Ad Hoc Networks 52 (2016) 17–27

Contents lists available at ScienceDirect

Ad Hoc Networks

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

Analytically modeling data dissemination in vehicular ad hoc networks

Xiaoyun Liu[a], Gongjun Yan[b,*]

[a] Anhui University, Hefei, Anhui 230601, China
[b] University of Southern Indiana, Evansville, IN 47712, USA

ARTICLE INFO

Article history:
Received 28 January 2016
Revised 7 September 2016
Accepted 9 September 2016
Available online 20 September 2016

Keywords:
VANETs
Data dissemination
Analytically models
Probability
Ad hoc networks

ABSTRACT

Without pre-defined infrastructures, vehicles connect to each other to form Vehicular Ad-Hoc Networks (VANETs) to deliver data among vehicles. Dissemination of messages, for example accident alert messages or congestion messages is critical in VANETs. People's lives may be at stake in accidents. Modeling and predicting VANETs' message dissemination opens an opportunity to adopt appropriate strategies to alert severe accidents, manage traffic, evacuate vehicles from disasters, or recover from accidents. However, modeling data dissemination is challenging due to vehicles' high mobility, which results in network topologies rapidly changing and data transmission being unstable. This paper presents analytical data dissemination models on VANETs. We model data dissemination as a new production adoption process and as a time-dependent stochastic process. The fact that information value or importance is decreasing with time and distance is also considered. The analytical models allow prediction and evaluation of information diffusion and enable intelligent traffic management.

1. Introduction

The past decade has witnessed a growing interest in vehicular ad hoc networking, a subset of mobile ad hoc computing. The initial vision of vehicular networking originally sees that radio-equipped vehicles can somehow network together and, by exchanging and aggregating individual views, can build the drivers informed about potential traffic safety risks, and can heighten their awareness of road conditions and other traffic-related events. In support of vehicular networking and, more generally, of traffic-related communications, the US Federal Communications Commission (FCC) has allocated 75 MHz of spectrum in the 5.850–5.925 GHz band for the exclusive use of Dedicated Short Range Communications (DSRC) [1]. As pointed out by a number of workers [2,3], the DSRC spectrum by far exceeds the needs of traffic safety applications. Therefore, a lot of applications have proposed [4,5]. There are mainly two categories [6,7]: safety related and Internet connectivity related applications.

Most, if not all, applications expect to know how many and how fast a message will be transmitted over VANETs. Therefore, modeling data dissemination is naturally important for applications such as, avoiding severe traffic accidents on highway, managing traffic, evacuating or recovering from disasters, etc.

But it is challenging to model data dissemination in vehicular networks. Vehicles are high mobile objects, which cause the topology of the network to constantly change. Inherently, data transmission among vehicles is unreliable and hard to predict. Yan, et al. [3,8] has modeled the communication linkage of one hop at micro level. Massive communication modeling has been an open field. Most models in VANETs are to model vehicular mobility [9,10].

Therefore, this paper analytically proposes analytical data dissemination models by extending the previous research [11,12]. In this paper, we mainly address two types of models: 1) We first analytically investigate the impact parameters on time-critical information diffusion by extending to the classic BASS diffusion model [11,13] which addresses the diffusion of new product among potential customers. 2) We probabilistically derive a data dissemination model where data diffusion is counted as a time-dependent stochastic process.

In this paper, we mainly investigate the spreading speed which is defined as the number of vehicles accepting the alert message per time unit. As an example, we model data dissemination in a road collision event scenario shown in Fig. 1. The spreading speed is a critical metric for traffic accident alert messages. We also show the probability of k vehicles receiving the alert message in the duration of time $[0, t]$.

The scope of this paper is mainly about dissemination simulation models in VANETs at the macro level. Our main assumptions include: 1) Vehicles can communicate to each other over the VANETs without considering connection challenges which have

* Corresponding author. Fax: +18122509264.
E-mail addresses: liujody@hotmail.com (X. Liu), gyan@usi.edu, yangongjun@gmail.com (G. Yan).

http://dx.doi.org/10.1016/j.adhoc.2016.09.010
1570-8705/

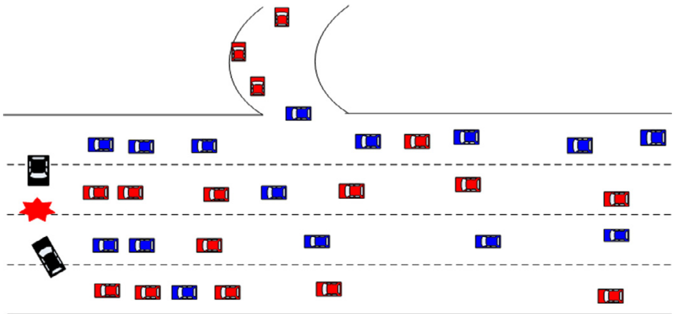


Fig. 1. The alert messages are disseminated among vehicles.

been addressed by many communication protocols [3,14–16]; 2) VANETs channel capacity [17,18] is satisfied by communication protocol surveyed in [14,15].

Our main contributions include:

1. Analyzing information diffusion affected by the value of time and distance. The distance influence can be obtained by analyzing the diffusion model based on time.
2. Deriving close forms for the time-dependent probability mass function of the number of adopters and of the expected time for information to reach a given threshold.
3. Deriving close forms of time-dependent stochastic process, which can better describe the dynamics of VANETs.

The remainder of the paper is organized as follows. Section 2 contains a succinct review of the state of the art. Section 3 discusses the details of the proposed data diffusion model by importing a new production adoption model. Therein, Section 3.1 offers a high-level discussion of the system model. In addition, Section 3.2 presents the details of the model at a macro level of data dissemination. Section 3.3 and Section 3.4 extend the new production model and discuss that value of messages can be devalued with time. Moreover, Section 4 offers the data dissemination model in a stochastic or dynamical way. Section 4.1 presents a fundamental stochastic model in which vehicles are dynamically receiving messages. Further, Section 4.2 considers a more sophisticated and realistic model in which vehicles can randomly leave the area-of-interest. Section 5 provides simulation and validation results. Finally, Section 6 offers concluding remarks and maps out directions for future work.

2. Related work

There is some research undertaken in message routing algorithm in VANETs. For example, a probabilistic routing protocol is based on the probability of wireless links [3,11]. The QoS based routing protocol has been proposed by balancing delay, throughput and cost [19]. Yan, et al. introduced more detailed routing protocols in VANETs [20].

However, little work has been done in the aspect of message diffusion model although some diffusion models have been studied. There are generally three types of diffusion models: stochastic epidemic models, virus diffusion models, and product diffusion models.

Yan [12] has proposed a data dissemination model by using a stochastic epidemic model. In the model, vehicles are mapped as individuals in the epidemic model. The transmission of message is mapped as virus contact among vehicles. In this model, parameters like arrival rate are time-dependent variables. A simple scenario where vehicles received an alert message and stayed in the system has been explored. To our best knowledge it started the attempt to model VANETs by using epidemic models. But it can be further extended by considering the vehicles “leaving” the system after receiving the alert message.

Arif and Olariu [21,22] have proposed an epidemic model which treats virus as a stochastic process. Individuals in contact with a virus will be infected and stay in infected status. The parameters of the model are time-dependent variables which can show the probability of new infected individuals is related to time t. Papadopouli and Schulzrinne [23] and Lindemann and Waldhorst [24] explore a simple stochastic epidemic model to analytically address the delay of the spreading of data. A pure birth process is considered as a stochastic epidemic model. The data diffusion process is represented as a continuous-time Markov process. Each node has two states: Infection and Susceptible. A small group of populations (5 nodes) are considered in the continuous-time Markov model. This is not directly applicable to VANETs because the population in VANETs is much bigger than just five nodes. Many diseases like cancer are either incurable or defy our persistent attempts at developing an effective vaccination and immunization regimen [25–27]. Epidemic models have been used to model the outbreak in AIDS epidemic in Vancouver [28], the effects of vaccination on the spread of seasonal outbreaks of influenza [29].

