Sybil Nodes Detection Based on Received Signal Strength Variations within VANET

Mohamed Salah Bouassida, Gilles Guette, Mohamed Shawky, and Bertrand Ducourthial
(Corresponding author: Mohamed Salah Bouassida)

Laboratoire Heudiasyc UMR CNRS 6599, Université de Technologie de Compiègne
BP 20529, 60205 Compiègne, France (Email: forname.name@utc.fr)
(Received Nov. 22, 2007; revised and accepted Apr. 9, 2008)

Abstract

A Vehicular Ad hoc Network is a collection of mobile hosts forming a temporary network without the aid of any established infrastructure. This flexibility in space and time induces new challenges towards the security needed to support secure communications. Indeed, VANET are subject to attacks due to their vulnerabilities; one of the most compromising attacks is the Sybil nodes attack. We present in this context a Sybil detection approach, based on received signal strength variations, allowing a node to verify the authenticity of other communicating nodes, according to their localizations. In addition, we define an estimated metric of the distinguishability degree between two nodes, allowing to determine Sybil and malicious ones within VANET. The applicability of our contributions is validated through geometrical analysis, simulations and real measurements.

Keywords: Distinguishability degree, RSSIs variations, sybil attack detection, VANET

1 Introduction

Mobile Ad Hoc Networks have undergone incredible growth of popularity during the last years. One of the most practical example of these networks is Vehicular Ad Hoc Network (VANET). The use of wireless communication in VANET implies an always increasing number of potential applications in these networks such as driving assistance, road traffic information or emergency braking alert. All these applications need to exchange data with other vehicles that may be related to the driver safety. The need of confident communications between such critical applications becomes obvious.

One possible threat is the creation of multiple fake nodes broadcasting false information. This attack is known as the Sybil attack [6]. Several security schemes based on keys management have been proposed for intrusion detection and intruder nodes revocation. However, these approaches are constraining within wireless networks, as there is no centralized administration ensuring nodes authentication. For this reason, we propose in this context a Sybil detection technique based on physical signal characteristics, easily mesurable by the commonly used wireless cards. Our technique allows to detect malicious and Sybil nodes within VANET by using received signal strength variations, localization verification and nodes distinguishability degree evaluation. We first show, via geometrical analysis, that an attacker should not increase its sending power. Then, by successively measuring the received signal strength variations, we obtain an estimation of relative nodes localization. This rough localization gives an accurate enough indication on the coherence of the received signal strengths and on how much a pair of nodes could be distinguished from each other, known as "the distinguishability degree". Our geographical localization technique takes into account the characteristics of the wireless networks, such as mobility and dynamicity of nodes, assuming that all messages are sent with the same signal strength, which is not particulary constraining as a Sybil attacker emitting with constant power level has more chances to remain covered up.

To present our contributions, this paper is structured as follows. Section 2 presents related works concerning Sybil nodes detection techniques. Section 3 provides a geometrical analysis and discussions of Sybil attacks. Section 4 presents our relative localization technique, based on received signal strength gradient. In Section 5, we present our technique to detect intruders and malicious nodes within VANET, by evaluating the distinguishability degree between nodes. Section 6 presents analysis, simulations and real tests in order to validate our contributions and finally Section 7 concludes this paper with our aimed future work.

2 Existing Sybil Detection Approaches

The Sybil attack was first described and formalized by Douceur in [6]. In a VANET, a node has knowledge about its neighborhood only with messages it receives. The Sybil attack consists in sending multiple messages from one node (the attacker) with multiple identities. Hence, the attacker simulates several nodes in the network. Different types of attacks that can be launched with Sybil nodes in sensor networks are described in [10]. Applications of the Sybil attack to Vehicular Ad-Hoc Networks have been discussed in [1, 13]. The goal of these attacks could be to give an illusion of a traffic jam to force other vehicles to leave the road to the benefit of the attacker. But the attack could be more dangerous, targeting directly human life for instance, trying to provoke collision in a vehicle platoon [1]. This shows the importance of detecting Sybil nodes in VANET.

