Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2016, Article ID 5159739, 10 pages
http://dx.doi.org/10.1155/2016/5159739

Research Article

IA^{2}P: Intrusion-Tolerant Malicious Data Injection Attack Analysis and Processing in Traffic Flow Data Collection Based on VANETs

Nan Ding, Guozhen Tan, and Wei Zhang

School of Computer Science and Technology, Dalian University of Technology, Dalian 116023, China

Correspondence should be addressed to Nan Ding; dingnan@dlut.edu.cn

Received 13 August 2015; Revised 4 January 2016; Accepted 28 April 2016

Academic Editor: Yu Wang

Several studies investigating data validity and security against malicious data injection attacks in vehicular ad hoc networks (VANETs) have focused on trust establishment based on cryptology. However, the current researching suffers from two problems: (P1) it is difficult to distinguish an authorized attacker from other participators; (P2) the large scale of the system and high mobility set up an obstacle in key distribution with a security-based approach. In this paper, we develop a data-centric trust mechanism based on traffic flow theory expanding the notion of trust from intrusion-rejecting to intrusion-tolerant. First, we use catastrophe theory to describe traffic flow according to noncontinuous, catastrophic characteristics. Next, we propose an intrusion-tolerant security algorithm to protect traffic flow data collection in VANETs from malicious data injection attacks, that is, IA^{2}P, without any security codes or authentication. Finally, we simulate two kinds of malicious data injection attack scenarios and evaluate IA^{2}P based on real traffic flow data from Zhongshan Road in Dalian, China, over 24 hours. Evaluation results show that our method can achieve a 94% recognition rate in the majority of cases.

1. Introduction

VANETs are emerging as an effective new tool to monitor the physical world [1]. They gather traffic flow data (GPS, speed measurements, etc.) from sensor platforms in vehicles and relay these data via vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication. With advances in wireless communication and sensing, VANETs can be used to solve installation and maintenance problems caused by traditional traffic monitoring infrastructure, such as loop detectors, cameras, and radar. Therefore, more and more studies have suggested expanding traditional traffic monitoring infrastructure to gather the traffic flow data with VANETs in Intelligent Transportation Systems (ITS).

VANETs are data-based networks [2], in which data quality and security are paramount. In VANETs scenarios, each participating vehicle or fixed roadside infrastructure is transformed into a wireless message transmitter or sensing terminal. Some studies indicate that the security of data could seriously influence the performance of VANETs in practice [3, 4]. In these security attacks, the malicious data injection attack can harm intelligent traffic systems. By injection attack, the malicious data are injected into VANETs and disrupt ITS applications. For example, sending false traffic flow data to emulate traffic jams or accidents may disrupt traffic signal control systems. This could increase accidents, compromising safety.

Conventional approaches against injection attacks are apt to adopt the traditional notions of trust. A variety of research contributions are based on designing cryptographic solutions to offer both Trusted Authority (TA) and Message Confidentiality (MC) for VANETs' applications. To use a cipher for TA or MC, every participator (vehicular or fixed roadside infrastructure) requires some kind of a shared secret, providing various methods of secret key distribution [5, 6]. However, these researches are suffering from the following problems: (1) it is difficult to distinguish an authorized attacker from other participators and (2) the large scale of the system and high mobility set up an obstacle in key distribution with a security-based approach.

Some studies have addressed data-based security mechanisms against the malicious data injection attacks in new ways [7, 8], which is more efficient in the fields of data-based VANETs applications, such as traffic congestion detection and traffic route guidance. The data-based security mechanisms are more resilient to attacks, coming quickly to the correct decision. However, these studies focus mainly on establishing frameworks for data-centric trust rather than linking data characteristics. It is well known that the characteristics of traffic flow data are distinct and regular. Regardless of data characteristics, data-based trust mechanisms are insufficient and impractical against injection attacks.

In this paper, we used traffic flow characteristics to develop an intrusion-tolerant security mechanism to protect traffic flow data collection in VANETs against injection attacks. This security mechanism can be applied in most data-based VANET scenarios. Our study is innovative because we

(1) develop an intrusion-tolerant security mechanism against injection attacks without security codes or authentication, IA^{2}P, and this extends the notion of security from intrusion-rejecting to intrusion-tolerant, and, therefore, this approach is more useful in practice than traditional trust establishment based on cryptology;

(2) expand cusp catastrophe theory to analyze traffic flow data profiling and this is more suitable for traffic flow data characteristics in most traffic scenarios, allowing for effective analysis of injection attacker's activities;

(3) integrate batch estimation filters with coefficient self-adjustment to meet traffic flow time-varying volatility in order to generalize injection attack analysis and processing.

The rest of this paper is organized as follows. Section 2 presents the basic principles of malicious data injection attacks and the model of traffic flow data based on catastrophe theory. Section 3 proposes the IA^{2}P mechanism for the injection attack analysis and processing and then improves it with the batch estimation filter for generalization. Section 4 demonstrates the performance of IA^{2}P through simulation. Section 5 focuses on related work.

2. Related Works and Problem Statement

Security studies have produced rich literature in VANETs. As with other applications in DSRC, mobile ad hoc network (MANET) and Peer-to-Peer (P2P), notions of security in VANET are mainly to build trust mechanisms against injection attacks.

Most state-of-the-art studies have focused on designing cryptographic solutions to offer both Trusted Authority and Message Confidentiality and thus protect VANET applications against malicious data injection attacks. These approaches have mainly considered two cases: certification and routing. For example, Lu et al. [6] proposed Trusted Authority with the authenticated recognition to each vehicle in VANETs. Sun et al. [9] proposed an identity-based security system by cryptography to VANETs. Wasef et al. [10] and Schoch et al. [11] proposed a scheme to complement the public key infrastructure to secure VANETs.

Meanwhile, Raya et al. [12] proposed data-centric security mechanisms for data-based trust establishment in ad hoc networks. They concluded that data-based trust mechanisms are more simple and practical than cryptographic trust mechanisms. Furthermore Aslam et al. [13] presented two approaches for reliable traffic information propagation: two-directional data verification and time-based data verification. With these two types of verification, traffic messages are sent through two (spatially or temporally spaced) channels. The recipient verifies message integrity by checking whether data received from both channels match.