Khelil et al. [30] proposed an information diffusion model by applying the virus diffusion model. Information is disseminated by a contact. The contact is defined as a process that a node (susceptible) meets (or communicates) with another infected node (infectious node). The information flows from infected node(s) to susceptible node(s). This infection process continues until all the nodes are infected. However, for time-critical information like accident alert messages, the value of information decreases with time. Time issues are not considered in Khelil’s model. In our model, the fact that value of information decreases with time is also addressed in the diffusion model.

A static trapping model is addressed by Papadopouli and Schulzrinne [31] as a data dissemination model. The servers are sinks for data like kitchen sinks for water. The mobile nodes diffuse messages within a certain space. If the mobile nodes are close to a server, their messages will be absorbed by the server. A fixed number of nodes are studied [31]. Important point to note here is that the population in VANETs is not fixed because of the dynamics of traffic.

Bass et al. proposed the well-known Bass model [13] in 1960s’. The model states a purchase process in a potential customer population. A group of customers will buy a new product without consulting any others. The rest of the customers will consult the previous buyers about the new product and decide to buy the product. A new product diffusion with price is discussed in Bass model [13]. The Bass model describes the process of how new products get adopted among users as an interaction process between adopted users and potential users. The innovation is the number of users that will directly buy the product without consulting anybody. The imitation is the number of users who will buy the product after consulting adopted users. In our previous work, dissemination of traffic-event notification [32,33] has been studied by adapting the BASS model in vehicular networks. The mathematical model is the following:

$$\frac{f(t)}{1-F(t)} = p + qF(t) \quad (1)$$

Where:

1. $f(t)$ is the rate of change of the installed base fraction
2. $F(t)$ is the installed base fraction
3. p is the coefficient of innovation
4. q is the coefficient of imitation

The growth rate of purchase $S(t)$ is the rate of change of installed base (i.e. adoption) $f(t)$ multiplied by the ultimate market

Table 1
Bass model versus alert message diffusion [11].

	Bass model	Alert message diffusion
m	The ultimate market potential	The total number of transmissions
$f(t)$	The rate of adoption fraction	The change of alert message dissemination
$F(t)$	The adoption fraction	The finished dissemination
p	The coefficient of innovation	The fraction/coefficient of innovators
q	The coefficient of imitation	The fraction/coefficient of imitators
$S(t)$	The rate of change of adoption	The rate of change of dissemination
$x(t)$	Price effect on product adoption	Alert message value effect on dissemination
$X(t)$	Cumulative of price effect	Cumulative of value effect

potential m.

$$S(t) = mf(t) \tag{2}$$

Since the Bass model is a business model, it is not directly applicable to VANETs. In this paper, we analogize the purchase process as a data diffusion process because of the similarity of the two models.

3. Data diffusion model as a new product adoption process

The diffusion of a new product among people [13,34] is similar with the diffusion of an alert message among vehicles. Both are processes in which given a contact something is communicated. When a new product spreads quickly and is distributed among people, it is called a purchase. Our simulations of alert messages in VANETs show alike behavior as well.

Existing mathematical models that describe a new product purchase can be as useful for use as they do for marketing science. Researchers in marketing use the Bass model to describe the process of how new products get adopted as an interaction between users and potential users. We use Bass model to describe the alert message dissemination in VANETs.

We adopt the new product adoption process to data diffusion model at macro level instead of micro level. We notice that a customer of products can decide to buy or not buy a product at time t. The customer's decision ultimately is affected by certain conditions or resources such as money, timing, personality, etc. Similarly, vehicles' adoption of a message is affected by certain conditions and resources such as network connection, distance, timing, etc. If we look at the two models at the micro level, there are many differences. However, if we compare the two models at macro level, they are very similar. At time t, the rate of adoption of the whole population is a function of time t. The coefficient of innovation is a percentage of the population. Likewise, the coefficient of imitation is a percentage of the population. Therefore, we can introduce the new product adoption process model into data diffusion model.

The mapping of parameters in the Bass model to our alert message dissemination is shown in Table 1 [11], along with the corresponding meaning for alert message dissemination.

Another important similarity between Bass model and alert message dissemination is the concepts of innovators and imitators when we compare the two models at macro level. The coefficients of innovation and imitators are percentages of the population. The innovators are the initial customers who buy the new product in Bass model. In our alert message dissemination, the innovators are the initial vehicles which receive alert message of an accident. The imitators in Bass model are the customers who are advised by the innovators. In our alert message dissemination, the imitators are the vehicles which receive alert message from innovator vehicles. Fig. 2 shows the concept of innovators and imitators in VANETs.

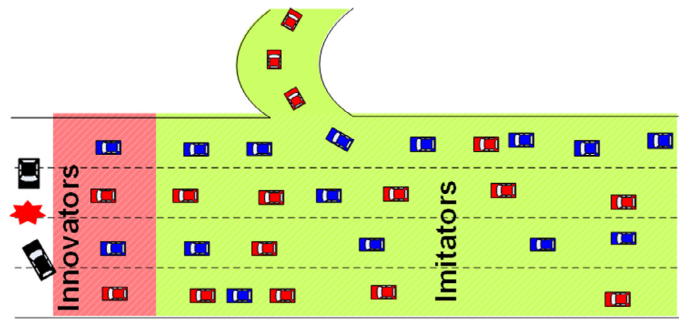


Fig. 2. The innovators and imitators in VANETs.

3.1. System model

We focus on large-scale VANETs composed of a large number of vehicles. A network considered within scope of this paper is a set of nodes in a big area (downtown) where vehicles can communicate with each other either directly or through multi-hops with a certain time interval. Other topologies of VANETs and other geometry or maps can be further explored in the future. A population refers to the set of vehicles that is part of the network. The vehicles form a population network through wireless transceivers by communicating either with other vehicles or with roadside infrastructures.

The mobility model of vehicles will affect the dissemination of alert messages. The mobility model includes velocity, accelerate/decelerate speed, car following etc. We view the vehicle mobility on a macro level and we use average speed for all vehicles, similar to the speed limit in a downtown or highway. Communication parameters like transmission range will affect the alert message dissemination as well. In our model, all vehicles can communicate each other directly or indirectly. In this paper, we consider a diffusion algorithm for alert messages. We are interested in modeling the dissemination rate of alert messages, i.e. the average alert message receivers in a considered time interval (0, t).

3.2. Technique details

The diffusion process of disseminating alert message among vehicles is similar to the diffusion process of disseminating a new product among people. The diffusion model of a new product was first proposed by Bass et al. (1969) [13]. We have a major assumption: the probability that an initial dissemination will be made at t given that no dissemination has yet been made is a linear function of the number of previous receivers. This assumption has been validated by many scenarios in marketing science. Thus, $Pr(t) = p + \frac{q}{m}Y(t)$, where p and $\frac{q}{m}$ are constants and $Y(t)$ is the number of previous receivers. Since $Y(0) = 0$, the constant p is the probability of an initial dissemination at $t = 0$ and its magnitude reflects the importance of innovators in the social system. It is possible to elect a unit of measure for time such that p reflects the fraction of all receivers who are innovators. The product $\frac{q}{m}Y(t)$ reflects the pressures operating on imitators as the number of previous receivers increases. Over time *T*, there will be *m* initial disseminations of the alert messages.

