A Study of the VANET Connectivity by Percolation Theory

Xin Jin*, Weijie Su†, YAN Wei*
*Department of Computer Science, †School of Mathematical Sciences, Peking University
E-mail: *{jinxin, yanwei@net.pku.edu.cn}, †suweijie444@gmail.com

***Abstract*—As deploying Vehicular Ad Hoc NETworks (VANETs) costs large amounts of resources, it is crucial that governments and companies make a thorough estimation and comparison of the benefits and the costs. The network connectivity is an important factor we should take care of, because it can greatly affect the performance of VANETs and further affect how much we can benefit from VANETs. We use percolation theory to analyze the connectivity of VANETs. Through theoretical deduction, we discover the quantitative relationship among network connectivity, vehicle density and transmission range. We show that there is a jump of the network connectivity when vehicle density or transmission range is big enough. Simulations conducted in a large scenario validate our theoretical results. Our results have great meanings in the deployment of VANETs in real world. Given vehicle density, our theorem can be used to calculate the minimum transmission range to achieve good network connectivity. As a large transmission range can cause serious collisions in wireless links, it is a tradeoff to choose a proper transmission range. Our analysis can give a hand to this tradeoff and guide the deployment of VANETs in real world.**

I. Introduction

A. Motivation

Vehicular communication system is a promising technology, which can provide customers with various services from safety alert to in-car entertainment. Due to its huge application potential, it attracts attentions both from academia and industry. Many works have been done to establish the foundations for future intelligent vehicular communication application. Regardless of its progress in the research field, large-scale deployments of Vehicular Ad Hoc NETworks (VANETs) are still missing. An important reason is that both governments and companies are doubtful about the performance of VANETs in large-scale deployment in real world. Because deploying such huge facility as VANETs costs large amounts of resources, it is crucial that governments and companies make a thorough estimation and comparison of the benefits and the costs of VANETs.

Nowadays, there are still many protocols competing with each other and we do not know which one will finally become the standard to real deployment. Such situation stresses the difficulty for us to estimate the performance of VANETs. However, we can circumvent this problem by analyzing the connectivity of VANETs. No matter which network protocol is used in VANETs, the performance of the protocol is closely related to the connectivity of the network. So if we can have a good understanding of the VANET connectivity, we can have a good estimation of the performance of VANETs.

B. Previous work

In MANETs, many works have been done to analyze the network connectivity theoretically[1][3][4]. However, in VANETs, such studies are still lacking. Kafsi et al. attempt to analyze the problem both through theoretical study and simulations [8]. But their analysis is too rough and their simulation area is too small. Other works mainly use simulations to analyze the connectivity of VANETs, such as [6][7][9][10]. All these works get interesting and meaningful results. But they are all lack of a theoretical and quantitative analysis to explain the reasons behind the phenomena. Furthermore, the simulation scenarios in these works are also too small, which only cover several roads. As the application of VANETs is aiming at a whole city, a large simulation scenario is more desirable. It is doubtful that whether results got from small simulation scenarios can be applied to large scenarios in real world.

C. Our contribution

In this work, we give a theoretical analysis of the connectivity of VANETs using percolation theory. Our main contributions are:

- We discover the quantitative relationship among network connectivity, vehicle density λ and transmission range r. Based on our analysis, there is a jump for the network connectivity when vehicle density or transmission range is big enough. And we can calculate the threshold values. We conduct simulations in a large scenario, whose results accord with our analysis. Our analysis gives us some insights about the properties of the network topology of VANETs.
- We discuss the application of our theorem in deployment of VANETs in real world. A large transmission range can have good network connectivity, but it can also cause serious collisions in wireless links. So it is a tradeoff for governments and companies to choose a proper transmission range in deployment. Given vehicle density, our theorem can give the minimum transmission range to achieve good network connectivity. Below the minimum transmission range, although collisions might be few, the overall performance of VANETs can be disappointing due to the bad network connectivity. So our theorem can help us do the tradeoff.

The rest of the paper is organized as follows. In section II, we give a theoretical analysis of the VANET connectivity. In section III, we conduct simulations to validate our analysis. Section IV gives a discussion of our results in application. Finally the conclusions are drawn in section V.

