in following sections.

194 **3. Modeling the capacity of VANET by IP models**

195 The number of successful concurrent transmissions in VANET reflects how many packets can be successfully transmitted
 196 simultaneously in VANET at any given snapshot time t . In other words, it demonstrates how fast an information can be dis-
 197 seminated in the VANET. Based on this, we define the capacity of VANET in this study as the maximum number of successful
 198 concurrent transmissions in VANET. Considering each successful transmission includes one transmitter as well as one recei-
 199 ver. To capture this objective, clearly, involves two aspects. On the one hand, we can maximize the number of successful
 200 concurrent receiving nodes (receivers). On the other hand, we can maximize the number of successful concurrent transmit-
 201 ting nodes (transmitters). A successful receiving node is governed by four conditions: (1) the node should be at the receiving
 202 state, (2) since the signal power degrades in terms of the distance, only the nodes in the transition range are possible to ob-
 203 tain the information packages successfully, (3) there must be a corresponding transmitting node, which sends packets to the
 204 receiving node and (4) the signal power at the receiving node should be higher enough than the interference. In other words,
 205 SINR at the receiving node should be higher than its required threshold. Correspondingly, a successful transmitting node is
 206 identified by other two conditions: (1) the node is in transmitting state and (2) there is at least one other node, which is
 207 successfully receiving its packets. In the following study, we build the IP models for: (i) maximum concurrent receivers
 208 and (ii) the maximum concurrent transmitters. Before we discuss the IP models, we first add two more notations:

209 • R : a parameters that denotes the transmission range and
 210 • z_{jk}^t : a decision variable which is defined below:

$$212 z_{jk}^t = \begin{cases} 1 & \text{node } k \text{ successfully receives packets of node } j \text{ at time } t, \\ 0 & \text{o.w.} \end{cases}$$

213 The IP model which maximizes the number of concurrent receiving nodes are provided in IP model 1, where the vehicular
 214 location variables, y_i^t , $i = 1, \dots, n$ are considered as input parameters. After changing the constraint (10) and the objective

215 function in Eq. (4), we obtain IP model 2, which maximizes the number of concurrent transmitting nodes. The two IP models
 216 are provided below separately.

217 *IP model 1: maximizing the number of successful concurrent receiving nodes*

$$\max \sum_{j=1}^n \sum_{k=1}^n z_{jk}^t, \quad (4)$$

$$\sum_{i=1(i \neq j)}^{k-1} e_i^t (y_i^t - y_k^t)^{-\alpha} - TH^{-1} e_j^t (y_j^t - y_k^t)^{-\alpha} \leq M(1 - z_{jk}^t), \quad j \neq k \quad (5)$$

$$R^{-\alpha} - e_j^t * (y_j^t - y_k^t)^{-\alpha} \leq (1 - z_{jk}^t), \quad k = 1, \dots, n, \quad j = 1, \dots, k-1, \quad (6)$$

$$e_k^t + \sum_{j=1}^{k-1} z_{jk}^t \leq 1, \quad (7)$$

$$\sum_{k=1}^n e_k^t \geq 1, \quad (8)$$

$$\sum_{j=k}^n z_{jk}^t = 0, \quad k = 1, \dots, n, \quad (9)$$

$$\sum_{k=j+1}^n z_{jk}^t \leq ne_j^t, \quad j = 1, \dots, n, \quad (10)$$

$$e_k^t = 0, \text{ or } 1, \quad k = 1, \dots, n, \quad (11)$$

$$z_{jk}^t = 0, \text{ or } 1, \quad k = 1, \dots, n, \quad j = 1, \dots, n, \quad (12)$$

220

221 *where*

222 • Constraint (5) presents the condition that if the information packets can be successfully received at a receiver then the
 223 signal power at the receiver is higher enough than the interference around.
 224 • Constraint (6) presents the condition that if information packets can be successfully received at a receiver then the receiver
 225 should be in the transmission range of the corresponding transmitter.
 226 • Constraint (7) presents the condition that a node can either be a transmitter or a receiver, but not both at the same time.
 227 • Constraint (8) presents one more requirement for the network that at least one node is transmitting information at each
 228 time.
 229 • Constraint (9) presents that each node can only get information from one direction (forward or backward).
 230 • Constraint (10) presents the condition that if a node is successfully receiving information packets, then there must be a
 231 corresponding transmitter.
 232 • Constraint (11) and Constraint (12) show e and z are binary variables.

233

234

235

IP model 2: maximizing the number of successful concurrent transmitting nodes

$$\max \sum_{j=1}^n e_j^t, \quad (13)$$

Eq. (5) ... to ... Eq. (9),

Eq. (11) ... to ... Eq. (12),

$$\sum_{k=j+1}^n z_{jk}^t \geq e_j^t, \quad j = 1, \dots, n, \quad (14)$$

237

238 *where*

239 • Objective function (13) maximizes the number of successful concurrent transmitters in VANET.
 240 • Constraint (14) presents the condition that if a node is considered as a successful transmitting node then there is at least
 241 one other node which is successfully receiving its packets.

242

243

244

245

246

247

248

When VANET is used to build decentralized advanced traveler information systems (ATIS), the capacity of VANET directly influences the information quality and quantity reaching to individual vehicles, which consequently affect driver's routing decision en-route. Therefore, it is necessary to integrate the capacity of VANET into the routing decision policy. To realize it, the capacity of VANET must be evaluated in real-time. Even though the above IP models can provide the capacity of a VANET with the given network topology (i.e. vehicular positions) at each time instant, two limitations of the IP models limits their potential applications for real-time traffic operation managements. First of all, based on our computational experiences,

249 for most topology of VANET, it will take a relative long time to obtain the optimal value. This is not consistent with the rapid
 250 topology change in VANET, since for each topology, the capacity of the VANET need to be re-computed. Second, the vehicular
 251 trajectory data used in IP models is the microscopic traffic data, which cannot explicitly capture the relationship between
 252 traffic flow features and the capacity of VANET. Motivated by the above points, we next investigate the statistical models
 253 which predict the capacity of VANET by the significant parameters in communication networks as well as macroscopic traffic
 254 flow characteristics.