The likelihood of transmission at time *t* given that no transmission has yet been made is

$$\frac{f(t)}{1-F(t)} = Pr[t] = p + \frac{q}{m}Y(t) = p + qF(t) \tag{3}$$

where $f(t)$ is the change of dissemination at time *t* and the finished dissemination $F(t)$ is defined as

$$F(t) = \int_0^t f(t)dt, F(0) = 0. \tag{4}$$

which means the summation of all changes over time [0, *t*]. *m* is the total number of transmissions during the period for which the density function was constructed,

$$Y(t) = \int_0^t S(t)dt = m\int_0^t f(t)dt = mF(t) \tag{5}$$

is the total number of transmission in the interval (0, *t*). *q* is the fraction of imitators. The meaning of $qY(t)$ is the total number of imitators. The probability of imitators, thus, is $\frac{q}{m}Y(t)$.

Therefore,

$$S(t) = mf(t) = Pr[t] \times (m - Y(t) \tag{6}$$

$$= \left(p + \frac{q}{m}\int_0^t S(t)dt\right)\left(m - \int_0^t S(t)dt\right). \tag{7}$$

Expanding the product we have

$$S(t) = pm + (q-p)Y(t) - \frac{q}{m}Y(t)^2. \tag{8}$$

This differential equation has been solved by Bass [13]. We directly apply the solution in this paper,

$$f(t) = \frac{(p+q)^2}{p}\frac{e^{-(p+q)t}}{(1+e^{-(p+q)t})^2} \tag{9}$$

Since $S(t) = mf(t)$, we obtain

$$S(t) = m\frac{(p+q)^2}{p}\frac{e^{-(p+q)t}}{(1+e^{-(p+q)t})^2} \tag{10}$$

There are two special cases of the above model. The first special case occurs when $q = 0$.

Lemma 1. *When* $q = 0$, *the model reduces to Exponential distribution.*

Proof. We can obtain $F(t)$ from $f(t)$,

$$F(t) = \int_0^t f(x)dx \tag{11}$$

$$= \int_0^t \frac{(p+q)^2}{p}\frac{e^{-(p+q)x}}{(1+e^{-(p+q)x})^2}dx \tag{12}$$

$$= \frac{1-e^{-(p+q)t}}{1+\frac{q}{p}e^{-(p+q)t}} \tag{13}$$

When $q = 0$,

$$F(t) = \frac{1-e^{-(p+0)t}}{1+\frac{0}{p}e^{-(p+0)t}} \tag{14}$$

$$= 1 - e^{-pt} \tag{15}$$

By definition, the model is Exponential distribution. □

When $q = 0$, the model reflects the behavior of innovators. This scenario indicates that no message relay is permitted. The alert message's dissemination depends on the contact of the alert source site/vehicles.

The second special case occurs when $p = 0$.

Lemma 2. *When* $q = 0$, *the model reduces to Logistic distribution.*

The model reduces to the Logistic distribution, when $p = 0$.

Proof. When $p = 0$,

$$\frac{f(t)}{1-F(t)} = p + qF(t) = qF(t) \tag{16}$$

Therefore,

$$f(t) = \frac{F(t)}{dt} = qF(t) - qF(t)^2 \tag{17}$$

The solution is

$$F(t) = \frac{1}{1+C\times e^{-qt}}. \tag{18}$$

To get *C*, we use the boundary condition $F(0) = 0$,

$$F(0) = 0 \tag{19}$$

$$= \frac{1}{1+C\times e^{-q\times 0}} \Rightarrow C = -1 \tag{20}$$

By definition, this is a logistic distribution. □

When $p = 0$, the model reflects the behavior of imitators. This scenario indicates that no direct message dissemination is permitted. The alert message's dissemination depends on the contact of the relay vehicles.

The traffic flow density ρ is related to *m*. We define the traffic flow density as vehicles per mile. We obtain

$$m = \rho \times d_s \tag{21}$$

where d_s is the safe distance of the accident. For an accident, there is a certain distance within which the drivers should be alerted, for example, 5 miles. The drivers, being far away d_s, will not care about the accident. In reality, d_s depends on many parameters such as road maps, the types of messages, etc. We use five miles just as an example of many possible values. *m* represents the total potential vehicles that can receive the message at macro level where all vehicles are in multi-hop transmission.

3.3. Alert message's value decreasing with time

The alert message contains the information about an accident. The value of the alert messages is different to vehicles. As shown in Fig. 3, we partition a road with an accident into three parts: *S*1, *S*2, and *S*3. The value of alert message is high for vehicles in *S*1 because these vehicles are close to the accident. The drivers have to be aware of the accident. The value of alert message is medium for vehicles in *S*2. Although the information of accident contained in the alert message is not urgent, these drivers in *S*2 aware of the accident and can make the decision to stay on the road or take an exit. For the vehicles in *S*3, the value of the accident is low because it takes a certain time to reach the accident. Therefore, these drivers will not take it as urgent.

It is noticeable that the number of vehicles in each of {*S*1, *S*2, *S*3} is not static because of vehicles' mobility and traffic's dynamics. The lengths of {*S*1, *S*2, *S*3} is actually dynamical and depends on the types of message, traffic density, and speed of vehicles. In this paper we are interested in the data diffusion at macro level, i.e. at time *t*, how many vehicles will receive this message. The design of {*S*1, *S*2, *S*3} can be explored in the future.

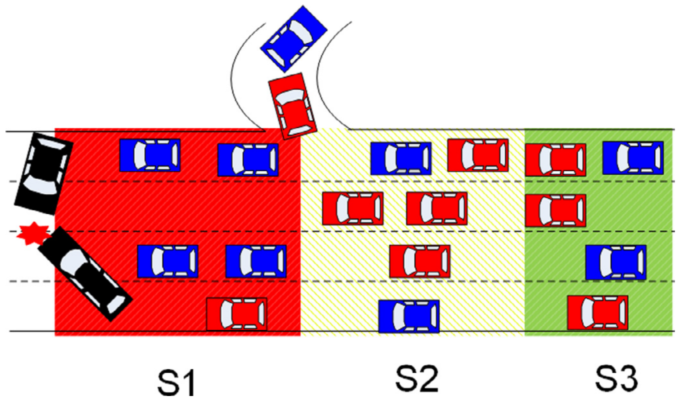


Fig. 3. The value of alert message decreasing with time (three stages).

3.4. Diffusion model with value of messages

We model the diffusion model with message value decreasing with time in this section. The diffusion model with message value is similar with the generalized Bass model (called GBM) developed by Bass et al. (1994) [35]. With consideration of the value of alert message, we write:

$$\frac{f(t)}{1-F(t)} = [p + qF(t)]x(t) \tag{22}$$

where $f(t)$ is the density function of the random variable t, the time to first dissemination of the alert message. $F(t)$ is the cumulative distribution function, p and q are the diffusion parameters, and $x(t)$ is the current value of the alert message. The value of the alert message is specified as the following:

$$x(t) = 1 + \beta \frac{V'(t)}{V(t)} \tag{23}$$

where $V(t)$ is value at time t and $V'(t)$ is the change in value at time t. We define

$$V(t) = \frac{V(0)}{(1-\eta)^t} \tag{24}$$

$$v(t) = V'(t) = \frac{V(0)(1-\eta)^t}{ln(1-\eta)} \tag{25}$$

where η is the percentage of value deduction per time units. Therefore, we obtained

$$x(t) = 1 - \beta \times \frac{1}{ln(1-\eta)} \tag{26}$$

We refer β as the diffusion value parameter and it reflects the effect of value in accelerating and decelerating the diffusion process. Based on Bass, β is expected as a negative value [35]. In simulation, we showed the impact of β to the system.