In VNETs, Wu et al. [7] proposed a Roadside-Unit Aided Trust Establishment (RATE) scheme to execute data-centric trust establishment. And Mazilu et al. [8] designed a data-trust security model designed for VANETs, based on social network theories, to compute a trust index for each message according to the relevance of the event, such as traffic congestion and safety warnings. Based on them, Sha et al. [14] proposed RD4, a data-detection and filtering mechanism, to detect false data in VANETs. They focused on false data generated from the unreliable components and untrustworthy data sources.

For literature in other fields, Liu et al. [15] proposed a theoretical model based on data characters to analyze false data injection attacks in the field of electric power state estimation. Roy et al. [16] present a verification algorithm to determine whether the aggregate includes any false data, which are used in wireless senor networks (WSNs).

In conclusion, our study found that the data-based security mechanism we developed is proficient in identifying false data generated by injection attacks. According to tests and simulations of [15, 16], protection against malicious data injection attacks is more efficient if the characteristics and disciplines of the data are considered. However, data-based security mechanisms considering the characteristics and disciplines of traffic flow data in VANETs against the injection attack were not detected.

2.1. Malicious Data Injection Attacks. VANETs are complex systems connecting vehicle-to-vehicle and vehicle-to-roadside infrastructures through transmission and distribution networks across local geographical area. As long as they have legal authority, the malicious vehicles can send messages or data to other vehicles, whether these are unreal or illegal. They can also modify other legal messages or data as relay nodes receiving and transferring from their neighbor nodes [17].

For example, as shown in Figure 1, a vehicle A sends a "Road clear" message to a malicious vehicle B (attacker) and B alters the message as "Traffic jam ahead" and sends it to a legitimate vehicle C. C transfers it to vehicle D. C and D will be affected by this message since they will change the road and be in trouble later on.

Unfortunately, most existing trust mechanisms cannot identify these attacker's illegal activities. Since they are authorized, attackers are often able to bypass safeguards.

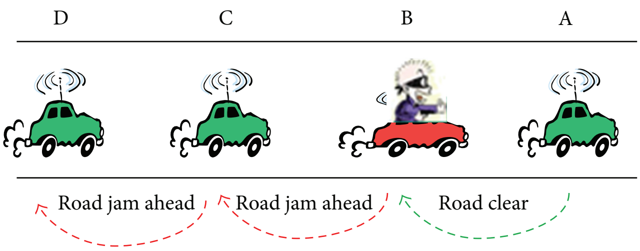


Figure 1: The scenario of VANETs with attackers.

There are two types of data injection attacks: random false data injection attacks and targeted false data injection attacks [19]. The aim of a random false data injection attack is to find any attack vector that can result in a wrong estimation of state variables. The aim of a targeted false data injection attack is to find an attack vector to inject a specific error into certain monitoring variables.

The attacker chooses any nonzero arbitrary vector as the attack target and then constructs malicious measurements. However, a traditional bad measurement detection approach cannot detect them. For example, an approach based on a 2-norm of the measurement residual is bypassed because the data appear to be valid. Fortunately, these data do not accord with traffic flow features, especially when they are analyzed with the multidimensional model.

Following are definitions based on descriptions of injection attacks [3, 4].

Definition 1 (malicious data). Malicious data is invalid data injected by an injection attacker. It can be divided into two categories: multirepeat data and fake data. (a) Multirepeat data (MRD) is copied directly from valid data and injected regularly into VANETs. Although appearing to be valid, the values of this data are fixed and constant. (b) Fake data (FD) is falsified from valid data or randomly generated. It is unfixed and variable outside of traffic flow laws.

Definition 2 (VANETs Participant ID (VP ID)). A VANETs Participant ID (VP ID) is a unique identification of each participant in VANETs and does not require special authorization. It can use a MAC or IP address. VANETs Participants include vehicles or Roadside Units, which compose VANETs and exchange messages. In this paper, one selected MAC as the VP ID.

Definition 3 (Injection Attacker List (O_{list})). An Injection Attacker List is a record of the attacker's VP ID, which is stored in each VANETs Participant, meaning that the participant holding this VP ID is an injection attacker and has sent malicious data to the O_{list} owner before. O_{list} of each VANETs Participant may be different.

2.2. Problem Formulation of Malicious Data Injection Attack Based on Catastrophe Theory. Nonlinearity is an inherent property of the traffic flow [20]. Gazis et al. improved nonlinear follow-the-leader models to describe the traffic flow in 1961 [21], which attracted the researching attention from then on. With the rapid development of information technology, more and more traffic flow data are collected by installing sensors (usually double induction loop detectors) along the road that measure flux and speed at a certain location. The nonlinearity of traffic flow has been proven, and more nonlinear theory and model, such as the fluid-dynamical model [22], are improved to describe the traffic flow.

Catastrophe theory is used to explain the natural and social phenomena that occur in the process of discontinuous changes and analyze the noncontinuous characteristics near the critical point. Navin [23] proposed a cusp catastrophe traffic model to explain sudden changes in traffic flow. Hall and others later demonstrated that traffic flow fits the cusp catastrophe surface [24–27]. According to basic cusp catastrophe theory, the total potential energy function of traffic flow ($W(v)$) is as follows:

$$W(v) = av^4 + bqv^2 + cpv. \tag{1}$$

Here, v is the vehicle speed, representing the state variable of W. As the control variables of W, q and p are traffic volume and occupancy, respectively. Parameters a, b, and c are coefficients. In our algorithm, these coefficients will be given which will be described in Section 3.

Based on (1), the manifold function and the bifurcate equation of cusp catastrophe are

$$\begin{gathered} 4av^3 + 2bqv + cp = 0, \\ 8b^3q^3 + 27ac^2p^2 = 0. \end{gathered} \tag{2}$$

Based on (2), the relationship of v, p, and q is developed. Let x represent the original measurements collected from VANETs, where $x = (v, q, p)$. To describe these measurements and represent their relationships, we define the Catastrophe Vector.