255 **4. Modeling the capacity of VANET by statistical models**

256 Our main idea is to estimate the capacity of VANET by using the statistical models. As presented in Fig. 1, we consider the
 257 IP models as a black box which has input data such as the transmission range, the SINR threshold, the network size, and the
 258 vehicular locations. The corresponding output data is indicated to VANET capacity. For a given road segment with fixed traf-
 259 fic flow characteristics, we solve the IP models at each time instant, and then build a relationship between the capacity of the
 260 VANET and the given parameters in the communication layer and the transportation layer. In the following paragraphs, we
 261 first discuss the experiment design, and then we demonstrate that the design helps in building an efficient statistical model.

262 *4.1. Experiment design*

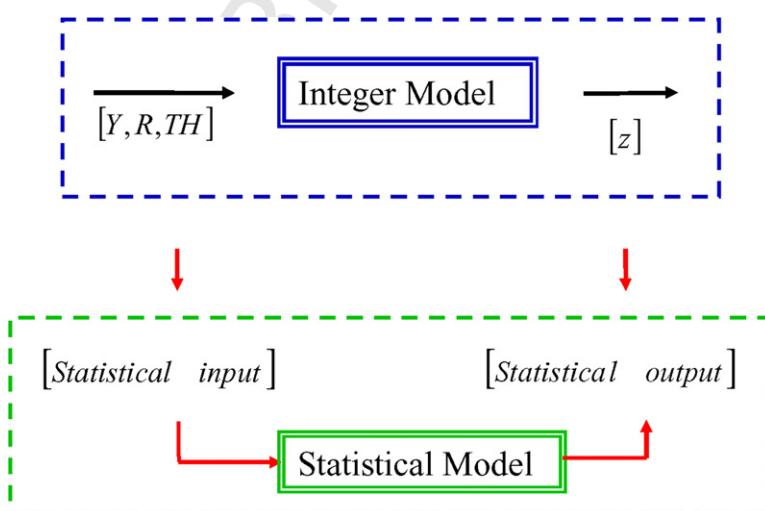
263 Our experiment procedure to investigate the statistical models is described in Fig. 2. In this experiment, we characterize
 264 the transportation network by parameters, such as traffic demand, the number of lanes, and free flow speed limitation, which
 265 are also the necessary parameters used to specify the uninterrupted road segments in MITSIMLab (MIT, 2002). At the same
 266 time, the communication network is featured by transmission range and SINR threshold. Based on the features in both trans-
 267 portation and communication network mentioned above, we develop our statistical models. In order to obtain the statistical
 268 model with high prediction capability, we have the following requirements for the experimental procedure:

269 1. Vehicular trajectory data is required as the input parameters for the IP **models** 1 and 2. A well tested, open source
 270 dynamic traffic simulation framework **MITSIMLab** (Ben-Akiva et al., 2002) is used as the simulation platform. It is devel-
 271 oped by MIT Intelligent Transportation System Program.
 272 2. Various combination among lanes, speed limit and traffic flow should be setup in the MITSIMLab simulation so that the
 273 vehicular trajectory data under different traffic flow conditions can be collected.
 274 3. Various reasonable transmission ranges and SINR thresholds should be involved in the experiments.

275 *4.1.1. Response surface methodology*

276 Since a multitude of **factor** considered in the statistical model, it is hard to have a limited data set covering all possible
 277 situations. Therefore, we use **response surface methodology** (RSM) to design a set of experiments.

278 RSM is a technique employed before the regression analysis so that the prediction capability of the relationship can be
 279 guaranteed even though limited experiments are conducted. That is, both the input variables and their values must be se-
 280 lected carefully so that the output obtained represents a good approximation of the population. Thus, the RSM involves the



281 **Fig. 1.** Exploring the relationship between the network characteristics and VANET performance.

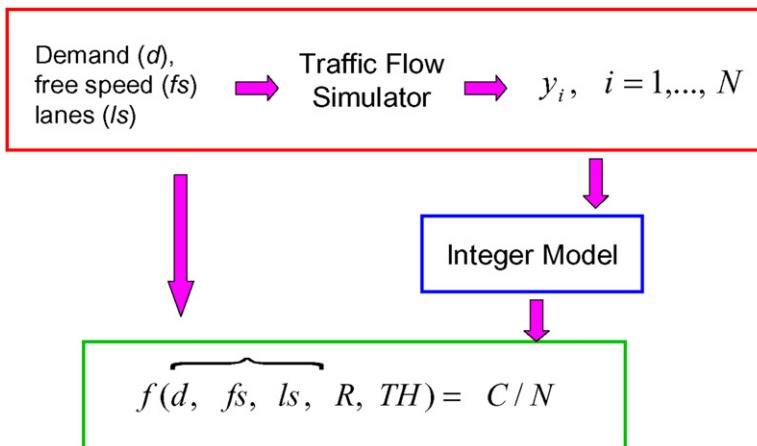


Fig. 2. Experiment procedure for developing the statistical models.

experiment design and the regression analysis to gain the understanding of the response variable using specific inputs. Regression analysis of RSM usually uses orthogonal arrays in the computation of the parameters (the most popular of these designs are the factorial 2^k) so that the coefficient estimates are uncorrelated and robust to small changes in the observed data. Moreover, RSM uses designs that must accomplish some specific properties (see Box and Draper (1987)). One of the principal properties is the rotatability or the distribution of information in the experimental region. A rotatable design has the same prediction variance in the response at any points that are equidistant from the design center. For more detail information about RSM experiment design, the reader is mainly referred to the books by Box and Draper (1987), Khuri and Cornell (1987), or Myers and Montogomery (2002).

In addition, before explaining the RSM approach to the VANET capacity problem, we will briefly introduce a class of RSM model called the **central composite design** which has been chosen because of its good rotatability and orthogonality properties for our experiment.

4.1.1.1. *Central composite design CCD*. The **central composite design** (CCD) is a design used in RSM and contains an imbedded factorial or fractional design with center points. To improve the experiment design space, it is augmented with a group of “star points” or “axial points” (\nearrow) that allow estimation of curvature. If the distance from the center of the design space to a factorial point is ≥ 1 unit for each factor, the distance from the center of the design space to the axial point is $\pm \alpha$ with $|\alpha| > 1$. The precise value of α depends on certain properties desired for the design and on the number of factors involved:

$$\alpha = (\text{number of runs})^{1/4}. \quad (15)$$

For example, for a **three** factor experiment at two levels $\alpha = (2^3)^{1/4} = 1.682$. The number of runs ($=2^k$) depends on the number of factors (k). Similarly, the number of center points to contain also depends on certain properties required for the design.