One important result shown in [6] is that without a logically centralized authority, Sybil attacks are always possible (*i.e.* may remain undetected) except under extreme and unrealistic assumption of resource parity and coordination among entities. That is to say, entities have the same resources constraints, all identities are validated simultaneously by all entities. We explore in the following a classification of the different defenses proposed.

Douceur [6] and Newsome and al. [10] propose resources testing as a defense against Sybil attack. This resource testing is based on the assumption that each physical entity is limited in some resource. The method described in [6] uses computational puzzles [9] to test nodes computational resources. In [10], the authors show that this approach is not suitable to ad-hoc networks, and hence typically VANET, because the attacker can have more computational resources than a *honest* node. Moreover, they emphasize a problem of network congestion due to the multiple requests/replies for identities checking. Instead, they propose a radio resource testing. However, in VANET the attacker can use multiple radio devices to overcome this detection method.

In [14], the authors try to solve the security problem of the Sybil attack with public key cryptography and authentication mechanism. The authors propose the use of a PKI for VANET (VPKI). They describe a complete solution to provide security of communications and they address the problem of key distribution and privacy. They also propose a mechanism for the most challenging problem: the key revocation. This solution is based on a set of three revocation protocols and a kind of base stations support to send revocation messages. As each vehicle may be authenticated with public key cryptography, the Sybil attack is always detected. Nevertheless, deploying PKI for VANET is a heavy and difficult solution that must be tested to assess its possible use in a real world due to the VANET characteristics. In a VANET, access to network infrastructure is not guaranteed and cryptographic processing may be too long to be usable (tests regarding the time required to sign typical VANET messages can be found in [12]).

Another possibility to defeat Sybil attack is to provide the security of the positioning system and the reliability of the position claimed by vehicles. In [15], the authors propose methods for determining a transmitting peer's node location using signal properties and trusted peers collaboration for identification and authentication purposes. The method uses characteristics such as signal strength and direction so it assumes directional antennas and node's cooperation.

In [5], the authors made a brief survey of positioning techniques and related papers and describe their vulnerabilities to distance enlargement or reduction for position spoofing. They also present a novel approach called *verifiable multilateration*, using distance bounding protocol [4] and base stations. They also assume that all network nodes can establish pairwise secret keys (this is difficult to establish in the mobile ad hoc network).

In [8], the authors propose an approach to evaluate the validity of VANET data. Data are correlated and scored; data with the higher score will be accepted. The proposed model rely on four assumptions. The second assumption is called local distinguishability and rely on the fact that nodes are equipped with specific devices allowing to tie a message with a physical sources. This model uses also short life public key pair generated by the node to extend the distinguishability allowing to authenticate messages coming from a node that keeps its public key during a given time.

To avoid the deployment of a public key infrastructure within VANET, or the addition of specific devices allowing to detect Sybil nodes, some research works [11, 16] use the received signal power to deduce some inconsistencies between the power of the signal and the claimed position. In [16], when a node received a beacon message, it collects signal strength measurement from this node and estimates its new position. A node is considered suspect if its claimed position is too far from the evaluated one. Note that [11] made very strong assumptions about devices and environment. In this context of Sybil attack detection using signal characteristics, we propose in this paper a Sybil detection approach, based on received signal strength gradient, allowing a node to verify the authenticity of other communicating nodes, according to their localizations, and thus to detect malicious and Sybil entities within VANET.

3 Geometrical Analysis of Sybil Attacks

In this section, we provide a geometrical characterization of the success area of a Sybil attack. We begin with some notations and the problem formalization.

3.1 Notations

Let us consider S and R be two mobile nodes such that S sends some messages received by R. We assume that the transmission of a single message is immediate, allowing to consider the positions of S and R as fixed points of the space at the time of transmission. We denote by $\mathrm{d}(S,R)$ the distance between S and R.

Let denote by P_{snd} the sending power of the node S. With an isotropic antenna of gain G_{snd} and for $\mathrm{d}(S,R)$ sufficiently large, the node R will receive a power P_{rcv} equals to:

$$P_{\mathrm{rcv}} = P_{\mathrm{snd}} \times \frac{G_{\mathrm{snd}} \times G_{\mathrm{rcv}} \times \lambda^2}{16\pi^2 \times \mathrm{d}^2(S,R)},$$

where λ denotes the wavelength of the radiation.