Definition 4 (Catastrophe Vector). The Catastrophe Vector (CV) is used to describe the traffic flow measurement with the cusp catastrophe model. The CV of measurement x is as follows, where $m = b/a$, $n = c/a$, and a, b, and c are the coefficients of (2):

$$\text{CV}(x) = \left(4v^3 + 2mqv + np, 8m^3q^3 + 27n^2p^2\right). \tag{3}$$

Based on CV, an injection detection model of traffic flow data can be proposed as

$$f(x) = h\text{CV}(x). \tag{4}$$

Here, h is the coefficient, whose value is suggested in (1–1.05) since the error tolerance limit of traffic flow data is ±5%, according to [18]. And this error tolerance limit is still used in some popular traffic signal control system, such as SCOOT and SCATS. As we know, the analysis method of the traffic flow data validity in these systems is to detect whether the change in the adjacent data from the same source is within the threshold range, which is similar in [18]. So, these are the reasons that we adopt ±5% as the threshold of IA^{2}P.

2.3. Evaluation Function of Malicious Data Injection Attacks. Malicious data z, $z = (v_z, q_z, p_z)$, is used for injection into x. Let x_z be the vector of observed measurements, where z has been injected into x.

According to the model of (4), each observed measurement x_z can generate $f(x)$. Then $f(x)$ can be projected at the two-axis Cartesian coordinates and regarded as a vector. Therefore, the evaluation function of injection attack is defined:

$$\begin{aligned} H(x_z) &= \|f(x_z)\| = h\|\mathrm{CV}_{xz}\| \\ &= h\sqrt{(4v^3 + 2mqv + np)^2 + (8m^3q^3 + 27n^2p^2)^2}. \end{aligned} \tag{5}$$

Based on this evaluation function, the conclusions of malicious data injection attack are as follows.

Conclusion 1. The measurement x_z is clean without injection attack, when

$$H(x_z) \le \varepsilon. \tag{6}$$

Conclusion 2. x_z is false data added by injection attacks, when

$$H(x_z) > \varepsilon. \tag{7}$$

Here, ε is the threshold of injection detection. The value of ε is given based on fluctuation of true traffic flow data x:

$$\varepsilon = \max\left(h\|v - \overline{v}\|, h\|q - \overline{q}\|, h\|p - \overline{p}\|\right). \tag{8}$$

Here, h is the coefficient in (4), and $\overline{v}$, $\overline{q}$, and $\overline{p}$ are the effective value according to the history data of the valid measurement x. For example, they could be expressed by the mean value of the valid measurements.

3. Malicious Data Injection Attack Analysis and Processing

In this section, we proposed a new malicious data injection attack analysis and processing algorithm, IA^{2}P. Firstly, we introduced how to self-adjust coefficients in the cusp catastrophe model based on batch estimation filter to make IA^{2}P more practical in most traffic scenarios. Based on it, we described the theory and procedure of the IA^{2}P algorithm.

3.1. Coefficients Self-Adaption Based on Batch Estimation Filter. The traffic flow model based on the catastrophe theory can describe the character of traffic flow. However, the traffic flow characteristics for each traffic scenario differ since roadbed construction, transportation infrastructure, and traffic signal patterns are distinct. These can influence variation in traffic flow model coefficients. In fact, adjusting the parameters manually for each traffic scenario does not work in this case. This uncertainty affects injection attack analysis and detection in VANETs.

To solve this problem and generalize injection attack detection, the batch estimation filter can be adopted to actively learn the coefficients of (4) online. According to the valid measurements judged by (4), m and n are calculated with each CV. Equation (9) is given:

$$\begin{aligned} \varepsilon_{i-1} &= \sqrt{{m_{i-1}}^2 - m'^2}; \\ \varepsilon_{i-2} &= \sqrt{{m_{i-2}}^2 - m'^2}; \\ m_i &= \frac{{\varepsilon_{i-2}}^2}{{\varepsilon_{i-1}}^2 + {\varepsilon_{i-2}}^2} m_{i-1} + \frac{{\varepsilon_{i-1}}^2}{{\varepsilon_{i-1}}^2 + {\varepsilon_{i-2}}^2} m_{i-2}; \\ {\varepsilon'}_{i-1} &= \sqrt{{n_{i-1}}^2 - n'^2}; \\ {\varepsilon'}_{i-2} &= \sqrt{{n_{i-2}}^2 - n'^2}; \\ n_i &= \frac{{{\varepsilon'}_{i-2}}^2}{{{\varepsilon'}_{i-1}}^2 + {{\varepsilon'}_{i-2}}^2} n_{i-1} + \frac{{{\varepsilon'}_{i-1}}^2}{{{\varepsilon'}_{i-1}}^2 + {{\varepsilon'}_{i-2}}^2} n_{i-2}. \end{aligned} \tag{9}$$

Here m_{i-1} and n_{i-1} and m_{i-2} and n_{i-2} are the parameters of $\mathrm{CV}(x_{i-1})$ and $\mathrm{CV}(x_{i-2})$, respectively. m' and n' are the parameters of $\mathrm{CV}(x_i) = 0$, if the measurement x_i provides good data. Based on (5), let $m = m_i$, which can be adapted based on m_{i-1} and m_{i-2}, and $n = n_i$, which can be adapted based on n_{i-1} and n_{i-2}.

It is notable that m' and n' may be more than one, so that results are calculated in $\mathrm{CV}(x_i) = 0$ based on x_i. The one, which deviates to m_i and n_i and is the least, should be selected and used in (9).

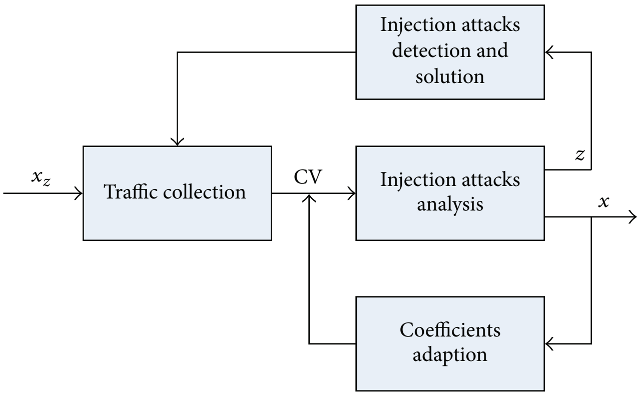

Injection attacks detection and solution
x_z
Traffic collection
CV
Injection attacks analysis
z
x
Coefficients adaption

Figure 2: Block diagram of IA^{2}P.

3.2. Procedures for Malicious Data Injection Attack Analysis and Processing. Based on the evaluation function of malicious data injection attack and the coefficient self-adaption, we propose a generalized method of the injection attack analysis and processing algorithm, that is, IA^{2}P, as shown in Figure 2.