4.1.2. Description of RSM procedure in this study

The experimental design for the VANET capacity problem is a half CCD with five factors at two levels: demand (d), free speed (fs), the number of lanes (ls), SINR threshold (TH) and transmission range (R). Reference values used to generate the experiment data set is shown in Table 1. The design (combinations) will be used to predict the ratio of the successful receiving nodes (C) to the network size (N) (or the ratio (S/N) of the successful transmitting nodes (S) to the network size). Then the ratio (C/N) can be adjusted by a response surface using the five factors chosen, say $C/N = f(d, fs, ls, TH, R)$.

Based on the reference data, we obtain the experiments shown in Table 2. Each parameter has **five** possible values corresponding to the levels $+1, -1$, the center point 0 and the axial points $+2, -2$ ($\alpha = [2^{5-1}]^{1/4}$). For example, the corresponding traffic flows are 1000 (-2), 1750 (-1), 2500 (0), 3250 ($+1$), 4000 ($+2$) v/h. The CCD consists in two separate components: the factorial design which allows the estimation of the linear and interaction terms, and the additional points (axial and central points) to estimate coefficients for the **second-order** variables. Therefore, the number of runs for the CCD design is a function of the number of runs of the factorial experiment. For instance, in our experiment, the total number of runs for the half CCD is

Table 1

Reference values of CCD experiment design

	Demand	Lanes	Free speed	R	TH
Lower level (-1)	1750	2	72.5	112.5	17.5
Upper level ($+1$)	3200	4	77.5	237.5	22.5

Table 2

CCD experiment design

<i>d</i>	ls	fs	<i>R</i>	TH	C/N	S/N	<i>d</i>	ls	fs	<i>R</i>	TH	C/N	S/N
1000	3	75	175	20	0.596511	0.272059	2500	3	75	175	20	0.684063	0.270510
1750	2	72.5	112.5	22.5	0.545913	0.311828	2500	3	75	175	20	0.694276	0.279070
1750	2	72.5	237.5	17.5	0.780279	0.258065	2500	3	75	175	20	0.688819	0.292683
1750	2	77.5	112.5	17.5	0.565577	0.258242	2500	3	75	175	20	0.692019	0.250000
1750	2	77.5	237.5	22.5	0.754056	0.291777	2500	3	75	175	20	0.676968	0.301394
1750	4	72.5	112.5	17.5	0.580614	0.272487	2500	3	80	175	20	0.692202	0.271646
1750	4	72.5	237.5	22.5	0.750006	0.254598	2500	5	75	175	20	0.662521	0.239687
1750	4	77.5	112.5	22.5	0.548842	0.266963	3250	2	72.5	112.5	17.5	0.604998	0.274279
1750	4	77.5	237.5	17.5	0.783397	0.250000	3250	2	72.5	237.5	22.5	0.784760	0.250223
2500	1	75	175	20	0.671612	0.240015	3250	2	77.5	112.5	22.5	0.569831	0.250000
2500	3	70	175	20	0.704380	0.275172	3250	2	77.5	237.5	17.5	0.838229	0.260882
2500	3	75	50	20	0.458840	0.296341	3250	4	72.5	112.5	22.5	0.563773	0.270468
2500	3	75	300	20	0.883917	0.278883	3250	4	72.5	237.5	17.5	0.810913	0.268519
2500	3	75	175	15	0.738219	0.258583	3250	4	77.5	112.5	17.5	0.601975	0.272156
2500	3	75	175	25	0.677866	0.268293	3250	4	77.5	237.5	22.5	0.785274	0.269330
2500	3	75	175	20	0.676876	0.269231	4000	3	75	175	20	0.708620	0.260714

Table 3

Range of parameters considered in the experiment design

	Demand	Lanes	Free speed	<i>R</i>	TH
Minimum	1000	1	70	50	15
Maximum	4000	5	80	300	25

314 32 which is obtained by $2^{5-1} + 10 + 6 = 32$ where 2^{5-1} is the runs of the half factorial design, 10 corresponds to the axial
 315 points (2 level times 5 factors) and 6 corresponds to the number of replicates of the center point. The replicates in the center
 316 point are used to estimate the curvature and provide a stable measure of the variability (see Myers and Montgomery (2002)).
 317 As a rough guide a number of (3-5) center points are usually added in factorial experiments. However, for CCD designs Myers
 318 et al. (1992) have recommendations about the number of center points. In our experiment we decided to use 6 which uses
 319 one more run than the recommendation.

320 For each input data set: *d*, *fs*, *ls*, TH and *R*, we will run simulation models and IP models to obtain the vehicular trajectory
 321 data and the corresponding successful concurrent transmissions under different time instants. Finally we obtain the average
 322 output ratio: C/N or S/N for each input data set (see Table 2).

323 Based on the data set in Table 2, we will develop our statistical models. Due to the limitation of the experiment data set,
 324 the applicability of the developed statistical models are restricted to the range of values presented in Table 3. In other words,
 325 within the values in Table 3 we can assure that the adjusted *R-square* will represent a good measure of "goodness of fit" of
 326 the model. For values out of these ranges, the measure would be the *R-square* of prediction (included in the regressions also)
 327 but in some cases these are good and in others it is poor.

328 4.2. Statistical model for maximum successful concurrent receivers

329 Based on the CCD experiment data, we first conduct a simulation by MITSIMlab to get the vehicle trajectory data at each
 330 time instance under given network configurations. Combining the vehicle trajectory data, signal-to-interference-plus-noise
 331 threshold and transmission range in the CCD experiments, IP model 1 strictly explores the maximum number of successful
 332 concurrent receivers under each network configurations. Finally, we develop the statistical model to estimate the ratio of the
 333 maximum number of concurrent receivers *C* to the network size *N*. We have evaluated two different statistical models: a
 334 linear model and a second-order polynomial model.

335 We first use the best subsets algorithm to find the best linear model evaluated by the *C_p* Mallows and the highest adjusted
 336 *R-sq* criterions (Neter et al., 2004). The corresponding results in Table 4 show that the best linear model corresponds to the
 337 one presented in Table 5, which includes the independent variables: *d*, *R* and SINR. Its *R-sq* and *R-sq* (adj) are 97.7% and
 338 97.5%, respectively.

339 Employing the same procedure above and further including the second-order variables, we obtain the final nonlinear
 340 model presented in Table 6. It includes the same variables in the linear model but with the addition of *d*². This improves
 341 the adjusted *R-sq* up to 98.2%.