II. Theoretical Analysis

A. Square bond percolation process

First of all, we assume that for each road, the incoming rate of vehicles follows Poisson distribution with parameter λ. Actually, it's not a strong assumption, as it is widely accepted in transportation engineering [13], [14]. Then we can use the parameter λ to denote the vehicle density. Secondly, every vehicle is equipped with electronic devices so that they can communicate with each other through wireless links. The transmission range is r, which means if the distance between two vehicles is no more than r, they are connected by a wireless link. Thirdly, the network should be large. As VANET is usually aiming at providing service to a whole city, the network will be large enough for us to conduct our analysis using percolation theory.

Based on the three assumptions, it is safe for us to employ a $square\ bond$ percolation process for the network. Every road segment between two intersections can be regarded as a bond. If the road segment is covered by a sequence of connected vehicles, the $bond$, which denotes the road segment, is $open$; otherwise, the bond is $closed$. We define a probability p as follows.

- p: A road segment between two intersection is covered by a sequence of connected vehicles with probability p. It also means a bond is open with probability p.

Then a bond is closed with probability $1-p$. From percolation theory, the connectivity of open bonds is closely related to p [15]. If $p < 0.5$, the network connectivity is good; if $p > 0.5$, the network connectivity is bad [15]. The state of the network connectivity changes dramatically when p jumps from below 0.5 to above 0.5 [15]. So if we can use vehicle density λ and transmission range r to calculate p, we can have an good estimation of the network connectivity.

B. The relationship among p, λ and r

Now, we begin to discover the relationship between p, λ and r. First of all, we define a term S as follows.

- S: For a certain vehicle on a road, S is the distance from this vehicle to the furthest vehicle that can be connected to it via one-hop or multi-hop wireless links.

Then using S we define a function $h(x)$ as the probability that S is larger than x. Formally,

$$h(x) = \mathbb{P}(\mathcal{S} > x). \tag{1}$$

In [1], similar definition is used to analyze the connectivity of MANETs. It also gives the precise expression of $h(x)$:

$$h(x) = \begin{cases} 1, & \text{if } 0 \leq x < r; \\ \displaystyle\sum_{i=0}^{\lfloor x/r \rfloor} \frac{(-\lambda e^{-\lambda r(x-ir)})^i}{i!} & \\ \displaystyle - e^{-\lambda r} \sum_{i=0}^{\lfloor x/r \rfloor -1} \frac{(-\lambda e^{-\lambda r(x-(i+1)r)})^i}{i!}, & \text{if } x \geq r. \end{cases} \tag{2}$$

In [1], it further shows that h satisfies the following integral equation:

$$h(x) = \lambda e^{-\lambda x} \int_{x-r}^{x} h(y) e^{\lambda y} dy, \tag{3}$$

which can be differentiated to a delay differential equation

$$h'(x) = -\lambda e^{-\lambda r} h(x-r) \tag{4}$$

for all $x > r$.

So, if we can find the relationship between $h(x)$ and p, we can use λ and r to calculate p. We guess that $h(x)$ is the same order with p^x when x grows to infinity. More precisely, we expect that there exists a positive number c such that

$$h(x) = (c + o(1))p^x, \tag{5}$$

which is equivalent to

$$\lim_{x \to +\infty} \frac{h(x)}{p^x} = c. \tag{6}$$

1) Proof of the relationship between p and $h(x)$: In order to prove Equation (5), we should firstly demonstrate that

$$\lim_{x \to +\infty} h(x)^{\frac{1}{x}} \tag{7}$$

exists. We need the following lemma.

Lemma 1. *For any $x, y \geq 0$, we have*

$$h(r + x + y) \geq h(r + x)h(r + y). \tag{8}$$

Proof: Let $A_z = \{\mathcal{S} \geq z\}$. By definition we have $\mathbb{P}(A_z) = h(z)$. It is clearly that $A_{r+x+y} \subset A_{r+x}$. Given A_{r+x}, we denote by w the distance from the beginning to the furthest vehicle whose distance from the beginning is no more than $x + r$. It is easy to see $x < w \leq x + r$. Therefore, conditioned on A_{r+x}, a random length between y and $y + r$ should be covered. Since h is a decreasing function, using the *Strong Markov Property* [2] of Poisson process we have

$$\mathbb{P}(A_{r+x+y}|A_{r+x}) \geq \mathbb{P}(A_{r+y}). \tag{9}$$

Therefore, we have

$$\begin{aligned} \mathbb{P}(A_{r+x+y}) &= \mathbb{P}(A_{r+x+y} \cap A_{r+x}) \\ &= \mathbb{P}(A_{r+x})\mathbb{P}(A_{r+x+y}|A_{r+x}) \\ &\geq \mathbb{P}(A_{r+x})\mathbb{P}(A_{r+y}). \end{aligned} \tag{10}$$

■

Now using Lemma 1, we can prove the existence of (7).