The algorithm is comprised of 4 parts: traffic collection, injection attack analysis, coefficient adaption, and injection attack processing.

(a) Each observed measurement $x_z(v, q, p)$ is collected in traffic collection. When x_z is collected it is first checked for the first type of malicious data, MRD, through comparison with adjacent data from the same vehicle. If it is fixed and constant, it is discarded as MRD. Then the other types of data are transformed to CV(x_z) and sent to injection attack analysis.

(b) In injection attack analysis, CV(x_z) is evaluated by the evaluation function (5). Malicious data can be detected, which may be either MRD or FD.

(c) x can be sent to coefficient adaption for adaption to the model's coefficients to fit the variation in local traffic flow.

(d) Malicious data are then sent for injection attack processing so that, for example, the traffic collection portion can add the attacker's ID to O_{list}, preventing further attacks.

Furthermore, we propose a state machine for IA^{2}P, as shown in Figure 3.

S0: the Station of Initialization mainly sets the coefficients of the model (h_0, m_0, n_0, ε, and so on) and then goes to S1.

S1: the Station of CV Transformation checks whether a measurement, $x_z(v, q, p)$, is MRD upon collection. If it is fixed and constant, it is discarded as MRD. Otherwise, it is transformed into CV(x_z). Then CV(x_z) is sent to S2 if ID_{x_z}(MAC) $\notin O_{\text{list}}$, meaning that x_z's sender is a valid participant.

S2: the Station of Injection Attack Analysis is based on evaluation function (5): if $H(x_a) > \varepsilon$, x_z is malicious data. This indicates that an injection attack is occurring, and x_z is sent to S4. However, if $H(x_a) \le \varepsilon$, x_z is safe and valid data. As x_z is the output of IA^{2}P, it is sent to S3.

S3: for the Station of System Update, the coefficients of the model are self-adapted based on (9) to retain the traffic flow pattern's variation. Then the station machine goes on to S1 to continue the next measurement transformation and injection attack analysis.

S4: in the Station of Injection Attack Processing, the injection attack is recognized. The measurement x_z is isolated, and the sender's ID is sent back to S1. The Injection Attacker List, O_{list}, is updated and then the station machine goes back to S1.

Considering the characteristics of traffic flow monitored in VANETs, the state machine should run a long time and process continuously. Therefore, the state machine is not arranged for the end state. In actual operations, the system must be stopped and restarted manually.

4. Simulation and Performance Analysis

4.1. Experiment Setup. In this section, we validate the malicious data injection attack analysis and processing through experiments using actual traffic data sets provided by the Dalian Department of Transportation. These data sets were archived from traffic flow data collected by inductive loop detectors on the Zhongshan Road in Dalian, China. Table 1 shows the details of these detectors. The archives contain traffic volume, speed, and occupancy measurements from 12:00 a.m. to 12:00 p.m. on May 7, 2010.

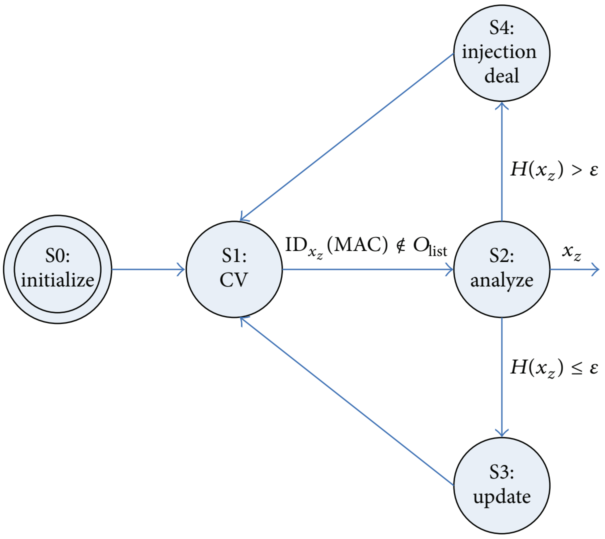
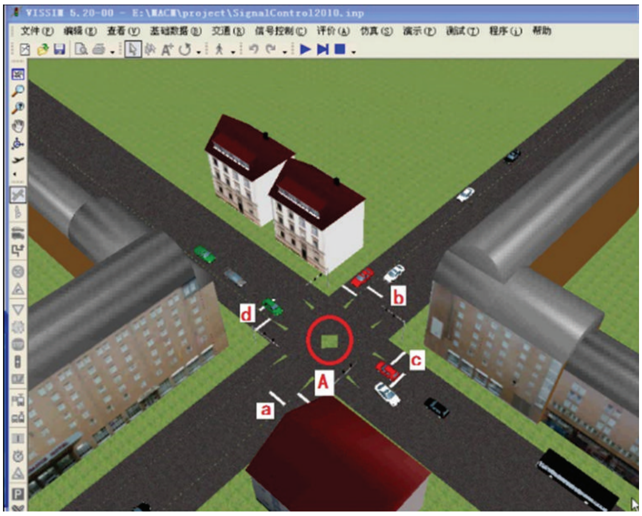


FIGURE 3: State machine for IA^{2}P.


VISSIM 5.20-00 - E:\MACM\project\SignalControl2010.inp
文件(F) 编辑(E) 查看(V) 基础数据(B) 交通(R) 信号控制(C) 评价(A) 仿真(S) 演示(P) 测试(T) 程序(I) 帮助
a
b
c
d
A

⧅ Loop sensor detector

🟩 Roadside Unit

FIGURE 4: Traffic scene reconstructed by VISSIM.

Based on these archives, we reinstated traffic flow using VISSIM, a type of simulation software. Figure 4 displays a simulation of the traffic scene.

In the simulation scene, a Roadside Unit (singed as A) is placed at the middle of the road to collect the traffic flow data according to passing vehicles, which are arranged according to actual traffic data sets. Participants in VANETs, vehicles, and the Roadside Unit are linked by Dedicated Short Range Communications (DSRC) and the communication distance is

Table 1: Inductive loop detector details.