By denoting $G_{\mathrm{SR}} = G_{\mathrm{snd}} \times G_{\mathrm{rcv}} \times \lambda^2/(16\pi^2)$ the gain of the link from S to R, the maximal power $P_{\mathrm{rcv}}^{\max}(dist(S,R))$ at distance $\mathrm{d}(S,R)$ from the sender can be rewritten as:

$$P_{\mathrm{rcv}}^{\max}(\mathrm{d}(S,R)) = P_{\mathrm{snd}} \times G_{\mathrm{SR}} \times \frac{1}{\mathrm{d}^2(S,R)} \quad (1)$$

By taking into account signal attenuation, the power received by R is smaller than $P_{\mathrm{rcv}}^{\max}$:

$$P_{\mathrm{rcv}}(\mathrm{d}(S,R)) = \alpha \times P_{\mathrm{rcv}}^{\max}(\mathrm{d}(S,R)) \qquad 0 \leq \alpha \leq 1, \quad (2)$$

where α depends on several parameters (distance $\mathrm{d}(S,R)$, λ, atmospheric conditions...).

We denote by $\mathrm{d}_{\min}$ the minimal distance between the antenna of a sender and the antenna of a receiver. We can define the maximal received power $P_{\mathrm{rcv}}^{\max}$ for a receiver close to a sender as:

$$P_{\mathrm{rcv}}^{\max} = P_{\mathrm{snd}} \times G_{\mathrm{SR}} \times \frac{1}{\mathrm{d}_{\min}^2} \quad (3)$$

3.2 Sybil Attacks and Hypotheses

We suppose that each vehicle is equipped with a standard embedded device, in such a way that antennas, gains and sending powers are fixed and known. We can legitimately assume that the future standard track defining the wireless communication in VANET will fix all these characteristics. Moreover we assume that each car periodically diffuses some *hello messages* containing their GPS position.

To create a Sybil node F, a sender S could give some false GPS positions in its messages. However, thanks to the previous equations, a receiver R may detect a mismatch between the measured received power P_{rcv} and the GPS positions inside the message. To complicate the Sybil node detection by the receivers, the sender may use a non standard equipment. This implies that its sending power could vary instead of being fixed.

We assume that all the mobile nodes (cars) are on the same two-dimensional Euclidean space, approximating the earth's surface. To estimate the number of potentially cheated cars, we can equivalently compute the area of the Euclidean plan where the Sybil attack cannot be detected.

3.3 Results

In this section, we consider a real propagation environment using a signal attenuation α. Such an attenuation depends on different parameters, including the distance from the sender to the receiver. However, it is not known and a receiver can not deduce the exact distance from the sender. Instead, it can only deduce the maximal distance from the sender, corresponding to an attenuation factor α of 1 (free space propagation).

We begin with some preliminary results needed to prove our main propositions.

Lemma 1. *Let S, F and O be three points on a line satisfying $\overline{SO} = \frac{\alpha}{\alpha-1} \times \overline{SF}$ (with $0 < \alpha$ and $\alpha \neq 1$). The set of points R satisfying $\sqrt{\alpha} \times \mathrm{d}(F,R) = \mathrm{d}(S,R)$ is the circle $\mathcal{C}_\alpha$ of radius $\frac{\sqrt{\alpha}}{|\alpha-1|} \times \mathrm{d}(S,F)$ and centered on O.*

Proof. Let $(\mathrm{S}, \frac{\overrightarrow{SF}}{\|SF\|}, \overrightarrow{j})$ be an orthonormal frame. The coordinates of the O point in this frame are $(\frac{\alpha}{\alpha-1}, 0)$ and the equation of $\mathcal{C}_\alpha$ is:

$$(x - \frac{\alpha}{\alpha-1})^2 + y^2 = \frac{\alpha}{(\alpha-1)^2}$$

We have:

$$\begin{array}{l} \alpha \times \mathrm{d}^2(F,R) = \mathrm{d}^2(S,R) \\ \Leftrightarrow \alpha \times (1-x)^2 + \alpha \times y^2 = x^2 + y^2 \\ \Leftrightarrow (\alpha-1)(x^2+y^2) - 2 \times \alpha \times x = -\alpha \\ \Leftrightarrow x^2 - \frac{2\times\alpha}{\alpha-1} \times x + y^2 = \frac{-\alpha}{\alpha-1} \\ \Leftrightarrow (x - \frac{\alpha}{\alpha-1})^2 + y^2 = \frac{\alpha}{(\alpha-1)^2} \end{array}$$