Theorem 2. $\lim_{x\to+\infty} h(x)^{\frac{1}{x}}$ *exists and lies in* $(0,1)$.

Proof: Firstly, we define a function:

$$g(x) = \log h(r+x). \tag{11}$$

Based on Lemma 1, we obtain

$$g(x+y) \geq g(x) + g(y). \tag{12}$$

For any positive numbers $x < y$, we can write y as $y = kx + z$, where $k = \lfloor y/x \rfloor$ and $0 \leq z < x$. By iteratively using Inequality (12), we have

$$g(y) \geq kg(x) + g(z), \tag{13}$$

which can be deduced to

$$\frac{g(y)}{y} \geq \frac{kx}{kx+z}\frac{g(x)}{x} + \frac{g(z)}{y}. \tag{14}$$

Since $g(z)$ is bounded, by letting $y \to +\infty$ we obtain

$$\liminf_{y\to+\infty} \frac{g(y)}{y} \geq \frac{g(x)}{x}, \tag{15}$$

for arbitrary x. Then we let $x \to +\infty$ and get

$$\liminf_{y\to+\infty} \frac{g(y)}{y} \geq \limsup_{x\to+\infty} \frac{g(x)}{x}. \tag{16}$$

As a consequence,

$$\lim_{x\to+\infty} \frac{g(x)}{x}$$

exists. We denote it by ω. Recall the definition of $g(x)$ in Equation (11). We have

$$\lim_{x\to+\infty} h(x)^{\frac{1}{x}} = \lim_{x\to+\infty} e^{\frac{g(x-r)}{x}} = e^{\omega}. \tag{17}$$

It is easily to see that

$$\omega \geq \frac{g(x)}{x} > -\infty \tag{18}$$

for each positive number x.

On the other hand, [1] gives a loose upper bound of h:

$$h(x) \leq (1-e^{-\lambda r})e^{-\lambda(x-r)e^{-\lambda r}}. \tag{19}$$

By using a limit argument we obtain $\omega \leq -\lambda e^{-\lambda r} < 0$. Therefore we have

$$\omega \in (-\infty, 0). \tag{20}$$

Thus we have proven the existence of $\lim_{x\to+\infty} h(x)^{\frac{1}{x}} = e^{\omega}$ and $\lim_{x\to+\infty} h(x)^{\frac{1}{x}} \in (0,1)$. ■

Now we begin to prove Equation (5). We need the following theorem from [16].

Theorem 3. *Each solution h of the delay differential equation (4), provided that*

$$\lim_{x\to+\infty} h(x)^{\frac{1}{x}} = p \tag{21}$$

exists and positive, has the asymptotic form:

$$h(x) = (c + o(1))p^x, \tag{22}$$

where c is a positive constant.

Remark 4. *This theorem is trivial at first glance. But actually it is not. For example, $h(x) = \frac{p^x}{x}$ also enjoys Equation (21), but does not have the property the theorem asserts.*

Theorem 5.

$$h(x) = (c + o(1))p^x. \tag{23}$$

Proof: Combining with Theorem 2 and Theorem 3 gives the desired result. ■

Our next step is to infer the relationship among p, λ and r from Equation (5).

2) Determination of p: To find the exact value of p, we need the following lemma.

Lemma 6. *Given $r, \lambda > 0$, when $r\lambda \neq 1$, the following equation has exactly two roots in the interval $(0,1)$:*

$$x^r \log x + \lambda e^{-\lambda r} = 0. \tag{24}$$

When $\lambda r = 1$, this equation has exactly one root in $(0,1)$.