Sign in Figure 4	Number of intersections	Number of detector	Type of detector
a	168	J83241	T4-16
b	167	J83240	T4-16
c	165	J87222	T8-8
d	166	J87221	T8-8

set to 200 m. Vehicles can broadcast their speed and location by DSRC as a special frequency. 10 Hz is recommended by the Vehicle Safety Communications Project, which is distributed by the U.S. Department of Transportation [28]. The Roadside Unit receives this information by DSRC. Meanwhile, four loop sensor detectors (signed as a, b, c, and d) are placed to collect data regarding traffic volume and occupancy. This data is sent to Roadside Unit A by a transmission interface, such as RS232 and RJ45. Data on vehicle speed is sent by vehicles to Roadside Unit A by DSRC. Then data on attacker activities is analyzed and processed on Roadside Unit A. IA^{2}P is installed on the Roadside Unit.

As the simulation running, the information of vehicles and loop sensor detectors is recorded in special files of NS3. The Roadside Unit extracts the data from these files and composes the traffic flow data sets. The whole processing of achieving the data sets is as follows.

When a sensor data package of a vehicle is achieved, Roadside Unit A identifies the vehicle's position according to the value of position in the package and picks up the value of speed (v) from the package. Meanwhile Roadside Unit A reads the volume (q) and occupancy (p) from the corresponding loop sensor detector. As a result, it records them as $x(v, p, q)$. So, the new traffic flow data sets $x(v, p, q)$ are achieved. $x(v, p, q)$ will be converted into CV(x). And the Catastrophe Vector sets are formed according to data series of CV(x).

In our experiments, we focus on simulating and analyzing the performance of IA^{2}P, so the delay of communication and multihop communication pattern are not considered in this paper although these factors could influence the data set building.

4.2. Analysis for Traffic Flow Character Based on the Cusp Catastrophe Model. This section focuses on the character analysis of traffic flow data based on the cusp catastrophe model. This is the theoretical basis of injection attack analysis and processing in this paper.

According to the aforementioned processing of the data sets, the traffic flow data are collected by the Roadside Unit from vehicles and loop detectors. The means of a, b, c, and d are shown in Figure 5. It is evident that the traffic flow character is nonlinear and a catastrophe.

Figure 6 displays a diagram of the speed-volume. It also displays the cusp catastrophe of traffic flow data. One traffic volume value is versus two speed values (v, v'), which means that $z(v, p, q)$ is collected when the vehicle is in the uncongested traffic flow state. The other (v', p', q) is collected when the vehicle is in the congested traffic flow state. The alteration of traffic flow from the uncongested to the congested state is not a gradual process, but an instant jump or catastrophe. As a result, according to the data sets, CV and the evaluation function based on the cusp catastrophe model are fitted.

Table 2: The performance of IA^{2}P.

	Valid data	Malicious data	
		MRD	FD
CMD set	1338	30	70
Result of IA^{2}P	1337	25	76

4.3. Analysis for Injection Attack Detecting and Processing. In this subsection, we mainly analyze the performance of IA^{2}P. Because we focus on detecting malicious data injected, we especially compare IA^{2}P with the method proposed in [18].

According to the definition of malicious data, we manually alter the data set shown in Figure 5. 1438 data values are picked up. 30 of them are altered to be MRD, and 70 are altered to be FD. So a new data set with malicious data is built, named collection with malicious data set, CMD set.

To be guaranteed that the picking method of MRD set and FD set would not affect the performance of IA^{2}P in simulations, we pick them randomly and repeat this process 100 times. At last, we build 100 data sets with malicious data.

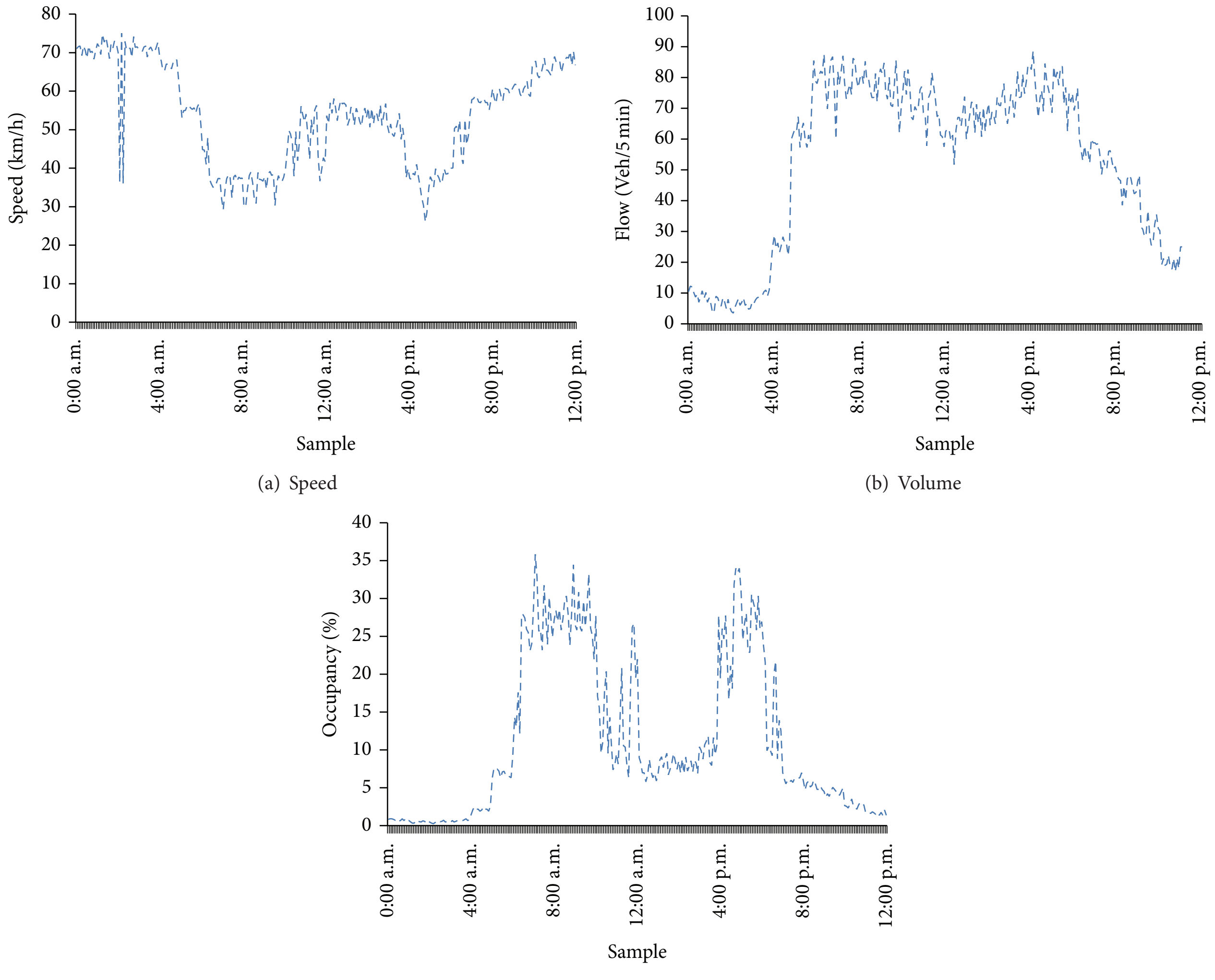
IA^{2}P is proposed to mainly recognize and process malicious data injected by attacker in VANETs. Using the CMD set, IA^{2}P is performed, and the results of one data set with malicious data are shown in Figure 7. There are three integer values predefined to represent the kind of data distinguished by IA^{2}P. Output = 1 represents the fact that the data is valid; output = 2 states that the data is MRD; output = 3 means that the data is FD.