Hence, all the points R satisfying $\sqrt{\alpha} \times \mathrm{d}(F,R) = \mathrm{d}(S,R)$ are on the circle $\mathcal{C}_\alpha$. □

Proposition 1. *Let O_α be a point of the (S,F) line such that $\overline{SO_\alpha} = \frac{\alpha}{\alpha-1} \times \overline{SF}$. Let $\mathcal{C}_\alpha$ be the circle of radius $\frac{\sqrt{\alpha}}{|\alpha-1|} \times \mathrm{d}(S,F)$ and centered on O_α.*

With an omni-directional antenna, a standard sending power P_{snd} and a signal attenuation $\alpha < 1$, the Sybil attack of S cannot be detected from the nodes R (i) outside the circle $\mathcal{C}_\alpha$ and (ii) inside the disk $\mathcal{C}_{\max}^{\alpha}$.

Proof. The receiver R does not know the value of α. Hence, when S sends a message with the sending power P_{snd}, R will measure a power $P_{\mathrm{rcv}}^{\alpha}$ and compute an erroneous distance $\mathrm{d}_\alpha(S,R)$ from S to R using the free space propagation model (see Equation 1):

$$\mathrm{d}_\alpha(S,R) = \sqrt{\frac{P_{\mathrm{snd}}}{P_{\mathrm{rcv}}^{\alpha}} \times G_{\mathrm{SR}}}.$$

By Equation 1, $P_{\mathrm{rcv}}^{\alpha} = \alpha \times P_{\mathrm{rcv}}$ and we have

$$\mathrm{d}_\alpha(S,R) = \sqrt{\frac{P_{\mathrm{snd}}}{\alpha \times P_{\mathrm{rcv}}} \times G_{SR}} = \frac{1}{\sqrt{\alpha}} \times \mathrm{d}(S,R).$$

In the message from the node S, the node R will read the position of the Sybil node F and will compute the

distance $\mathrm{d}(F,R)$. To be cheated by the Sybil attack, the receiver node R must satisfy $\mathrm{d}_\alpha(S,R) > \mathrm{d}(R,F)$, that is $\mathrm{d}(S,R) > \sqrt{\alpha} \times \mathrm{d}(F,R)$. By Lemma 1, the cheated nodes are then on the circle $\mathcal{C}_\alpha$.

Moreover, R cannot receive a message if it is outside the sender's range equal to $\sqrt{\alpha} \times \mathrm{d}_{\max}$. Hence, the cheated receiver nodes are those (i) on the circle $\mathcal{C}_\alpha$ and (ii) inside the disk $\mathcal{C}^\alpha_{\max}$. □

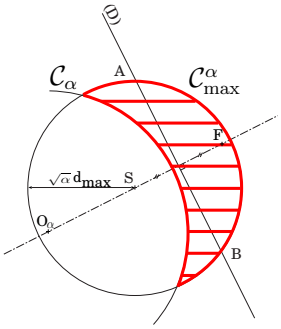

Figure 1: Omni-directional antenna and standard sending power with attenuation. Cheated nodes are outside the circle $\mathcal{C}_\alpha$ and inside the disk $\mathcal{C}^\alpha_{\max}$ (hatched area).

Lemma 2. *Assuming a free space signal propagation model, let S be the sender of a message such as its sending power is equal to $\beta \times P_{\mathrm{snd}}$, $\beta > 1$. The received signal power of this message is larger than $P^{\max}_{\mathrm{rcv}}$ if and only if the receiver is inside the disk $\mathcal{C}^\beta_{\min}$ of radius $\sqrt{\beta} \times \mathrm{d}_{\min}$ and centered on S.*