Solving the differential Eq. 22 and using the boundary condition $F(0) = 0$, we obtain

$$F(t) = \frac{1 - e^{-(X(t)-X(0))(p+q)}}{1 + \frac{q}{p} e^{-(p+q)(X(t)-X(0))}} \tag{27}$$

$$f(t) = F'(t) \tag{28}$$

$$= x(t) \frac{(p+q)^2}{p} \frac{e^{-(p+q)(X(t)-X(0))}}{(1 + \frac{q}{p} e^{-(p+q)(X(t)-X(0))})^2} \tag{29}$$

where $X(t) = \int_0^t x(\tau)d\tau$, is referred to as the cumulative value of the alert message. From Eq. 23. it is easy to see that

$$X(t) = t(1 - \beta \times \frac{1}{ln(1-\eta)}) \tag{30}$$

Therefore,

$$S(t) = mf(t) \tag{31}$$

$$= m \times x(t) \times \frac{(p+q)^2}{p} \frac{e^{-(p+q)X(t)}}{(1 + \frac{q}{p} e^{-(p+q)X(t)})^2} \tag{32}$$

where m is the total number of ultimate transmission of the alter messages and we assume $X(0) = 0$.

4. Data diffusion model as a stochastic process

In this section, we extend epidemic models [21,22] to VANETs. We assume a data diffused on a closed system with a population of N vehicles. In this system, there is initially one vehicle *received ("infected")* a message and all the other $N - 1$ vehicles are *unaware* of the message. Unaware vehicles that come in *"contact"* with received vehicles can adopt the message and, in time, can spread it to other vehicles. In the model we adopt this approach, adopted vehicles will remain aware of the message. More formally, consider a stochastic process $\{N(t) | t \geq 0\}$ of continuous parameter t, where for every positive integer $k, (1 \leq k \leq N)$, the event $\{N(t) = k\}$ occurs if the population contains k adopted vehicles at time t.

We assume that at time $t = 0$ there is a *single* adopted vehicle. Further, we let $P_k(t)$ denote the probability that the event $\{N(t) = k\}$ occurs. In other words,

$$P_k(t) = \Pr[\{N(t) = k\}].$$

4.1. Simple model

It is fairly obvious that for $k > 1$ and for a small $h > 0$, $P_k(t+h)$ has the following components:

- $P_k(t)[1 - h\lambda_k + o(h)]$, which describes the probability of staying at state k.
- $P_{k-1}(t)[h\lambda_{k-1} + o(h)]$, which describes the probability of reaching to the state k from state $k - 1$.

This allows us to write

$$P_k(t+h) = P_k(t)[1 - h\lambda_k + o(h)] + P_{k-1}(t)[h\lambda_{k-1} + o(h)] + o(h) \tag{33}$$

$$= P_k(t)[1 - h\lambda_k] + P_{k-1}(t)h\lambda_{k-1} + o(h). \tag{34}$$

Transposing $P_k(t)$ and dividing with h yields

$$\frac{P_k(t+h) - P_k(t)}{(t+h) - t} = -\lambda_k P_k(t) + \lambda_{k-1} P_{k-1}(t) + \frac{o(h)}{h}.$$

Taking limits on both sides of this equality as $h \to 0$ yields the differential equation

$$\frac{dP_k(t)}{dt} = -\lambda_k P_k(t) + \lambda_{k-1} P_{k-1}(t) \tag{35}$$

with the boundary condition $P_k(0) = 0$.

Proceeding similarly, when $k = 1$, we write

$$P_1(t+h) = P_1(t)[1 - h\lambda_1 + o(h)] + o(h) \tag{36}$$

$$= P_1(t)[1 - h\lambda_1] + o(h). \tag{37}$$

Transposing $P_1(t)$ and dividing with h yields

$$\frac{P_1(t+h) - P_1(t)}{(t+h) - t} = -\lambda_1 P_1(t) + \frac{o(h)}{h}.$$

Taking limits on both sides of this equality as $h \to 0$ yields the differential equation

$$\frac{dP_1(t)}{dt} = -\lambda_1 P_1(t) \tag{38}$$

with the boundary condition $P_1(0) = 1$. This latter equation can be easily solved to obtain

$$P_1(t) = e^{-\lambda_1 t}. \tag{39}$$

Having obtained $P_1(t)$, it is easy to compute $P_2(t)$ and $P_3(t)$. To summarize, we state the following general result.

Theorem 1. *For all* $t \geq 0$,

$$P_k(t) = \begin{cases} \mu_1\mu_2\cdots\mu_{k-1}\sum_{i=1}^{k}\dfrac{e^{-\mu_i\int_0^t \alpha(u)\,\mathrm{d}u}}{\prod\limits_{\substack{j=1\\ j\neq i}}^{k}(\mu_j-\mu_i)} & \text{for } 1 \leq k \leq N-1 \\ 1-\sum_{i=1}^{N-1}\left(\prod\limits_{\substack{j=1\\ j\neq i}}^{N-1}\left[\dfrac{\mu_j}{\mu_j-\mu_i}\right]e^{-\mu_i\int_0^t\alpha(u)\,\mathrm{d}u}\right) & \text{for } k = N. \end{cases} \tag{40}$$

As it turns out, the iterative technique outlined above together with an obvious inductive argument leads to the following general result.

Theorem 2. *For* $1 < k < N$ *and* $t \geq 0$,

$$P_k(t) = \lambda_1\lambda_2\cdots\lambda_{k-1}\sum_{i=1}^{k}\frac{e^{-\lambda_i t}}{\prod_{\substack{1\leq j\leq k\\ j\neq i}}(\lambda_j-\lambda_i)}. \tag{41}$$

For the details of these derivations, the reader is referred to the supplemental file that accompanies this paper.

We also can obtain the closed form of $P_N(t)$.

$$P_N(t) = 1 - \prod_{\substack{1\leq j\leq N-1\\ j\neq i}}\left[\frac{\lambda_j}{\lambda_j-\lambda_i}e^{-\lambda_i t}\right]. \tag{42}$$

4.2. Data dissemination model with leave

In this section, we develop a data dissemination model with considering vehicles leaving the area of interest. The model is a little similar to the Kermack and McKendrick SIR (Susceptible, Infected, Removed) model [36]. To make the mathematical derivations more manageable, we set $P_k(t) = 0$ for $k < 0$ and $k > N$. Thus, $P_k(t)$ is well defined for all integers $k \in (-\infty, \infty)$ and for all $t \geq 0$. In particular, the assumption about the initial population containing a $n_0 = 1$ vehicle at $t = 0$ translates into $P_1(0) = 1$ and 0 otherwise.

Let $t, (t \geq 0)$, be arbitrary and let h be sufficiently small that in the time interval $[t, t+h]$ the probability of two or more "infections" or recoveries which will leave from the system, or of a simultaneous infection and recovery, is $o(h)$. The recoveries are the vehicles which receive an alert message and take actions such as exiting the highway. With h chosen as stated, the probability $P_k(t+h)$ that the number of alerted vehicles $k, (0 \leq k \leq N)$, at time $t+h$ has the following components:

- $P_k(t)[1 - hk\alpha - kh\beta + o(h)]$
- $P_{k-1}(t)[h(k-1)\alpha + o(h)]$
- $P_{k+1}(t)[h(k+1)\beta + o(h)]$.