342 Examining the statistics pertinent to predict the performance of the linear and nonlinear models above, we find that the
 343 linear model in Table 5 is not significant since the *P-value* indicates a fitness test value² which is less than the significant size

² $H_0 : E\{y\} = \beta_0 + \beta_1 x_1 + \cdots + \beta_{p-1} x_{p-1}$; $H_a : E\{y\} \neq \beta_0 + \beta_1 x_1 + \cdots + \beta_{p-1} x_{p-1}$

Table 4

Best subsets for C/N

Vars	R-sq	R-sq (adj)	Mallows C-p	S	d	<i>l</i> <i>s</i>	<i>f</i> <i>s</i>	R	T H
1	92.4	92.2	59.1	0.027246					×
1	3.2	0.0	1086.2	0.097464	×				
2	95.6	95.3	24.3	0.021057	×				×
2	94.5	94.1	37.0	0.023576					×
3	97.7	97.5	2.2	0.015477	×				×
3	95.7	95.2	26.1	0.021382	×	×			×
4	97.7	97.4	4.0	0.015694	×	×			×
4	97.7	97.4	4.2	0.015761	×		×	×	×
5	97.7	97.3	6.0	0.015992	×	×	×	×	×

Table 5Regression analysis: *avgratio* versus *d*, *R*, *TH*

Predictor	Coefficients	SE coefficients	T	P	VIF
The regression equation is <i>avgratio</i> = 0.445 + 0.000026 <i>d</i> + 0.00170 <i>R</i> – 0.00640 <i>TH</i>					
Constant	0.44474	0.02890	15.39	0.000	
<i>d</i>	0.00002640	0.00000421	6.27	0.000	1.0
<i>R</i>	0.00170370	0.00005055	33.70	0.000	1.0
<i>TH</i>	-0.006404	0.001264	-5.07	0.000	1.0
<i>R</i> -sq = 97.7 and <i>R</i> -sq (adj) = 7.5% PRESS = 0.00938855 and <i>R</i> -sq (pred) = 96.81					
Source	DF	SS	MS	F	P
<i>Analysis of variance</i>					
Regression	3	0.287681	0.095894	400.33	0.000
Residual error	28	0.006707	0.000240		
Lack of fit	11	0.004612	0.000419	3.40	0.012
Total	31	0.294388			

Table 6Regression analysis: *avgratio* versus *d*, *R*, *TH*, *d*²

Predictor	Coefficients	SE coefficients	T	P	VIF
The regression equation is <i>avgratio</i> = 0.358 + 0.000101 <i>d</i> + 0.00170 <i>R</i> – 0.00640 <i>TH</i> – 0.00000001 <i>d</i> ²					
Constant	0.35817	0.03474	10.31	0.000	
<i>d</i>	0.00010067	0.00002151	4.68	0.000	36.6
<i>R</i>	0.00170370	0.00004269	39.91	0.000	1.0
<i>TH</i>	-0.006404	0.001067	-6.00	0.000	1.0
<i>d</i> ²	-0.00000001	0.00000000	-3.50	0.002	36.6
<i>R</i> -sq = 98.4% and <i>R</i> -sq (adj) = 98.2% PRESS = 0.00824379 and <i>R</i> -sq (pred) = 97.20					
Source	DF	SS	MS	F	P
<i>Analysis of variance</i>					
Regression	4	0.289775	0.072444	424.03	0.000
Residual error	27	0.004613	0.000171		
Lack of fit	10	0.002518	0.000252	2.04	0.094
Total	31	0.294388			

344 $\alpha = 0.05$. Therefore, we select the nonlinear model as our final statistical model to predict the ratio of the successful concurrent
 345 receivers to the network size of the VANET. Based on the regression model, we obtain the following insights, which are consistent
 346 with our intuition:

347 • The negative sign of SINR threshold, *TH* demonstrates that increasing SINR threshold degrades the proportion of successful concurrent
 348 receivers in VANET. This is consistent with our intuition that high SINR threshold results in less successful
 349 receivers in VANET.

350 • The positive sign of transmission range, *R* shows that large transmission range improves the proportion of successful receiving nodes in VANET. If the system only improves the transmission range, *R* of individual vehicles, then each transmitting node will cover more nodes. Due to the interference between concurrent transmissions, the optimal situation is that more nodes are in receiving state, consequently, the optimal ratio of receiving nodes in VANET will be improved. Therefore, the positive sign of *R* in the regression model is reasonable.

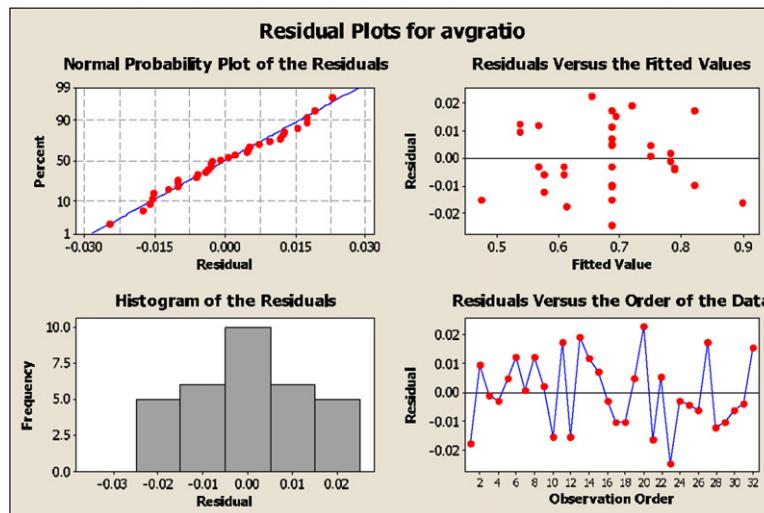


Fig. 3. Residual plot for average receiver ratio in VANET.

355 • The influence of traffic demand, d on the ratio of the successful receiving nodes in VANET to the network size has two
 356 sides. On the one hand, if only the traffic demand of VANET increases, and any other configurations in both the road net-
 357 work and the communication network do not change, then the traffic density will increase. This will make each transiting
 358 node cover more receiving nodes. Consequently, successful concurrent receiver ratio in the VANET will increase. This is
 359 consistent to the positive sign of demand variable in our regression model. On the other hand, if the traffic demand is
 360 increased a lot, then the vehicles will be distributed very densely. Consequently, the interference between concurrent
 361 transmission will degrade successful concurrent receiving rate. This phenomenon supports the negative sign of demand
 362 square in the selected regression model.