Proof. Let S be a sender and R be a receiver such that the sending power of S is $\beta \times P_{\mathrm{snd}}$, $(\beta > 1)$ and the received power of the node R is P^β_{rcv}. Then, we have:

$$\begin{array}{l} P^\beta_{\mathrm{rcv}} \leq P^{\max}_{\mathrm{rcv}}; \\ \beta \times P_{\mathrm{snd}} \times G_{SR} \times \frac{1}{\mathrm{d}^2(S,R)} \leq \frac{P_{\mathrm{snd}} \times G_{SR}}{\mathrm{d}^2_{\min}}; \\ \sqrt{\beta} \times \mathrm{d}_{\min} \leq \mathrm{d}(S,R). \end{array}$$

Hence, if the transmission power of S is equal to $\beta \times P_{\mathrm{snd}}$, $\beta > 1$, every node R inside the circle $\mathcal{C}^\beta_{\min}$ received a signal power greater than $P^{\max}_{\mathrm{rcv}}$. □

Proposition 2. *Let O_β be a point of the (S,F) line such that $\overline{SO_\beta} = \frac{\beta}{\beta-1} \times \overline{SF}$. Let $\mathcal{C}_\beta$ be the circle of radius $\frac{\sqrt{\beta}}{|\beta-1|} \times \mathrm{d}(S,F)$ and centered on O_β.*

With an omni-directional antenna and a non standard sending power $\beta \times P_{\mathrm{snd}}$ with $\beta > 1$, the Sybil attack of S cannot be detected from the nodes R (i) inside the circle $\mathcal{C}_\beta$, (ii) inside the disk $\mathcal{C}^\beta_{\max}$ and (iii) outside the disk $\mathcal{C}^\beta_{\min}$ (Figure 2).

Proof. For Conditions (i) and (ii), see Proposition 1. Condition (iii) is given by Lemma 2. □

We now consider the combined action of the attenuation factor α and the tuning factor β. We denote by γ the product $\alpha \times \beta$.

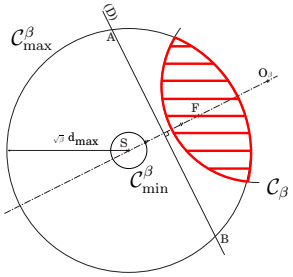

Figure 2: Omni-directional antenna and non standard sending power (drawing with $\beta = 2$). Cheated nodes are inside the disk $\mathcal{C}_\beta$ and inside the disk $\mathcal{C}^\beta_{\max}$ and outside the disk $\mathcal{C}^\beta_{\min}$.

Proposition 3. *Let O_γ be a point of the (S,F) line such that $\overline{SO_\gamma} = \frac{\gamma}{\gamma-1} \times \overline{SF}$. Let $\mathcal{C}_\gamma$ be the circle of radius $\frac{\sqrt{\gamma}}{|\gamma-1|} \times \mathrm{d}(S,F)$ and centered on O_γ.*

With an omni-directional antenna, an attenuation factor α and a tuned sending power $\beta \times P_{\mathrm{snd}}$, the Sybil attack cannot be detected from the nodes R.

- *(i) inside the circle $\mathcal{C}_\gamma$ and (ii) inside the disk $\mathcal{C}^\gamma_{\max}$ and (iii) outside the disk $\mathcal{C}^\gamma_{\min}$ if $\gamma = \alpha \times \beta > 1$ (Figure 3).*
- *(ii) outside the circle $\mathcal{C}_\gamma$ and (ii) inside the disk $\mathcal{C}^\gamma_{\max}$ if $\gamma < 1$ (as on Figure 1).*
- *(iii) belonging to the semi-plan defined by the intersection of the mediator line (D) and the disk $\mathcal{C}_{\max}$ and that does not contain S if $\gamma = 1$.*

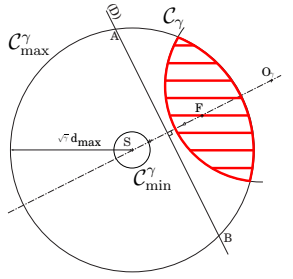

Figure 3: Omni-directional antenna and non standard sending power with attenuation (drawing with $\gamma = 2$). Cheated nodes are inside the disk $\mathcal{C}_\gamma$, outside the circle $\mathcal{C}^\gamma_{\min}$ and inside the disk $\mathcal{C}^\gamma_{\max}$.