We repeat this process 100 times according to 100 data sets with malicious data. The performance analysis results of 100 data sets are shown in Table 2. The mean recognition rate of MRD is 83.33%. And the mean recognition rate of FD is 100%, while 6 valid data values are recognized as the FD by mistake. As a result, the mean recognition rate of the malicious data is 95%.

Because of lack of the similar researches to detect the injection attacks from the view of the traffic flow theory, we compare the performance of IA^{2}P with the method proposed in [18] which was used to analyze the accuracy of measure data. Ki et al. [18] use the method with the filter to process traffic data. Based on the traffic flow theory, the data was recognized bad data if the data deviation was more than 5% with the former one collected.

Similarly, the CMD set is used, and two integer values were predefined as output to identify whether the data is valid or not; the results of the same set as Figure 7 are shown in Figure 8. Output = 0 means the data is valid. And output = 1 means the data is malicious data.

Also repeating this process using those 100 data sets, there are mean 181 data values out of 1438 considered to be



(a) Speed

(b) Volume

(c) Occupancy

Figure 5: A test example of 24-hour traffic data set from the on-road simulating result.

Table 3: The performance of the method in [18].

	Valid data	Malicious data
CMD set	1338	100
Result of [18]	1257	181

Table 4: The details of malicious data identified in [18].

Valid data but recognized as malicious data	Malicious data		
	MRD	FD	
CMD set	0	30	70
Result of [18]	131	6	55

malicious data, which is shown in Table 3. To have more details, in the 181 data values, 6 of 30 MRD values in the CMD set are recognized as malicious data exactly, which mean recognition rate is 20%. And 55 of 70 FD values are recognized as malicious data exactly, which mean recognition rate is 78.6%. Unfortunately, 131 data values of valid data are recognized as malicious data, which are mainly caused by the nonlinearity and catastrophe of traffic flow. They are shown in Table 4.

This subsection covers verification of the coefficient self-adjustment of generalized IA^{2}P, based on the batch estimation filter. Here, we focus on m and n in (9) since their values shift with traffic flow patterns. In practice, it is an impossible mission to manually set and adjust the coefficients of IA^{2}P for each traffic flow pattern. As the simulations proceeded, we found that this factor could influence IA^{2}P's performance.

Based on the above analysis, we found that it was necessary to carry out this procedure. First, we set the initialization of (m, n) and then recorded their value after each self-adjustment with the IA^{2}P running. Results are shown in Figure 9. It is evident that (m, n) gradually trends towards a steady state.

5. Conclusion and Future Work

VANETs are just like an Achilles heel. On one hand, VANETs are considered as a more efficient and convenient method to

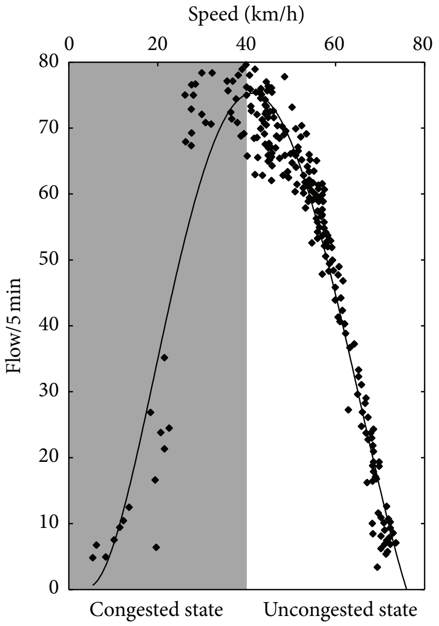


Figure 6: Speed-volume scatter plots for the traffic flow.

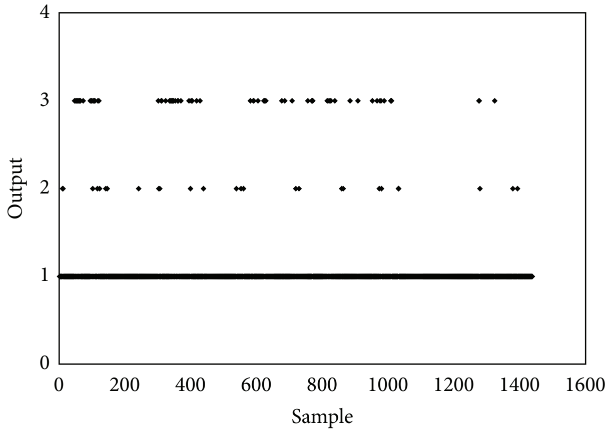


Figure 7: The results distinguished by IA^{2}P.

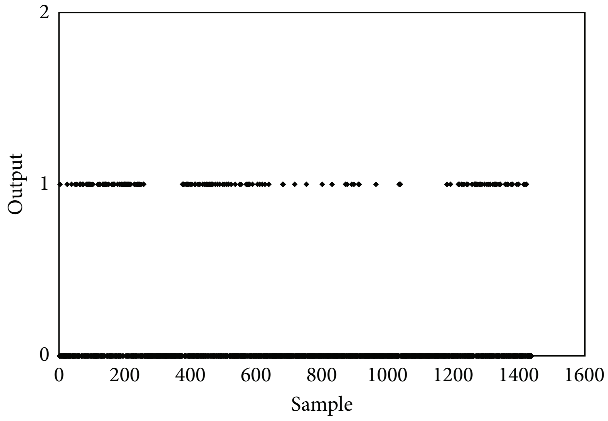


Figure 8: The results distinguished by the method in [18].