364 4.2.1. Regression model test

365 To examine the suitability of the final regression model in Table 6, we further test the model from several other aspects.
 366 Both graphic and formal statistical tests are applied.

367 • *Test for constancy of residual variance:* the plot of “residuals vs. fitted values” in Fig. 3 shows that the residuals randomly
 368 fluctuate around the reference line (residual = 0), which indicates that the residuals have constant variance.
 369 • *Test for independence of residuals:* the plot of “residuals vs. the order of the data” in Fig. 3 indicates the residual is indepen-
 370 dent on the order of the data since the residuals are randomly distributed around the reference line. We further apply Dur-
 371 bin–Watson statistic to test the correlation between adjacent residuals. The results in Table 6 show Durbin–Watson
 372 statistic is equal to 2.35494, which is greater than the upper bound, 1.67.³ Hence, there is no correlation between adjacent
 373 residuals.
 374 • *Test for normality:* the normal probability plot in Fig. 3 shows an approximately linear pattern consistent with a normal
 375 distribution. Fig. 4 shows the result of the standard normality test (Kolmogorov–Smirnov test⁴) which also indicates a nor-
 376 mal distribution of the residuals.
 377 • *Lack of fit test:* the P -value (=0.0984) of lack of fit test shown in Table 6 is larger than usual test size, 0.05, so we accept H_0
 378 that this model is appropriate.

380 4.2.2. Calibrating the prediction capability

381 In this study, we finally select the nonlinear model to predict the maximum ratio of successful concurrent receivers in
 382 VANET. The estimated regression coefficients, variance analysis and other statistics pertaining to the fitted model are shown
 383 in Table 6. Now, we further demonstrate the validity of the nonlinear model.

384 We first note that in the nonlinear model, the PRESS (prediction sum of squares) is equal to 0.00824379 is reasonably
 385 close to the SSE (=0.004613). This supports the validity of the fitted model. It also demonstrates that the MSE is a valid indi-
 386 cation of the prediction capability of the regression model.

³ Upper bound = 1.67 and lower bound = 1.46. Refer to <http://www.csus.edu/indiv/j/jensena/mgmt105/durbin.htm>.

⁴ Refer to http://en.wikipedia.org/wik/Kolmogorov-Smirnov_test.

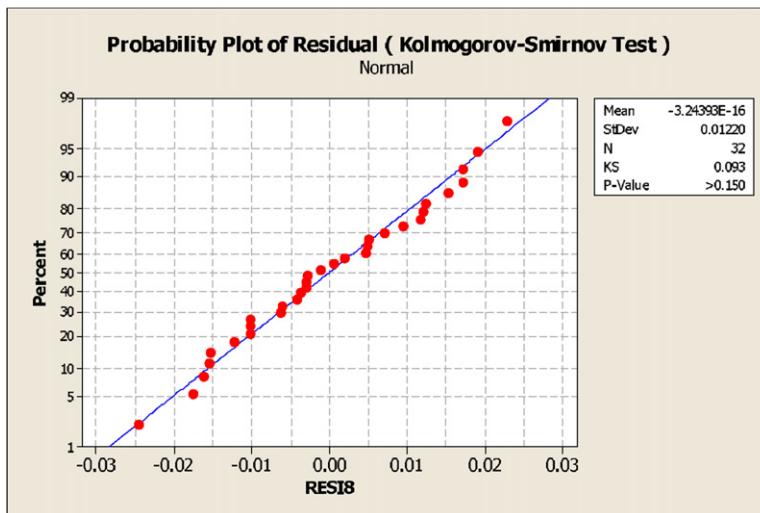


Fig. 4. Probability plot of residual for the regression model used to predict the maximum receiver ratio in VANET.

387 In addition, we collected another data set with 15 combinations in **Table 7** to test the predictive capability of the nonlin-
 388 ear regression model. For each data set, the capacity of the VANET measured by the regression model and the IP model are
 389 denoted by \hat{c} and c , respectively, in **Table 7**. Comparing the values of \hat{c} and c , we find that the average absolute error is about
 390 8% and the average relative error is 15%.

391 Therefore, we arrive at the conclusion that the prediction performance of the selected nonlinear regression model is rea-
 392 sonably good.

393 The validation results strongly support our idea that the significant parameters in both the transportation networks and
 394 the communication networks, such as the average traffic demand, the transmission range, and the signal-to-interference-
 395 plus-noise ratio are good indicators to predict the performance of VANET, for instance, the maximum ratio of the successful
 396 concurrent receivers in VANET.

397 4.3. Statistical model for maximum successful concurrent transmitters

398 Using the same idea and experiments shown in **Table 2**, we further explore the corresponding statistical model used for
 399 predicting the ratio (S/N) of the maximum successful concurrent transmitters to the network size in VANET.

400 IP model 2 is used to rigorously investigate the maximum number of successful concurrent transmitters in VANET and
 401 further provides the ratio (S/N). The results for the cases in the experiments are demonstrated in **Table 2**. Our regression

Table 7

New data for validation

Sample	Demand	Lanes	Free speed	R	TH	\hat{c}_i	c_i	\hat{s}_i	s_i
1	1550	2	72	195	24	0.668425	0.666872	0.298544	0.269231
2	3450	2	72	165	19	0.746325	0.716518	0.265129	0.259797
3	1500	2	77	160	16	0.656600	0.772115	0.263517	0.256250
4	1400	2	80	260	19	0.800200	0.881280	0.311925	0.275000
5	2075	2	76	60	18	0.511319	0.556250	0.290483	0.268750
6	3000	3	72	200	23	0.763800	0.594118	0.269055	0.286925
7	1500	4	73	225	19	0.747900	0.618889	0.257963	0.268750
8	3500	3	78	145	16	0.733100	0.615409	0.282924	0.269097
9	3100	2	72	200	18	0.799800	0.725000	0.266810	0.268750
10	2475	4	76	190	17	0.760919	0.818750	0.266506	0.281250
11	3600	5	70	140	16	0.727600	0.698903	0.293254	0.290025
12	1375	3	73	130	25	0.538969	0.428475	0.312075	0.279070
13	3150	1	76	60	25	0.518925	0.337500	0.240653	0.21875
14	1850	3	74	180	22	0.675825	0.587762	0.283980	0.227325
15	2500	4	79	70	16	0.564600	0.518750	0.263423	0.276741

\hat{c}_i is the predicted ratio of receivers in the VANET with the configuration in the i th new data set using the regression model.

c_i is the ratio of receiver in the VANET with the configuration in the i th new data set. c_i is calculated by IP model 1.