3.4 Discussions

We can reasonably assume that a standard for vehicular communication would fix the transmission power of each vehicles. Such a standard transmission power would then be used by all honest transmitting nodes. The only vehicles that may voluntarily bypass this rule are then the attacking nodes.

Increasing the sending power allows to increase the area of successful attacks. However it also increase the area of reception. And when a receiver node is not cheated, it detects the attack. Then, by vehicle cooperation, the attack has more chances to fail. Hence, increasing the sending power could decrease the impact of the attack and could also be compromising for the attacker. There is then a tradeoff between the area of successful attack, and the area of detection.

To evaluate this tradeoff, we study the ratio *area of successful attack over area of reception*. Figure 4 shows the ratio depending on the factor γ (product of attenuation α and tuning factor β) and on the distance between the attacker node S and the Sybil node F. We can see that the ratio is maximal for short values of γ and for short distance d(S, F).

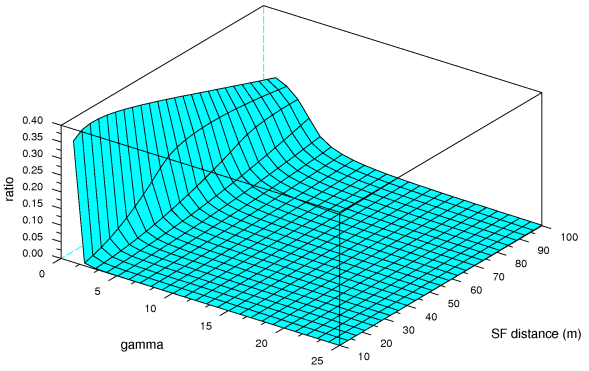

Figure 4: Ratio of the successful attack area by the attacker's signal reception area.

This means that increasing the transmission power is of limited interest for an attacker. The attacker should just tune its transmission power in the aim of compensating the signal attenuation (to obtain $\gamma = 1$). But this is not easily determined. We can conclude that it is better for an attacker to use also the standard sending power P_{snd}.

Note also that if the Sybil nodes should be near the attacker, it will be less easy to simulate a traffic jam by means of many Sybil nodes, because the area of successful attack is reduced.

4 Relative Node Localization Using Signal Strength Gradient

Our Sybil detection approach is based on a relative localization technique using received signal strength variations, under the assumption that all messages are sent with the same signal power. This assumption is legitimate as we have validated that malicious nodes should use a steady transmission power (cf. Section 3.4). We present in the following this localization technique [3], then we describe in the next section our Sybil detection approach.

We present in this section our technique allowing a mobile receiver to localize a fixed sender within its range, with only 3 RSSI measurements, in LoS (Line of Sight) environment and assuming that sent messages are with the same signal strength. The path loss model we use to evaluate the distance between a sender and a receiver is the Friis Free Space Path Loss Model [7], which represents the signal attenuation when there is a clear line of sight between the transmitter and the receiver. This model stipulates that:

$$PL(d)[dB] = 2.PLfs(d0)[dB] + 10.n.log_{10}(d/d0).$$

Where:

- PL is the path loss, PLfs is the path loss within a free space environment.
- λ is the wavelength of the propagation wave. λ is evaluated as $\lambda = c/f$, c is the light speed (3.10^8 m/sec) and f is the frequency of the signal.
- d is the distance between the transmitter and the receiver and $d0$ is a received power reference point. We can choose $d0 = 1m$ without loss of generalization.
- n is the path loss exponent which represents the increase of path loss with the increase of the distance between the transmitter and the receiver. For free space, n is equal to 2, but it would be better to calibrate this parameter, depending on each network characteristics [3].

Our relative localization technique uses the received signal strengths gradient, to estimate distances between two nodes at different positions (according to the Friis model). From these distances is then deduced an angle between the two nodes, as illustrated in Figure 5, where the receiver needs to localize the sender by determining the angle β between them. For that, the receiver starts by evaluating the received strength from the sender, at positions P_1 and P_2. The distance between these positions is L. Then, using the Friis path loss model, the receiver can evaluate the distances d_1 and d_2 at positions P_1 and P_2.