Visibly,

$$P_k(t+h) = P_k(t)[1 - h(k)\alpha - kh\beta] + P_{k-1}(t)h(k-1)\alpha + P_{k+1}(t)(k+1)h\beta + o(h) \tag{43}$$

$$= P_k(t)\left[1 - h\frac{N-k}{N}\lambda(t)\right] + P_{k-1}(t)h\frac{N-k+1}{N}\lambda(t) \tag{44}$$

$$+ (k+1)h\mu(t)P_{k+1}(t) + o(h). \tag{45}$$

After transposing $P_k(t)$ and dividing by h we have

$$\frac{P_k(t+h) - P_k(t)}{(t+h) - t} = -[k\alpha + k\beta]P_k(t) + (k+1)\beta P_{k-1}(t) + P_{k+1}(t)(k+1)\beta + \frac{o(h)}{h}.$$

Taking limits on both sides as $h \to 0$ yields the differential equation

$$\frac{\mathrm{d}P_k(t)}{\mathrm{d}t} = -[k\alpha + k\beta]P_k(t) + (k-1)\alpha P_{k-1}(t) + (k+1)\beta P_{k+1}(t) \tag{46}$$

with the initial condition $P_k(0) = 1$ for $k = n_0$ and 0 otherwise. Using standard technique this differential equation is very hard to solve and getting a closed form for expected value is not practically possible. Bailey [37] shows a complicated process to generate first few state probabilities but he does not go beyond $k = 4$. We used numerical solution which in turn gave us all the state probabilities P_k. Thus we can calculate the expected value from the state probabilities.

5. Numeric results and simulations

5.1. Simple diffusion model

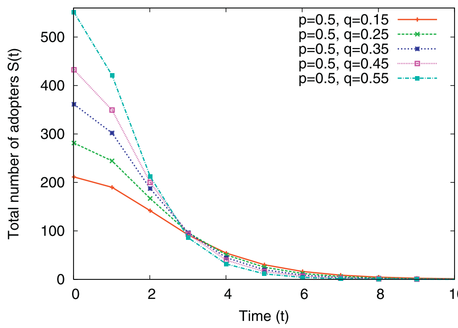
Using the results in Section 3, in this section we select various values of the system parameters, and we subsequently evaluate the growth rate of messages disseminated among vehicles for both the simple diffusion model and the diffusion with value model.

Approach: In the simple diffusion model, given m, we assume that there exists some time t ranging in (0, t_u] where t_u is the upper bound of the time, $t_u = 100$ or $t_u = 20$. Subsequently, we select various p values in [0.05, 0.4]. For each such value, we select various q values in [0.1, 0.4]. Then we evaluate the growth rate $S(t)$ of vehicles which received the alert message. Given p and q, we select various m values in [400, 1000] and we evaluate the values of $S(t)$. Each combination of these parameters generates one set of results. In the diffusion model, given p, q, m, we vary time values in (0, t_u]. Given m, η, we select various p values in [0.05, 0.4]. For each of these values, we select various q values in [0.1, 0.4]. We obtain a curve for each of the combination of these parameters. The constant β is set as -0.7973 and $\eta = 0.2$ initially.

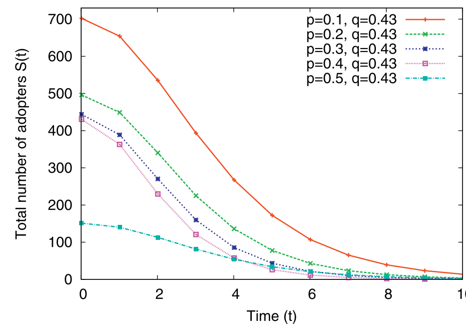
The rate p of the simple diffusion model is generally low, as compared to that of q, because the amount of initial vehicles witnessing or experiencing the accident is low. We selected a p value equal to 0.5, $m = 1000$ and varied q values in a set {0.15, 0.25, 0.35, 0.45, 0.55}. The result is shown in Fig. 4(a). When $p < q$, the imitators are dominating the system. The bigger q, the more contribution imitators make. The value of $S(t)$ increases at the beginning because more imitators are involved. Once the peak is reached, the value of $S(t)$ sharply drops because the rest of the vehicles that do not receive alert messages decrease quickly. If $p \geq q$, the innovators will be the major strength to contribute the value of $S(t)$. When we select various values of p in a set {0.1, 0.2, 0.3, 0.4, 0.5}, given $q = 0.43$ (shown in Fig. 4(b)), the value of $S(t)$ decreases monotonously because the innovator is dominating. We found m could affect the value $S(t)$ in Fig. 4(c). The bigger the m, the larger the value of $S(t)$.

When we selected $p = 0.5, q = 0.05$, $m = 1000$ and varied β, we found β could affect the value $S(t)$ in Fig. 5. The value of $S(t)$ decreases monotonously. Smaller β results in bigger number of imitators at the initial stage of the model.

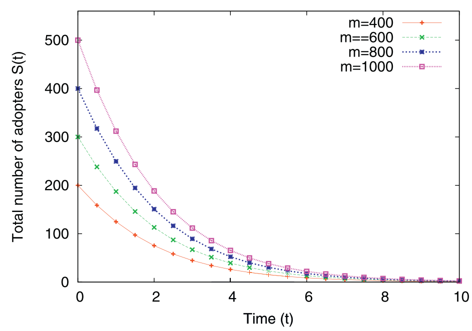
In earlier Section 3, we addressed the relationship between traffic flow density and m. The numerical analysis gives researches the relationship between traffic density and $S(t)$. In Fig. 6(a), we show the percentage deduction of the value of the alert messages. We



(a) Case 1.



(b) Case 2.



(c) Case 3.

Fig. 4. The growth rate of propagated message $S(t)$.

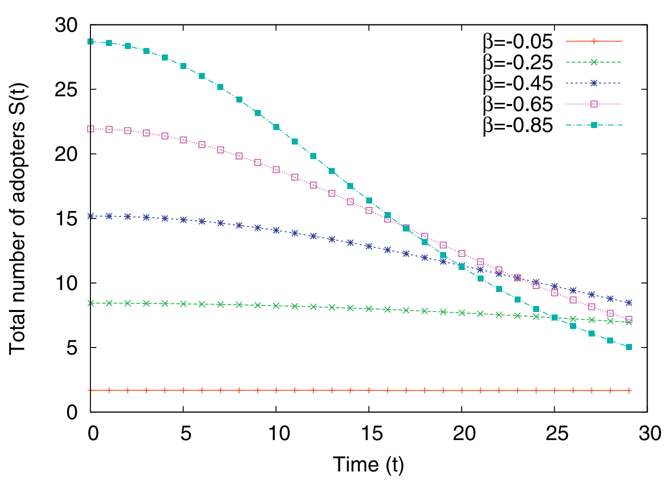


Fig. 5. The total number of message recipients $S(t)$ vs. time t in diffusion model with value depreciation. $p = 0.5, q = 0.05, m = 1000$.

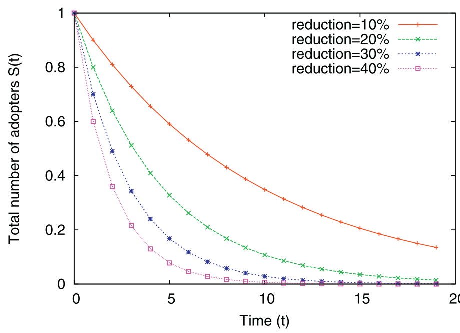
check how the percentage deduction affects the value $S(t)$ based on the result of Fig. 6(a). The values of p, q, m are set to 0.01, 0.1 and 1500 respectively. As expected, the shape of $S_v(t)$ becomes more flat when η increases in Fig. 6(b). When $p < q$, the imitators contribute more to the diffusion. We set m and p as 1500 and 0.01, ranging q as 0.05, 0.1, 0.2, and 0.3. The result shown in Fig. 4 validates the imitators' contribution discussed earlier. Since we have two models, we compare the performance of these two models. The constant parameters are set to $p = 0.05, q = 0.2, m = 1500, \eta = 0.1$. We found the curve of simple diffusion model is above the one of diffusion model with message value, as shown in Fig. 6(c). This is because we will drop the message if the value of the message is less than a threshold. Therefore, some vehicles will not be counted even if they receive alert messages with low value.