\hat{s}_i is the predicted ratio of senders in the VANET with the configuration in the i th new data set using the regression model.

s_i is the ratio of senders in the VANET with the configuration in the i th new data set. s_i is calculated by IP model 2.

Table 8

Response surface regression: *avgratio* versus coded variables, such as *d*-code and *ls*-code

Term	Coefficients	SE coefficients	T	P
<i>Estimated regression coefficients for avgratio</i>				
Constant	0.268330	0.002308	116.251	0.000
<i>d</i> _code	-0.005899	0.003398	-1.736	0.098
<i>ls</i> _code	-0.002619	0.003398	-0.771	0.450
<i>fs</i> _code	-0.004014	0.003398	-1.181	0.251
<i>R</i> _code	-0.005244	0.003398	-1.543	0.138
<i>TH</i> _code	-0.000118	0.003398	-0.035	0.973
<i>ls</i> _code * <i>ls</i> _code	-0.029835	0.006161	-4.842	0.000
<i>R</i> _code * <i>R</i> _code	0.029181	0.006161	4.736	0.000
<i>d</i> _code * <i>ls</i> _code	0.030238	0.008322	3.633	0.002
<i>d</i> _code * <i>TH</i> _code	-0.030547	0.008322	-3.670	0.002
<i>fs</i> _code * <i>R</i> _code	0.030572	0.008322	3.673	0.002
<i>R</i> -sq = 83.0%, <i>R</i> -sq = 74.5% and PRESS = 0.00361149				
Source	DF	Seq SS	Adj SS	Adj MS
<i>Analysis of variance for avgratio</i>				
Regression	10	0.006777	0.006777	0.000678
Residual error	20	0.001385	0.001385	0.000069
Lack of fit	16	0.001173	0.001173	0.000073
Total	30	0.008162		1.38
Term	Coefficients	Term	Coefficients	
<i>Estimated regression coefficients for avgratio using data in uncoded units</i>				
Constant	0.85413	<i>d</i>	0.000047	
<i>ls</i>	0.018245	<i>fs</i>	-0.009363	
<i>R</i>	-0.004364	<i>TH</i>	0.010159	
<i>ls</i> * <i>ls</i>	-0.007459	<i>R</i> * <i>R</i>	0.000002	
<i>d</i> * <i>ls</i>	0.00001	<i>d</i> * <i>TH</i>	-0.000004	
<i>fs</i> * <i>R</i>	0.000049			

The analysis was done using coded units.

402 analysis concludes with the final regression model with *R*-sq = 83.0% and *R*-sq (adj) = 74.5%. The coefficients, variance analysis,
 403 and related statistics are shown in Table 8. In the following section, we briefly discuss about the signs of the coefficients
 404 and the residual test results.

405 4.3.1. Regression model test

406 The coefficients of the first order variables in the regression model are intuitive. For example, the sign of variable *ls* is
 407 negative. The reason is that when there are more lanes on the road, the vehicles are distributed more sparse. If all other con-
 408 figurations, such as transmission range and traffic demand, do not change, then the proportion of the successful concurrent
 409 transmitters will degrade since some transmitters cannot have receivers in its transmission range. In addition, the regression
 410 model shows that the combinations, *d* * *ls*, *d* * *TH*, and *fs* * *R* significantly impact the successful concurrent transmitters in
 411 VANET since they have the relatively large coefficients. It is hard to present the intuitive meaning of the *second-order* vari-
 412 ables in the regression models.

413 Besides checking the sign of the variables in the regression model, we further evaluate the regression model by testing
 414 several statistical features of the residuals. The plots in Fig. 5 show that the residuals have good performance in constancy
 415 variance, independence, and normality. For instance, the normal probability plot in Fig. 5 demonstrates an approximately
 416 linear pattern, which is consistent with a normal distribution. In addition, Fig. 6 shows the results of the standard normality
 417 test (Anderson–Darling test), which also indicates a normal distribution of the residuals (*P*-value > 0.05).⁵

418 4.3.2. Calibrating the regression model

419 The selected regression model also demonstrates good validation features, such as PRESS = 0.00361149, and
 420 SSE = 0.001385. They are reasonably similar to one other. This indicates that the MSE is an valid indictor of the prediction
 421 capability of the regression model. We also validate the regression model by the new data sets shown in Table 7. Each data
 422 set represents a configuration in both the road network and the communication network for the VANET. The capacity of the
 423 VANET with the given configurations is estimated by the selected regression models as well as the rigorous IP model 2,
 424 respectively. The results are reported in Table 7 by column \hat{s} (the results of the regression model) and s (the results of IP
 425 model 2). We compare the values of \hat{s} and s , and find that on the average, the absolute error between \hat{s} and s is about 2%,
 426 and the average relative error between \hat{s} and s is 7%. These validation results show that our selected regression model per-
 427 forms reasonable good.

⁵ H_0 : the data follows a normal distribution and H_a : the data does not follow a normal distribution.

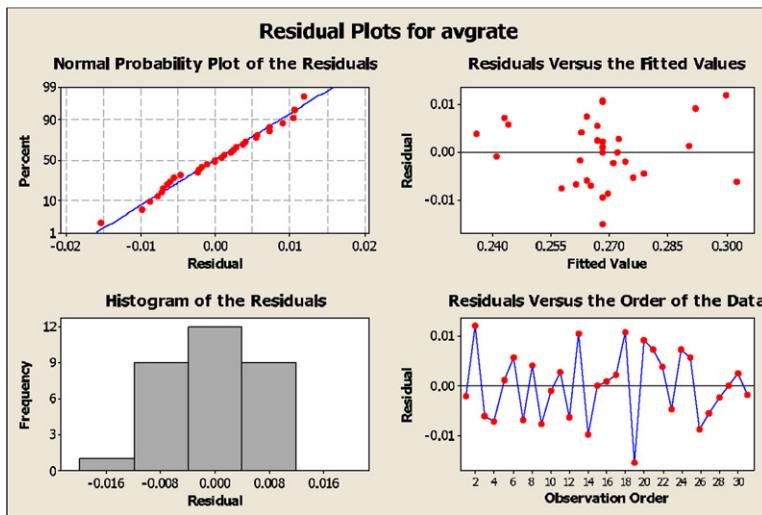


Fig. 5. Probability plot of residual for the regression model used to predict the maximum transmitter ratio in VANET.