Proof: Denote $f(x)$ by $f(x) = x\log x + \lambda r e^{-\lambda r}$. It is easy to see that x is a root of f if and only if x is a solution of Equation (24). When $x \to 0$ or $x \to 1$, f will tend to $\to \lambda r e^{-\lambda r} > 0$. Observe that

$$f''(x) = \frac{1}{x} > 0,$$

which implies that f is strictly convex on $(0,1)$. Let the derivative of f be zero, then we find f reaches minimum at $p = e^{-1}$:

$$f(e^{-1}) = -e^{-1} + \lambda r e^{-\lambda r} \leq 0,$$

which will be a strict inequality when $\lambda r \neq 1$. By the continuum and convexity properties of f we can conclude that f has one root in $(0, e^{-1})$ and the other root in $(e^{-1}, 1)$ respectively, when $\lambda r \neq 1$. When $\lambda r = 1$, e^{-1} is the only root of f. ■

Remark 7. *It is easy to see that Equation (24) has a root $e^{-\lambda}$. We denote by p the other root when $\lambda r \neq 1$. When $\lambda r = 1$, $e^{-\lambda}$ is the only root.*

Now we present our main theorem.

Theorem 8.

$$p = \lim_{x\to+\infty} h(x)^{\frac{1}{x}}$$

is a root of Equation (24). Furthermore, we have

$$\begin{cases} p > e^{-\frac{1}{r}}, & \text{if } \lambda r > 1 \\ p = e^{-\lambda}, & \text{if } \lambda r = 1 \\ p < e^{-\frac{1}{r}}, & \text{if } \lambda r < 1 \end{cases}. \tag{25}$$

Proof: From Equation (3) and Equation (22), we find that

$$(c+o(1))(pe^{\lambda})^x = \lambda(c+o(1))\int_{x-r}^{x} (pe^{\lambda})^y dy, \tag{26}$$

which leads to

$$\begin{cases} p^r \log p + \lambda e^{-\lambda r} = 0, p \neq e^{-\lambda}, & \text{if } \lambda r \neq 1 \\ p = e^{-\lambda}, & \text{if } \lambda r = 1 \end{cases} \tag{27}$$

by letting $x \to +\infty$. ■

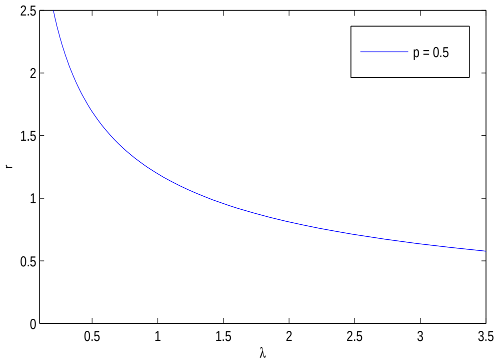


Fig. 1. r vs. λ when $p = 0.5$. r is unified by the distance between two intersections.

C. Theoretical discussion

From Theorem 8, we conclude that p is a function of r, λ:

$$p = p(r, \lambda). \tag{28}$$

Though this is an implicit function, it is easy to calculate the value of p by many methods, such as Newton-Raphson method. Intuitively, p is an increasing function of both r and λ, which can be proved rigorously. Figure 1 shows all pairs of (λ, r) that satisfies $p = 0.5$. All the theoretical analysis, as well as Figure 1, ensures us for each λ, there exists a minimum transmission range r_0 such that for any $r > r_0$, p will be larger than 0.5, and vice versa. For each r, similar conclusion can be easily obtained.