Based on the initial simulator [38], we use a tailored Java simulation engine. We assume a constant arrival rate λ, a service rate μ, a potential amount of vehicles m, and a percentage value decreasing per time unit η. Vehicles' mobility model is taken as IDM [38]. At the beginning of the simulation run, few nodes witnessed an accident happened (thus, they are innovators). The number of witnesses is determined by the value of $p \times m$. Then innovators will start to distribute alert messages. Other vehicles which receive this message and adopt it become imitators. The percentage of adoption of these vehicles is q. The value of the alert message decreases with time unit. Once the value of alert messages decreases to 10% of original message, it is dropped. The total number of adoption of alert message is of interest. In the first simulation, we set constant $p = 0.05, q = 0.2, m = 1500$ and change time t in the interval from 0 to 20. As expected, the simulation results match with the theoretical values, as shown in Fig. 7.

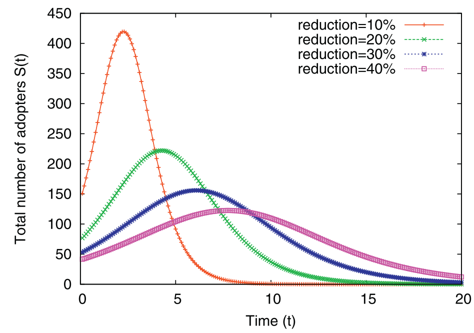
In our second simulation, we set constant $p = 0.01, q = 0.1, m = 1500$, and $\eta = 20\%$ and change time t in the interval 0–20. The simulation results are close to the theoretical values, as shown in Fig. 8. From the results shown in Figs. 7 and 8, we observe that the peak of $S(t)$ are both near time 5. This means the alert messages disseminate at a quick speed. Faster speed of transmission of alert messages means more time availability for other drivers on the road to make a decision.

5.2. Stochastic model simulation

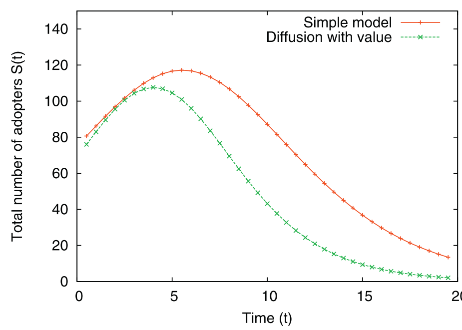
In this section, we are interested in computing the expected number of adopters $S(t)$ and the probability $p_k(t)$. The results are numerical and are obtained with aid of computers. We assumed that there is an accident appended and initially is only one vehicle sending out the alert message. We varied the value of time t from



(a) Case 1.



(b) Case 2.



(c) Case 3.

Fig. 6. The number of message $S(t)$. To differentiate to Fig. 4, we use $S_v(t)$ to represent the number of vehicles received alert message in the second model: diffusion model with value.

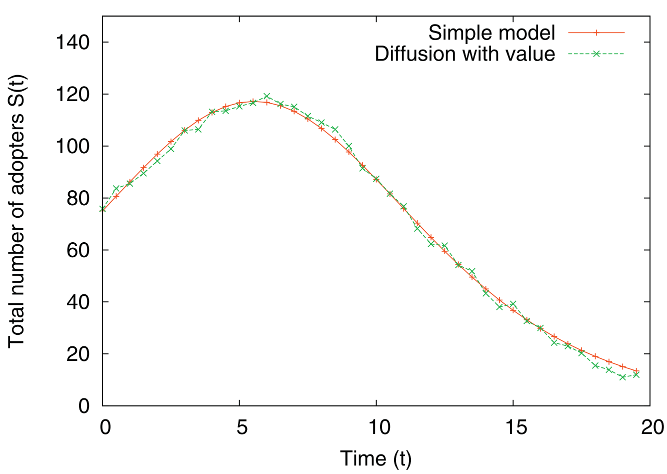


Fig. 7. The total number of message recipients $S(t)$ vs. time t in simple diffusion simulation. $p = 0.05, q = 0.2, m = 1500$.

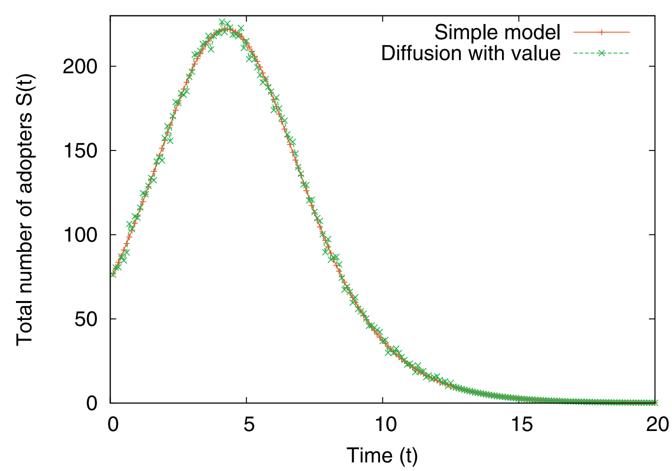


Fig. 8. The total number of message recipients $S(t)$ vs. time t in simple diffusion simulation. $p = 0.01, q = 0.1, m = 1500$, and $\eta = 20\%$.

0 to 10 s. The independent system parameters are the time t, the rate λ, the service rate μ, the potential amount of vehicles m, and the percentage of value decreasing per time unit η. The infection rate of an alert message is also varied from 1.91 to 2.91 and the recovery rate is varied from 0.56 to 1.86. We then collected the expected number of adopters. We first computed the probability value of each state, i.e. the value of $P_k(t)$. Once the state probabilities are computed, we computed the total expected number, $S(t)$ or $E[N(t)]$, of vehicles at any given time t, by the well-known formula

$$S(t) = E[N(t)] = \sum_{k=1}^{N} kP_k(t). \tag{47}$$

Other configuration settings are discussed in the previous Section 5.1.

We plotted $P_k(t)$ as shown in Fig. 9. As expected, we noticed the peak value of $P_k(t)$ decreases when k increases. That means the high value of $P_k(t)$, i.e. the likelihood that k vehicles receive the alert message, will drop when k increases. The larger the number of vehicle adopters of the message, the smaller the probability values of $P_k(t)$. Interestingly, we noticed that vehicles will not receive the alert message when time t is big enough as Fig. 9 shows that $t \approx \infty \Rightarrow P_k(t) \approx 0$. This also makes sense. All vehicles in our zone of interest will receive the alert message if t is big enough. When $t <$ 1 and k increases, the time t where the peak of $P_k(t)$ is associated

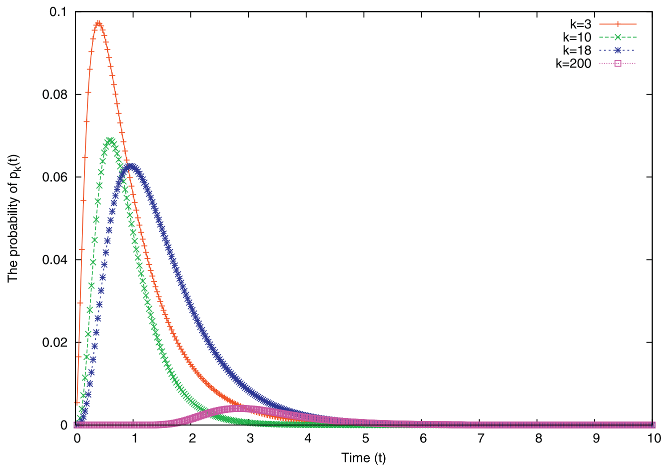


Fig. 9. The probability of $p_k(t)$.