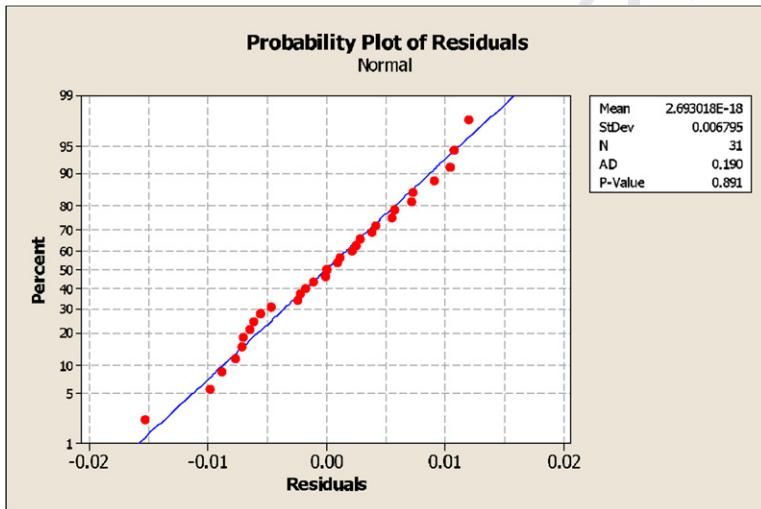


Fig. 6. Residual plot for average transmitter ratio in VANET.

428 The above two final statistical models integrate the parameters in both the transportation networks and the communication
 429 networks to estimate the broadcast capacity of VANET. For a given VANET with known network size, applying the re-
 430 ported statistical models we can approximate the maximum number of concurrent transmissions and therefore, know how
 431 many messages can be successfully transmitted in the network. This provides us more insight about the information prop-
 432 agation opportunities in VANET. They bring two significant benefits for VANET application in transportation. On the one
 433 hand, the explicit relationship between the broadcast capacity and the network configurations of VANET provides a guidance
 434 to design more efficient VANET for traffic managements by considering the traffic flow features. On the other hand, com-
 435 pared with the rigorous IP models, the statistical models provide a rapid way to approximate the broadcast capacity. There-
 436 fore, they have great potential to be used in the real-time application in the near future.

437 5. Conclusion

438 Due to the vehicular mobility, VANET is a special case of MANET. The simulation results in literature (Artimy et al., 2004,
 439 2005a,b, 2006; Blum et al., 2004; Moreno et al., 2005) have shown that the performance of VANET is strongly impacted by
 440 the traffic flow features such as traffic density and traffic speed. Therefore, to fully understand the properties of information
 441 propagation in the systems, it is important to study the relationship between the measurements which are used to evaluate
 442 the performance of VANET and the parameters which is used to characterize the features of the traffic flow networks.

443 In this study, we are interested in exploring the concurrent transmissions in VANET since it reflects how fast a new message
 444 such as a traffic incident can be disseminated among vehicles in VANET. Specifically, we wish to estimate the maximum
 445 successful concurrent transmissions in VANET, which is defined as the capacity of VANET in this study. Technically, the max-
 446 imum successful concurrent transmissions are represented by the maximum successful concurrent receivers or the maxi-
 447 mum successful concurrent transmitters. IP models are first built to rigorously explore the maximum concurrent
 448 transmissions in both ways. However, the IP models cannot be solved efficiently as the network size increases. We, there-
 449 fore, further develop the statistical models to approximate the concurrent transmissions. MITSIMLab is used as the simula-
 450 tion tool to generate the traffic flow data. CCD experiment design and linear regression are applied to develop the statistical
 451 models.

452 Our final regression models show that the ratio of maximum successful concurrent receivers in VANET is the function of
 453 average traffic demand and transmission range; the ratio of maximum successful concurrent transmitters in VANET is the
 454 function of average traffic demand, traffic speed limit, the number of lanes, SINR threshold, and transmission range. Further-
 455 more, we obtain several insights about interaction between traffic flow characteristics and the capacity of VANET. For exam-
 456 ple, large vehicular transmission range will improve the ratio of successful receivers and degrade the ratio of successful
 457 transmitters in VANET. Increasing traffic demand will improve the ratio of successful receivers and degrade the ratio of suc-
 458 cessful transmitter in VANET.

459 The main contributions of this work are in three-folders: first, this study investigates the interweave between the suc-
 460 cessful concurrent transmissions in VANET and traffic flow features such as traffic demand, the number of lanes and free
 461 speed limit. To our knowledge, this is the first time to study the capacity of VANET from the transportation points of view.
 462 Second, the methodology developed in this study can be further applied to investigate other relationships, in which the traf-
 463 fic flow features are represented by other parameters set or the configuration of the communication network is modified. For
 464 example, once we change the antenna features in VANET to be two-directional or omnidirectional, one immediate extension
 465 of this study is to address the new relationship between the traffic flow features and the capacity of VANET by the provided
 466 methodology. Finally, the provided statistical models in this study can be involved to building optimal online routing poli-
 467 cies, in which the capacity of VANET is considered as one of significant factors to evaluate the information quality. In con-
 468 clusion, the entire study in this paper strongly supports the implementation of the efficient decentralized ATIS in the near
 469 future.

470 Acknowledgements

471 The authors would like to acknowledge the support of the National Science Foundation (NSF) under Grant No. CNS-
 472 0627039 and the Blitman Career Development Chair Professorship.

473 References

474 Artimy, M., Robertson, W., Phillips, W., 2004. Connectivity in inter-vehicle ad hoc network. In: CCECE 2004 – CCIEI 2004, Niagara Falls, NY, USA, May 2–5.

475 Artimy, M., Phillips, W., Robertson, W., 2005a. Connectivity with static transmission range in vehicular ad hoc networks. In: Proceedings of the 3rd Annual
 476 Communication Networks and Services Research Conference, Washington, DC, USA, May 16–18.

477 Artimy, M., Robertson, W., Phillips, W., 2005b. Assignment of dynamic transmission range based on estimation of vehicle density. In: VANET'05, Cologne,
 478 Germany, September 2.

479 Artimy, M., Robertson, W., Phillips, W., 2006. Minimum transmission range in vehicular ad hoc networks over uninterrupted highways. In: The 9th
 480 International IEEE Conference on Intelligent Transportation Systems, Toronto, Canada, September 17–20.