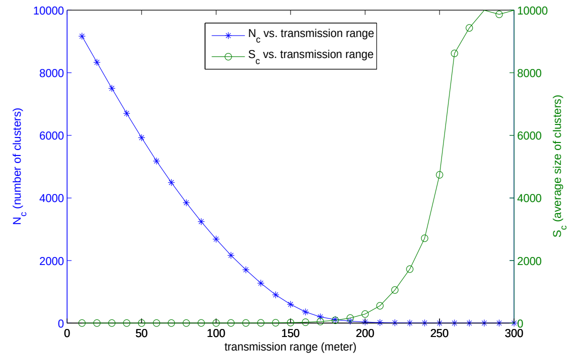
III. Simulation

A. Simulation settings and metrics

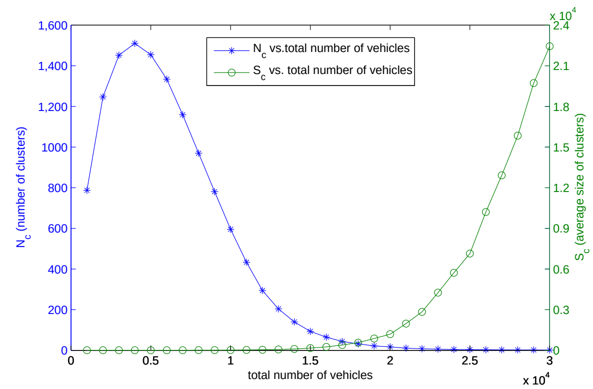
We conduct simulations in a large scenario. The area has 50 horizontal roads and 50 vertical roads. The distance between two parallel roads is 250 $meters$. The simulation area is about 150 $square\ kilometers$ large, nearly a small city. The mobility model we use is similar to Manhattan Mobility Model [5]. The max velocity is 30 m/s. For each second, each vehicle chooses a random velocity smaller than the max velocity. Vehicles can choose to turn left, right or go straight at intersections. The choice is probabilistic: the probability of going straight is 0.5 and the probability of turning left and right is both 0.25. The simulation time is 300 $seconds$. We use N_c to measure network connectivity, which is defined as follows.

- N_c: number of clusters. A cluster consists of a group of vehicles in which every vehicle is connected to other vehicles by single-hop or multi-hop wireless links. Then N_c is the number of different clusters in the network.
- S_c: average size of clusters. The size of a cluster is the number of vehicles in the cluster. And S_c is the average size of all the clusters in the network.

In each simulation, we calculate number of clusters and average size of clusters in every second and use the average of them as N_c and S_c for the simulation.



(a) Network connectivity vs. transmission range r. The total number of vehicles is fixed to 10000.



(b) Network connectivity vs. total number of vehicles. The transmission range r is fixed to 150 m.

Fig. 2. Simulation Results

B. The impact of transmission range

In the first group of simulations, we fix the vehicle density λ and change the transmission range r. The total number of vehicles in the simulation area is fixed to 10000, which means λ is about 2 $vehicles/lane$. r varies from 10 $meters$ to 300 $meters$. The results are shown in Figure 2(a). We can easily see that when r increases, N_c decreases and S_c increases. It means an increase of transmission range can result in better network connectivity. Furthermore, we observe that when r is about 200 $meters$, there is a jump both for N_c and S_c. From Theorem 8, we can calculate p is about 0.5 when r is 200 $meters$ and λ is 2 $vehicles/lane$. So our simulation result here is consistent with our analysis.

C. The impact of vehicle density

In the second group of simulations, we fix the transmission range r and change the vehicle density λ. r is fixed to 150 $meters$. The total number of vehicles varies from 1000 to 30000, which means λ varies from 0.20 $vehicles/lane$ to 4.08 $vehicles/lane$. The results are shown in Figure 2(b). When the total number of vehicles increases, number of clusters N_c first increases and then decreases. This is because when the vehicle is few, the network connectivity is not so good. More vehicles only result in more isolated clusters.

However, when the number of vehicles increases to some degree, the small isolated clusters aggregate to become big clusters. So N_c begins to decrease. For S_c, it increases along with the total number of vehicles. Moreover, when the total number of vehicles increases to about 16000, there is a jump for both N_c and S_c, after which the network connectivity becomes so good. When the total number of vehicles is 16000, the vehicle density is about $3.26\ vehicles/lane$. The probability p is about 0.5 when r is $150\ meters$ and λ is $3.26\ vehicles/lane$ based on Theorem 8. The simulations again accord with our theoretical analysis that when p is 0.5, there will be a jump for the network connectivity.

IV. Application in Real World

Our results have great meanings in real world, which can guide the deployment of VANETs.

If there are not many vehicles in a city, which means the vehicle density λ is low, the transmission range r should be large in order to get a good network connectivity. So when deploying VANET in the city, government and companies should be aware of such situation and choose electronic devices with large power. We can calculate the minimum transmission range to achieve good network connectivity using Theorem 8. On the other hand, we know that a large transmission range can cause serious collisions in wireless links, which can reduce the performance of wireless networks. So it's a tradeoff to choose a smaller transmission range with worse network connectivity but fewer collisions and a larger transmission range with better network connectivity but more collisions. At least, our theory tells us that there is a minimum transmission range, below which the overall performance of the network might be disappointing due to the bad network connectivity. Further discussion about the tradeoff will be a part of our future wok.