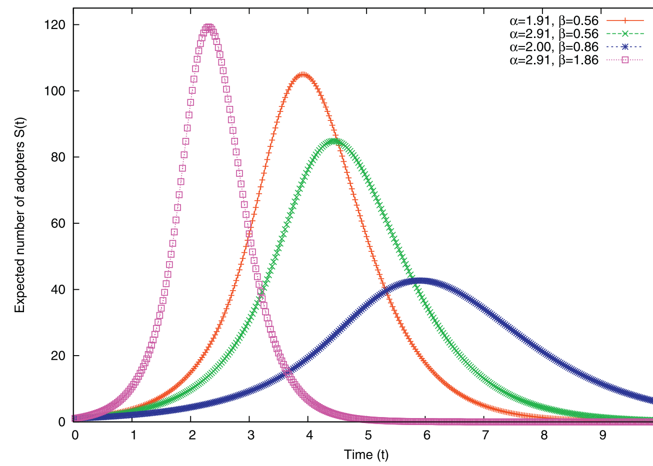
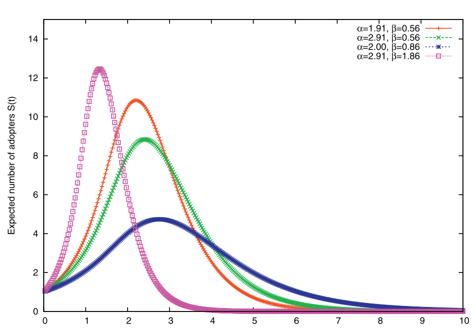


Fig. 10. The expected number of adopters $S(t)$ when the population is $N = 500$.

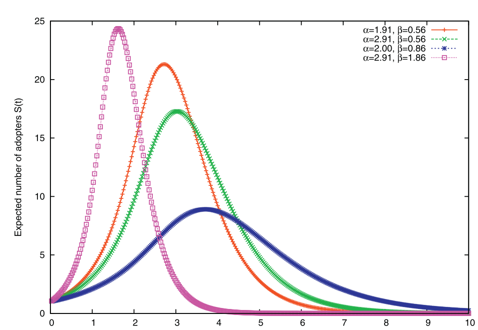
increases. The more vehicles receive the alert message at [0, t], the longer time it will take.

The total expected number of adopters $S(t)$ of the alert message is shown in Fig. 11. The discussion of the stochastic process in a complex situation can be referred in Section 4.2. We obtained values of $S(t)$ by computer approximation. We varied the value of time and the parameters of α and β but made n as a fixed number, i.e. 500. To make the values comparable, we intentionally selected four sets of results as shown in Fig. 10. As expected, we noticed that all sets of data showed a similar pattern. When time t increases, $S(t)$ increases initially, then reaches the peak value and drops to very small value near zero. It shows that the total number of new vehicles receiving the alert message first increases and then drops because the pool of unaware vehicles of the message decreases when time passes by quickly. We also noticed that when β is a constant value 0.56, the higher value of α will result lower peak value of $S(t)$. But the higher value of α will cause larger value of time t when $S(t)$ reaches a peak. On the other hand, if α is a fixed value 2.91, the higher value of β will result in higher value of $S(t)$. But the higher value of β will cause the longer time t when $S(t)$ reaches a peak. Therefore, the impacts of α and β are different.

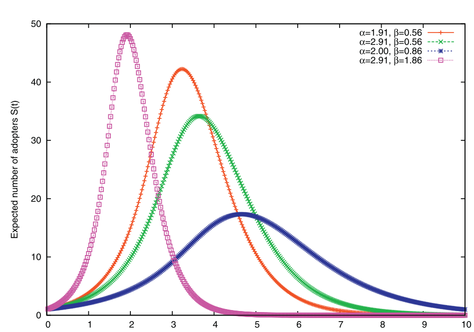
The time when $S(t)$ reaches its peak value is vital as half of the total recipients have received the alert message in Fig. 10. We explored cases in which the peak value of $S(t)$ will be reached by varying the number N as 50, 100 and 200. We plotted three sets of values of $S(t)$ where $N = 50, 100, 200$ respectively. We noticed that all three sets of $S(t)$ values showed a similar pattern. When time t increases, $S(t)$ increases initially, then reaches the peak value and drops to very small value near zero. We also noticed that when N increases, the peak value of $S(t)$ increases and the time t when $S(t)$



(a) Case 1. N=50



(b) Case 2. N=100



(c) Case 3. N=200

Fig. 11. The expected number of adopters $S(t)$, the number of vehicles received alert message in the stochastic model.

reaches the peak increases as well. The results make sense because it takes longer time to allow larger number of vehicles to receive the alert message. The larger the number of vehicles receiving the alert messages means the higher peak value of expected the number vehicles receiving the message at any time t.

6. Conclusions and future work

A brief idea on the limitations of the model and their influence on the analysis, performance and differences with respect to a more realistic scenario should be drawn.

We demonstrated mainly two data dissemination analytical models for VANETs. We started by modeling VANETs data dissemination as a new product adoption model. To make our model more realistic, we considered the fact that the value of alert messages decreases with time in the model. This model only works at a macro level instead of micro level. Therefore, this model can be used to predict the speed of data dissemination or the total vehicles receiving a message. Applications that expect precise prediction of peer-to-peer communication will work with this model.

Additionally we adopted a time-dependent stochastic epidemic model. We can probabilistically capture the spread of traffic messages. This model can be applied on applications that expect probability metrics such as expected number of vehicles receiving a message at time t, the probability that a message can reach to n vehicles at time t, etc. This model works at a macro level. So applications expecting to find the probability of communications at a micro level can be referred to our previous work [11,32].

In the future, we will compare the proposed models with real traffic data to validate our analytical models. Other topologies of VANETs and other geometry or maps can be further explored in the future as well. In addition, we are in the process of showing the relevance of our model to growth phenomena in software engineering, in macroeconomics, in the dissemination of traffic-event notification and the spread of computer viruses in social networks or smart phones. This promises to be an exciting area for future investigations.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.adhoc.2016.09.010