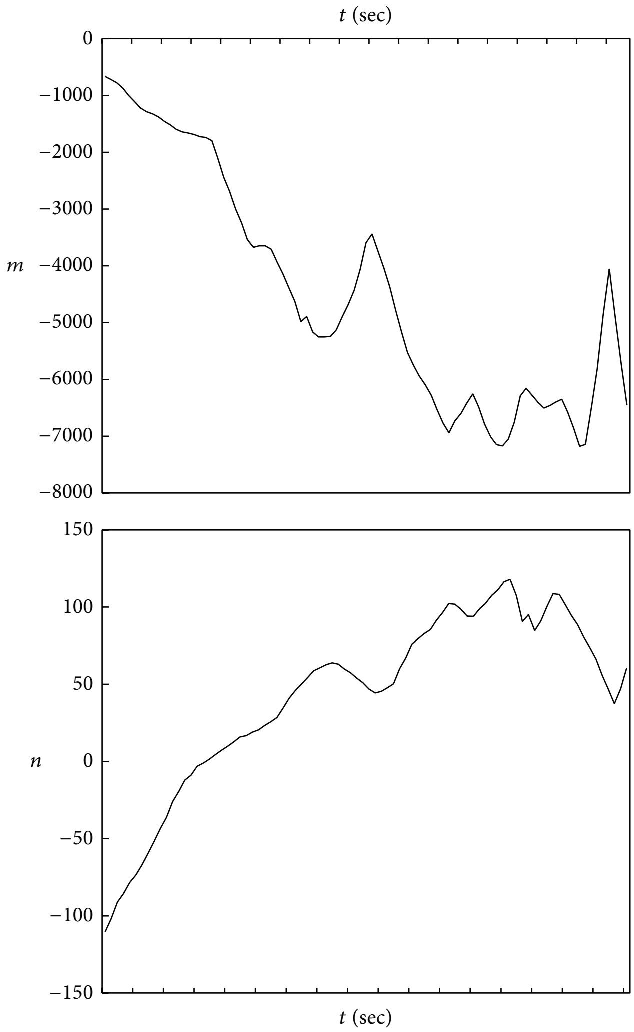


Figure 9: Variation curves of coefficients (m, n).

collect the traffic flow data for ITS application, comparing with the traditional methods. On the other hand, considering the security, a formidable set of abuses and attacks becomes possible and harmful for VANETs, because their networks are wirelessly accessed and exoteric for each participant.

In this paper, we firstly identify a previously unknown vulnerability in the current techniques aimed at security establishment against the malicious data injection attack in VANETs. Then, we investigate the mechanism of this vulnerability, especially for two kinds of the malicious data injection attack: multirepeat data injection attack and fake data injection attack. And then, we propose an intrusion-tolerant security mechanism based on the theory of traffic flow and the model of cusp catastrophe, IA^{2}P, to protect the traffic flow data collection in VANETs. At last, the simulation results show that the recognition rate of the malicious data is 94%, which is more useful and more practical than the existing methods.

In our future work, we would like to extend our results with thinking about vehicle privacy, because the MAC of participant in VANETs is exposed in this paper, and it is still dangerous for VANETs. So our research will have a focus on the way to express the participant's identification without exposing the vehicle privacy.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This work was supported in part by the National High Technology Research and Development Program of China (863 Program) under Grant no. 2012AA111900, the National Natural Science Foundation of China no. 61471084, and the Fundamental Research Funds for the Central Universities (DUT15QY02).

References

[1] F. Terroso-Sáenz, M. Valdés-Vela, C. Sotomayor-Martínez, R. Toledo-Moreo, and A. F. Gómez-Skarmeta, "A cooperative approach to traffic congestion detection with Complex Event processing and VANET," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 914–929, 2012.

[2] J. Zhang, F.-Y. Wang, K. Wang, W.-H. Lin, X. Xu, and C. Chen, "Data-driven intelligent transportation systems: a survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1624–1639, 2011.

[3] T. Leinmüller, R. K. Schmidt, E. Schoch, A. Held, and G. Schäfer, "Modeling roadside attacker behavior in VANETs," in *Proceedings of the IEEE Globecom Workshops (GLOBECOM '08)*, pp. 1–10, New Orleans, La, USA, November 2008.

[4] M. A. Moharrum and A. A. Al Daraiseh, "Toward secure vehicular ad-hoc networks: a survey," *IETE Technical Review*, vol. 29, no. 1, pp. 80–89, 2012.

[5] A. Wasef, R. Lu, X. Lin, and X. Shen, "Complementing public key infrastructure to secure vehicular ad hoc networks," *IEEE Wireless Communications*, vol. 17, no. 5, pp. 22–28, 2010.

[6] R. Lu, X. Lin, H. Zhu, P.-H. Ho, and X. Shen, "ECPP: efficient conditional privacy preservation protocol for secure vehicular communications," in *Proceedings of the 27th IEEE Communications Society Conference on Computer Communications (INFOCOM '08)*, pp. 1229–1237, April 2008.

[7] A. Wu, J. Ma, and S. Zhang, "RATE: a RSU-aided scheme for data-centric trust establishment in VANETs," in *Proceedings of the 7th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM '11)*, pp. 1–6, Wuhan, China, September 2011.

[8] S. Mazilu, M. Teler, and C. Dobre, "Securing vehicular networks based on data-trust computation," in *Proceedings of the 6th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, pp. 51–58, Barcelona, Spain, October 2011.

[9] J. Sun, C. Zhang, Y. Zhang, and Y. Fang, "An identity-based security system for user privacy in vehicular ad hoc networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, no. 9, pp. 1227–1239, 2010.

[10] A. Wasef, Y. Jiang, and X. Shen, "DCS: an efficient distributed-certificate-service scheme for vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 2, pp. 533–549, 2010.

[11] E. Schoch, B. Bako, S. Dietzel, and F. Kargl, "Dependable and secure Geocast in vehicular networks," in *Proceedings of the 7th ACM International Workshop on Vehicular Inter-Networking (VANET '10)*, pp. 61–68, September 2010.