481 Behzad, A., Rubin, I., 2006. High transmission power increases the capacity of ad hoc wireless network. *IEEE Transactions on Wireless Communications* 5 (1),
 482 156–165.

483 Ben-Akiva, M., Davol, A., Toledo, T., Koutsopoulos, H.N., Burghout, W., Andréasson, I., Johansson, T., Lundin, C., 2002. Calibration and evaluation of mitsimlab
 484 in stockholm. In: Proceedings of 81st Transportation Research Board Meeting, Washington, DC, USA, January 13–17.

485 Blum, J., Eskandarian, A., Hoffman, L., 2004. Challenges of intervehicle ad hoc networks. *IEEE Transactions on Intelligent Transportation Systems* 5 (4), 347–
 486 351.

487 Box, G., Draper, N., 1987. *Empirical Model Building and Response Surfaces*. John Wiley and Sons, New York.

488 Du, L., Ukkusuri, S., Kalyanaraman, S., 2007. Geometric connectivity of vehicular ad hoc networks: analytical characterization. In: The 4th ACM Workshop on
 489 Vehicular Ad Hoc Networks (VANET), Montreal, Canada, September 10.

490 Fleetnet, Inter-vehicle Communication. Available from: <<http://www.et2.tu-harburg.de/fleetnet/>> (accessed 05.05.08).

491 Gastpar, M., Vetterli, M., 2002. On the capacity of wireless networks: the relay case. In: IEEE INFOCOM, New York City, NY, USA, June 23–27.

492 Grossglauser, M., Tse, D., 2002. Mobility increase the capacity of ad hoc wireless network. *IEEE/ACM Transactions on Networking* 10 (4), 477–486.

493 Gupta, P., Kumar, P., 2000. Capacity of wireless networks. *IEEE Transactions on Information Theory* IT 46, 338–404.

494 Huang, X., Wang, J., Fang, Y., 2007. Maximum flow problem in wireless ad hoc networks with directional antennas. *Optimization Letters* 1 (1), 71–84.

495 Its-international-new Artical. Available from: <<http://www.itsinternational.com/news/article.cfm?recordID=12649>> (accessed 05.05.08).

496 Jain, K., Padhye, J., Padmanabhan, V.N., Qiu, L., 2005. Impact of interference on multi-hop wireless network performance. *Wireless Networks* 11 (4), 471–
 497 487.

498 Jamieson, K., Balakrishnan, H., 2008. Exploiting Concurrent Transmissions in Wireless Sensor Networks. Technical Report, MIT Computer Science and
 499 Artificial Intelligence Laboratory (CSAIL), 2006. Available from: <<http://publications.csail.mit.edu/abstracts/abstracts06/asn.shtml>> (accessed:
 500 February).

501 Jin, W., Recker, W., 2006. Instantaneous information propagation in a traffic stream through inter-vehicle communication. *Transportation Research Part B*
 502 40 (3), 230–250.

503 Khuri, A., Cornell, J., 1987. *Response Surfaces: Design and Analyses*. Marcel Dekker Inc., New York.

504 Koskinen, H., Virtamo, J., 2005. Probability of successful transmission in a random slotted-aloha wireless multihop network employing constraint
 505 transmission power. In: MSWiM'05, Montreal, Quebec, Canada, October 10–13.

506 Li, J., Blake, C., DeCouto, D.S.J., Imm Lee, H., Morris, R., 2001. Capacity of ad hoc wireless network. In: ACM MobiCom 2001, Long Beach, CA, USA, October 4–5.

507 User's Guide for MITSIMlab and Road Network Editor (RNE). MIT Intelligent Transportation Systems Program, November 2002. Available from: <<http://web.mit.edu/its/papers/Manual.pdf>> (accessed 04.07).

508 Moreno, M., Killat, M., Hartenstein, H., 2005. The Challenges of Robust Inter-vehicle Communications, Dallas, TX, USA, September 25–28.

509 Myers, R.H., Montgomery, D.C., 2002. Response Surface Methodology: Process and Product Optimization Using Designed Experiments. Wiley-Interscience, NY, USA.

510 Myers, R., Montgomery, D., 2002. Response Surface Methodology. John Wiley and Sons Inc., New York.

511 Myers, R.H., Vining, G.G., Giovannitti-Jensen, A.N.N., Myers, S.L., 1992. Variance dispersion properties of second-order response surface designs. *Journal of Quality Technology* 24 (1).

512 Negi, R., Rajeswaran, A., 2004. Capacity of power constrained ad-hoc networks. In: IEEE INFOCOM 2004, Hong Kong, China, March 7–11.

513 Nekovee, M., 2007. Sensor network on the road: the promises and challenge of vehicular ad hoc networks and grids, April 2007. Available from: <<http://www.semanticgrid.org/ubinesc/ubi-v1.pdf>>.

514 Neter, J., Kutner, M.H., Nachtsheim, C.J., 2004. Applied Linear Regression Models. McGraw-Hill/Irwin, Boston, NY, USA.

515 Now: Network on Wheels. Available from: <<http://www.network-on-wheels.de>> (accessed 05.05.08).

516 Son, D., Krishnamachari, B., Heidemann, J., 2006. Experimental study of concurrent transmission in wireless sensor networks. In: SenSys'06, Boulder, CO, USA, November 1–3.

517 Ukkusuri, S., Du, L., *in press*. Robust connectivity of vehicular ad hoc network under different traffic scenarios on freeway segments. *Transportation Research Part C* ^{Q1} _A.

518 Vehicle Infrastructure Integration (VII), VII Architecture and Functional Requirements. Available from: <<http://www.vehicle-infrastructure.org/documents/VII>> (accessed 05.05.07).

519 Wang, X., 2007. Modeling the process of information relay through inter-vehicle communication. *Transportation Research Part B* 41 (6), 684–700.

520 Wu, H., Fujimoto, R., Riley, G., 2004. Analytical models for information propagation in vehicle-to-vehicle networks. In: IEEE Vehicular Technology Conference, Los Angeles, CA, USA, September 26–29.

521 Yi, S., Pei, Y., Kalyanaraman, S., 2003. On the capacity improvement of ad hoc wireless networks using directional antennas. In: Proceeding of the 4th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc 2003), Annapolis, MD, USA, June 1–3.

522

523

524

525

526

527

528

529

530

531