References

[1] U. D. of Transportation, Standard specification for tele- communications and information exchange between roadside and vehicle systems - 5 ghz band dedicated short range communications (dsrc) medium access control (mac) and physical layer (phy) specifications, Aug, 2003,
[2] M. Abuelela, S. Olariu, Content delivery in zero-infrastructure vanet, Vehicular Networks: From Theory to Practice, Taylor & Francis, 2009, pp. 8.1–8.15.
[3] G. Yan, S. Olariu, A probabilistic analysis of link duration in vehicular ad hoc networks, IEEE Trans. Intell. Transp. Syst. 12 (4) (2011) 1227–1236.
[4] H. Hartenstein, K.P. Laberteaux, A tutorial survey on vehicular ad hoc networks, IEEE Commun. Mag. 46 (6) (2008) 164–171.
[5] S. Yousefi, M.S. Mousavi, M. Fathy, Vehicular ad hoc networks (vanets): challenges and perspectives, in: ITS Telecommunications Proceedings, 2006 6th International Conference on, IEEE, 2006, pp. 761–766.
[6] Y. Toor, P. Muhlethaler, A. Laouiti, Vehicle ad hoc networks: applications and related technical issues, IEEE Commun. Surv. Tut. 10 (3) (2008) 74–88.
[7] F. Li, Y. Wang, Routing in vehicular ad hoc networks: a survey, IEEE Veh. Technol. Mag. 2 (2) (2007) 12–22.
[8] W. Chen, R.K. Guha, T.J. Kwon, J. Lee, Y.-Y. Hsu, A survey and challenges in routing and data dissemination in vehicular ad hoc networks, Wireless Commun. Mobile Comput. 11 (7) (2011) 787–795.
[9] A.K. Saha, D.B. Johnson, Modeling mobility for vehicular ad-hoc networks, in: Proceedings of the 1st ACM International Workshop on Vehicular Ad Hoc Networks, ACM, 2004, pp. 91–92.
[10] J. Härri, F. Filali, C. Bonnet, Mobility models for vehicular ad hoc networks: a survey and taxonomy, IEEE Commun. Surv. Tut. 11 (4) (2009) 19–41.
[11] G. Yan, S. Olariu, S. Salleh, A probabilistic routing protocol in vanet, in: Proceedings of the 7th International Conference on Advances in Mobile Computing and Multimedia (MoMM2009), Kuala Lumpur, Malaysia, 2009.
[12] G. Yan, S. Arif, D.B. Rawat, Modeling message dissemination in ad hoc networks, in: 2014 International Conference on IET Information and Communications Technologies (ICT 2014), 2014, pp. 1–5.
[13] F.M. Bass, A new product growth for model consumer durables, Manage. Sci. 50 (12 Supplement) (2004a) 1825–1832.
[14] Y.-W. Lin, Y.-S. Chen, S.-L. Lee, Routing protocols in vehicular ad hoc networks: a survey and future perspectives., J. Inf. Sci. Eng. 26 (3) (2010) 913–932.
[15] K.C. Lee, U. Lee, M. Gerla, Survey of routing protocols in vehicular ad hoc networks, in: Advances in Vehicular Ad-Hoc Networks: Developments and Challenges, IGI Global, 2010, pp. 149–170.
[16] D.B. Rawat, Y. Zhao, G. Yan, M. Song, Crave: cognitive radio enabled vehicular communications in heterogeneous networks, in: Radio and Wireless Symposium (RWS), 2013 IEEE, IEEE, 2013, pp. 190–192.
[17] M. Nekoui, A. Eslami, H. Pishro-Nik, Scaling laws for distance limited communications in vehicular ad hoc networks, in: Communications, 2008. ICC'08. IEEE International Conference on, IEEE, 2008a, pp. 2253–2257.
[18] M. Nekoui, A. Eslami, H. Pishro-Nik, The capacity of vehicular ad hoc networks with infrastructure, in: Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks and Workshops, 2008. WiOPT 2008. 6th International Symposium on, IEEE, 2008b, pp. 267–272.
[19] G. Yan, S. Olariu, S. El-Tawab, Ticket-based reliable routing in vanet, in: Proceedings of the First International Workshop on Intelligent Vehicular Networks (InVeNET 2009), Macau, China, 2009.
[20] G. Yan, N. Mitton, X. Li, Reliable routing protocols in vehicular ad hoc networks, in: Proceedings of the Seventh Workshop on Wireless Ad hoc and Sensor Networks (WWASN2010), Genoa, Italy, 2010.
[21] S. Arif, I. Khalil, S. Olariu, On a versatile stochastic growth model, Int. J. Comput. Intell. Syst. 5 (3) (2012) 472–482.
[22] S. Arif, S. Olariu, Efficient solution of a stochastic SIR epidemic system, J. Supercomput. 62 (3) (2012) 1385–1403.
[23] M. Papadopouli, H. Schulzrinne, Seven degrees of separation in mobile ad hoc networks, in: IEEE Global Telecommunications Conference, 2000. IEEE GLOBECOM'00, vol. 3, 2000, pp. 1707–1711, doi:10.1109/GLOCOM.2000.891928.
[24] C. Lindemann, O.P. Waldhorst, Modeling epidemic information dissemination on mobile devices with finite buffers, in: SIGMETRICS '05: Proceedings of the 2005 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems, ACM, New York, NY, USA, 2005, pp. 121–132. http://doi.acm.org/10.1145/1064212.1064227.
[25] N. Becker, How does mass immunization affect disease incidence, Proc Workshop on Mathematical Modeling of Infectious Diseases: Dynamics and Control, Singapore, October, 2005.
[26] H.W. Hethcote, The mathematics of infectious diseases, SIAM Rev. 42 (4) (2000) 599–653.
[27] D.J. Daley, J. Gani, J.M. Gani, Epidemic Modelling: An Introduction, vol. 15, Cambridge University Press, 2001.
[28] C.W. McInnes, E. Druyts, S.S. Harvard, M. Gilbert, M.W. Tyndall, V.D. Lima, E. Wood, J.S. Montaner, R.S. Hogg, Hiv/aids in Vancouver, British Columbia: a growing epidemic, Harm Reduct. J. 6 (5) (2009).
[29] K.L. Nichol, K. Tummers, A. Hoyer-Leitzel, J. Marsh, M. Moynihan, S. McKelvey, Modeling seasonal influenza outbreak in a closed college campus: impact of pre-season vaccination, in-season vaccination and holidays/breaks, PLoS One 5 (3) (2010) e9548.
[30] A. Khelil, C. Becker, J. Tian, K. Rothermel, An epidemic model for information diffusion in manets, in: MSWiM '02: Proceedings of the 5th ACM international workshop on Modeling analysis and simulation of wireless and mobile systems, ACM, New York, NY, USA, 2002, pp. 54–60. http://doi.acm.org/10.1145/570758.570768.
[31] M. Papadopouli, H. Schulzrinne, Effects of power conservation, wireless coverage and cooperation on data dissemination among mobile devices, in: MobiHoc '01: Proceedings of the 2nd ACM International Symposium on Mobile Ad Hoc Networking & Computing, ACM, New York, NY, USA, 2001, pp. 117–127.
[32] G. Yan, S. Olariu, A probabilistic analysis of link stability in vehicular ad hoc networks, IEEE Trans. Intell. Transp. Syst. (2010).
[33] G. Yan, S. Olariu, M.C. Weigle, Providing VANET security through active position detection, Comput. Commun. 31 (12) (2008) 2883–2897.
[34] O. Diekmann, Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation, Wiley & Sons, New York, 2000.
[35] F.M. Bass, Comments on “a new product growth for model consumer durables”, Manage. Sci. 50 (12 Supplement) (2004) 1833–1840.
[36] W.O. Kermack, A.G. McKendrick, A contribution to the mathematical theory of epidemics, in: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 115, The Royal Society, 1927, pp. 700–721.
[37] N.T. Bailey, The Elements of Stochastic Processes with Applications to the Natural Sciences, vol. 25, John Wiley & Sons, 1990.
[38] M. Treiber, Microsimulation of road traffic, 2005, (http://www.traffic-simulation.de).

Xiaoyun Liu received Master's degree in Management Information System in 1994 and is currently working for Anhui University, Hefei, Anhui, China. Her research interests include information system, information systems, and e-commerce management. She authored and published many books and research papers. She also worked on several research grants from Department of Education and Department of Technology in Anhui province.

Gongjun Yan received his Ph.D. in Computer Science from Old Dominion University in 2010. He is currently an Assistant Professor in University of Southern Indiana and has been working on the issues surrounding Vehicular Ad-Hoc Networks, Sensor Networks and Wireless Communication. His main research areas include intelligent vehicles, security, privacy, and routing. In years, Dr. Yan applies mathematical analysis to model behavior of complex systems and integrates existing techniques to provide comprehensive solutions. He has published more than 40 journal papers, book chapters, and conference papers and has been the best paper award winner in international conferences (BWCCA and SOLI). He has served as grant reviewers, article/paper reviewers, session chairs and keynote speakers in international conferences and editor in journals. He is associate editor in IEEE Transactions on Intelligent Transportation System.