For cities with large amounts of vehicles, our theory states that even a small transmission range is enough to obtain good network connectivity. The exact value of the minimum transmission range can be calculated by Theorem 8. Devices with a small transmission range need less power and save energy, which also meets the requirement of "GREEN EARTH". However, there is also a pitfall here. Large amounts of vehicles do not mean large amounts of vehicles equipped with these electronic devices. In our analysis, we assume that all vehicles are equipped with the devices and use vehicle density to denote them. But it is natural that for a new technology, there will be a long process until reaching a high market penetration rate. So although a city might have a large quantity of vehicles, the exactly vehicle density equipped with the electronic devices can still be very low. In such situation, a large transmission range is still desirable in order to get good network connectivity.

Another thing we should take into account is that even in the same city, the vehicle density is different at different time. Usually the vehicles running on the roads in the daytime are much more than those at night. So the minimum transmission range varies during the day. If the transmission range is fixed to provide good network connectivity all the day, it will be too large for the daytime, which not only wastes energy but also brings more collisions. In such situation, the software radio technology can be used to adjust the transmission range according to different vehicle density at different time.

V. Conclusion

In this paper, we studied the connectivity of VANETs both using theoretical analysis and simulations. Firstly, we use percolation theory to analyze the problem theoretically. We find the quantitative relationship between network connectivity, vehicle density λ and transmission range r. Furthermore, we conduct simulations in a large scale scenario to validate our theorem, whose results accord with our analysis well. Our results have great meanings in real world. Given vehicle density, we can calculate the minimum transmission range to achieve good network connectivity, which can guide the deployment of VANETs.

In the future, we will take the impact of RSUs into consideration. Furthermore, we will study the impact of transmission range and vehicle density on collisions in wireless links. Given vehicle density, we want to give a proper choice of transmission range to obtain a good tradeoff of network connectivity and collision, which will greatly help governments and companies in planning and deploying of VANETs.

References

[1] Olivier Dousse, Patrick Thiran and Martin Hasler, *Connectivity in Ad-Hoc and Hybrid Networks*. IEEE INFOCOM, 2002.

[2] Richard Durrett, *Probability: Theory and Examples (2 edition)*. Duxbury Press, 1995.

[3] P.Gupta and P.R.Kumar, *The Capacity of Wireless Networks*. IEEE Transactions on Information Theory, IT-46(2), pp.388-404, March 2000.

[4] P.Gupta and P.R.Kumar, *Critical Power for Asymptotic Connectivity in Wireless Networks*, Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H.Fleming, pp.388-404, 1998.

[5] Fan Bai, N. Sadagopan and A. Helmy, *IMPORTANT: A framework to systematically analyze the impact of mobility on performance of routing protocols for adhoc networks*. IEEE INFOCOM, 2003.

[6] I. W. H. Ho, K. K. Leung, J. W. Polak, Rahul Mangharam, *Node Connectivity in Vehicular Ad Hoc Networks with Structured Mobility*. IEEE LCN, pp.635-642, 2007.

[7] Marco Fiore and Jerome Harri, *The networking shape of vehicular mobility*. ACM MobiHoc, 2008.

[8] M. Kafsi, P. Papadimitratos, O. Dousse, T. Alpcan and J. P. Hubaux, *VANET connectivity analysis*. IEEE Autonet, December 2008.

[9] M. Artimy, W. Phillips and W. Robertson, *Connectivity in inter-vehicle ad hoc networks*. Engineering Canadian Conference on Electrical and Computer, 2004.

[10] M. Artimy, W. Phillips and W. Robertson, *Connectivity with static transmission range in vehicular ad hoc netwroks*. CNSR, 2005.

[11] Michel Ferreira, Hugo Conceicao and Ozan K. Tonguz, *Stereoscopic aerial photography: an alternative to model-based urban mobility approaches*. ACM VANET, 2009.

[12] Natarajan Meghanathan, *A simulation study on the impact of mobility models on the network connectivity, hop count and lifetime of routes for ad hoc networks*. Informatica vol(34), pp.207-221, 2010.

[13] A. S. Narasimha Murthy and Henry R. Mohle, *Transportation Engineering Basics*. American Society of Civil Engineers, 2000.

[14] Roger P. Roess and Elena S. Prassas and William R. McShane, *Traffic Engineering (3rd Edition)*. Pretice Hall, 2004.

[15] Geoffrey Grimmett, *Percolation(1 edition)*. Springer-Verlag, 1989.

[16] Richard Bellman and Kenneth Cooke, *Differential-difference equations*. Academic Press, 1963.