We suppose that the distance x between the receiver and the sender is equal to the average between the distances d_1 and d_2, where $d_1, d_2 >> L$: $x = \frac{d_1+d_2}{2}$. The angle β is computed, through geometrical relations, as follows:

$$\beta = arccos\left(\frac{d_2 - d_1}{L}\right).$$

Let the coordinates of the receiver at the position P_0 be (x_0, y_0), and the coordinates of the sender (x_s, y_s). Because of $cos(x) = cos(-x)$, two localizations of the sender are possible, verifying the equation $\beta = arccos((d2 - d1)/L)$. Thus:

$$\begin{pmatrix} x_s = x_0 + d.sin\beta \\ y_s = y_0 + d.cos\beta \end{pmatrix} or \begin{pmatrix} x_s = x_0 - d.sin\beta \\ y_s = y_0 + d.cos\beta \end{pmatrix}$$

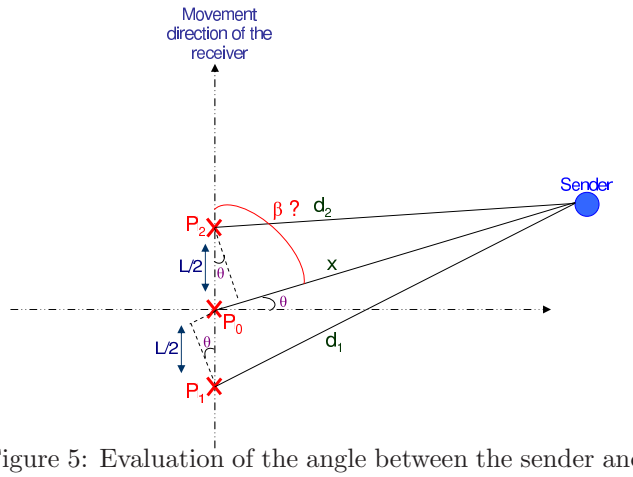

Figure 5: Evaluation of the angle between the sender and the receiver

To be able to decide which position to choose for the sender, the receiver can measure the received signal strength from the sender, in the direction of one of the two localizations. Depending on the increase or the decrease of the received signal strength, the receiver decides which localization to choose for the sender.

Calibration of Our Node Localization Technique. To avoid errors on the RSSI measurements, we should calibrate the exponent loss factor n used in the Friis Loss equation presented above. With $d0 = 1$, we have $PL(d)[dB] = 80 + 10n.log_{10}(d)$. Thus, the distance d and the loss factor n are computed as follows:

$$d = 10^{(PL(d)[dB]-80)/10n}.$$
$$n = (PL(d)[dB] - 80)/10.log_{10}(d).$$

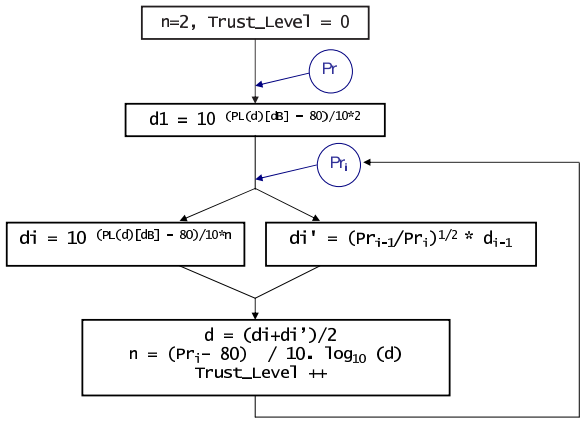

Figure 6: Calibration of the path loss factor

To calibrate n, we use a second formulation of the Friis Model, which stipulates that: $P_r/P_t = (\lambda/4\pi d)^2$; where P_r is the received signal strength and P_t is the transmitted signal strength. For two successive received signals from a transmitter, we show that $P_{r1}/P_{r2} = (d_2/d_1)^2$. We thus have:

$$d_2 = d_1.\sqrt{P_{r1}/P_{r2}}.$$

Our calibration algorithm, illustrated in Figure 6, consists of computing the distance between a transmitter and a receiver as the average between the two values produced by using the two formulations of the Friis Loss equation. The exponent loss factor n is then computed as a function of the computed average distance. We add to our calibration algorithm a trust_level, which represents the number of calibrations carried out by a node in the network. The trust_level is incremented at each calibration. Thus, the higher the trust_level is, the more precise are the measurements.