[12] M. Raya, P. Papadimitratos, V. D. Gligor et al., "On data-centric trust establishment in ephemeral ad hoc networks," in *Proceedings of the IEEE International Conference on Computer Communications (IEEE INFOCOM '08)*, pp. 1238–1246, Phoenix, Ariz, USA, April 2008.

[13] B. Aslam, S. Park, C. C. Zou, and D. Turgut, "Secure traffic data propagation in vehicular ad hoc networks," *International Journal of Ad Hoc and Ubiquitous Computing*, vol. 6, no. 1, pp. 24–39, 2010.

[14] K. Sha, S. Wang, and W. Shi, "RD4: role-differentiated cooperative deceptive data detection and filtering in VANETs," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 3, pp. 1183–1190, 2010.

[15] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security*, vol. 14, no. 1, article 13, 2011.

[16] S. Roy, M. Conti, S. Setia, and S. Jajodia, "Secure data aggregation in wireless sensor networks," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 1040–1052, 2012.

[17] M. S. Al-Kahtani, "Survey on security attacks in Vehicular Ad hoc Networks (VANETs)," in *Proceedings of the 6th International Conference on Signal Processing and Communication Systems (ICSPCS '12)*, pp. 1–9, December 2012.

[18] Y.-K. Ki and D.-K. Baik, "Model for accurate speed measurement using double-loop detectors," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 4, pp. 1094–1101, 2006.

[19] I. Martinovic, N. Gollan, L. Cappellaro, and J. Schmitt, "Chaotic communication improves authentication: protecting WSNs against injection attacks," *Security and Communication Networks*, vol. 2, no. 2, pp. 117–132, 2009.

[20] G. Orosz, R. E. Wilson, and G. Stépán, "Traffic jams: dynamics and control," *Philosophical Transactions of the Royal Society A: Mathematical,Physical and Engineering Sciences*, vol. 368, no. 1928, pp. 4455–4479, 2010.

[21] D. C. Gazis, R. Herman, and R. W. Rothery, "Nonlinear follow-the-leader models of traffic flow," *Operations Research*, vol. 9, no. 4, pp. 545–567, 1961.

[22] P. Wagner, "Fluid-dynamical and microscopic description of traffic flow: a data-driven comparison," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 368, no. 1928, pp. 4481–4495, 2010.

[23] F. P. D. Navin, "Traffic congestion catastrophes," *Transportation Planning and Technology*, vol. 11, no. 1, pp. 19–25, 1986.

[24] F. L. Hall, "An interpretation of speed-flow-concentration relationships using catastrophe theory," *Transportation Research Part A: General*, vol. 21, no. 3, pp. 191–201, 1987.

[25] X. Lignos, G. Ioannidis, and A. N. Kounadis, "Non-linear buckling of simple models with tilted cusp catastrophe," *International Journal of Non-Linear Mechanics*, vol. 38, no. 8, pp. 1163–1172, 2003.

[26] J. Guo, X.-L. Chen, and H.-Z. Jin, "Research on model of traffic flow based on cusp catastrophe," *Control and Decision*, vol. 23, no. 2, pp. 237–240, 2008.

[27] N. Ding, G. Tan, W. Zhang, and H. Ge, "Distributed algorithm for traffic data collection and data quality analysis based on wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2011, Article ID 717208, 9 pages, 2011.

[28] US Department of Transportation, *Vehicle Safety Communications Project Task 3 Final Report*, US Department of Transportation, Washington, Wash, USA, 2005.

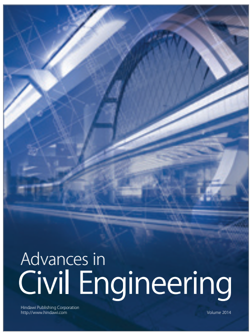
Journal of
Engineering

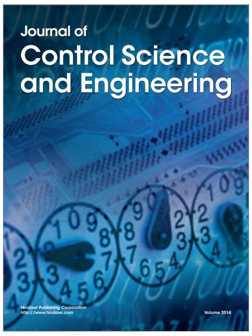
The Scientific World Journal

International Journal of
Rotating Machinery

Journal of
Sensors

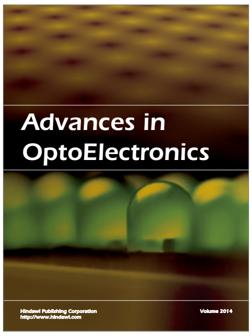
International Journal of
Distributed Sensor Networks


Advances in
Civil Engineering


Journal of
Control Science and Engineering

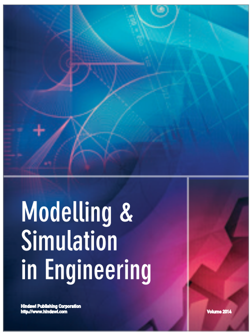
Journal of
Robotics

Journal of
Electrical and Computer Engineering

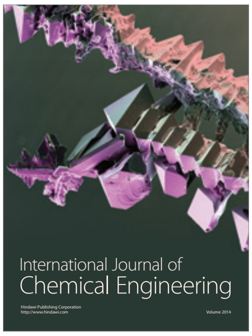

Advances in
OptoElectronics

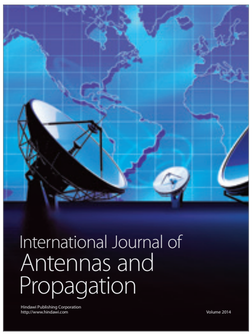
VLSI Design

International Journal of
Navigation and Observation

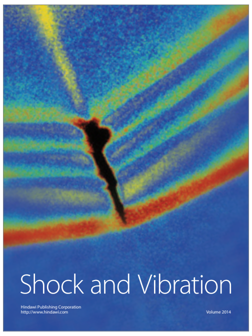

Modelling & Simulation in Engineering

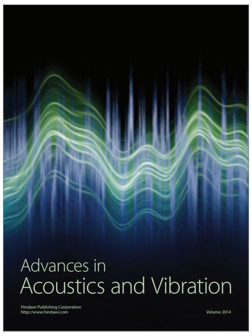
International Journal of
Aerospace Engineering


International Journal of
Chemical Engineering


International Journal of
Antennas and Propagation

Active and Passive
Electronic Components


Shock and Vibration


Advances in
Acoustics and Vibration