Discussions. The main advantages of our localization technique are the following:

- No additional equipment has to be added to the wireless nodes. Our localization technique uses only the history of received signal strength, to deliver a reliable and fast localization estimation.
- The node which wants to localize itself can move within the network and does not need to be fixed, as for other localization techniques based for example on triangulation mechanisms.
- The higher is the number of measurements of the received signals strength, the more is the localization precision. Indeed, the measurements of RSSI can calibrate the path loss attenuation model used to compute the distance between the sender and the receiver.

5 Sybil Nodes Detection Approach

Our Sybil detection approach is composed of two complementary techniques. The first one is a localization verification technique based on received signal strength. This technique allows a node to verify the authenticity of another node by estimating its future geographical localizations, and compare them to its evaluated localizations. When a node is detected suspect (incoherent signal strengths gradient), our second technique should be used. This technique is a Sybil detection mechanism, based on the definition of a distinguishability degree metric. This mechanism can be launched individually by every node in the network in order to detect Sybil and malicious ones based on their geographical localizations.

5.1 Coherence Verification of RSSIs Measurements

We present in this section the first technique of our intrusion detection approach. This technique is based only on

RSSIs measurements variations, and allows a node to detect malicious nodes within the network. We divide this section into two cases, according to the mobility of the verifier node.

Fixed Verifier Node. Our objective is to allow a fixed node A to estimate (or predict) the signal strength received from a node B, based on the previous RSSI measurements. Such approach can detect the intrusion of a malicious node in the network, which is trying to usurpate the identity of another node.

- Step 1:

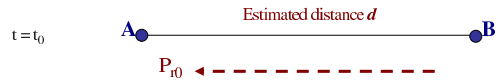

Figure 7: Step 1

Figure 7 shows two communicating nodes A and B at $t = t_0$. The node A measures P_{r0} the strength of the received signal at t_0. Using the Friis Attenuation Model, it can compute the distance d between it and B. From previous RSSI measurements, the node A can evaluate the average speed V of the node B.

- Step 2:

Figure 8 illustrates the possible locations of the node B at $t = t_0 + T$. These positions form a circle whose center is the old position of the node B and of radius equal to $V.T$. It is clear that node A measures the maximum received signal strength when the node B is at the position P_1 (the nearest position to A), and the minimum received signal strength when the node B is at the position P_2 (the most distant position from A).

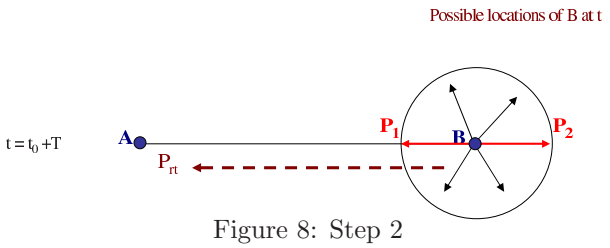

Figure 8: Step 2

- Step 3:

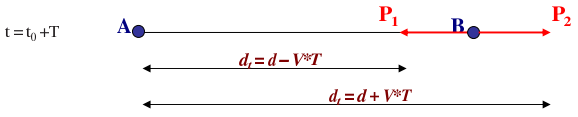

Figure 9: Step 3

In this step, we compute the received signal strength measured by A, when the node B is at the position P_1 and P_2, to delimit the estimated value of P_{rt} (cf. Figure 9). From the Friis Attenuation Model, we have:

$$\frac{P_{rt}}{P_{r0}} = (\frac{d_0}{d_t})^2 \Rightarrow P_{rt} = P_{r0}.(\frac{d_0}{d_t})^2$$

$$\Rightarrow P_{r0}.\left(\frac{d_0}{d_0+V.T}\right)^2 \leq P_{rt} \leq P_{r0}.\left(\frac{d_0}{d_0-V.T}\right)^2$$

Thus, we can conclude that at $t = t_0 + T$, the RSSI measured by the node A should belong to the interval $[P_{r0}.(\frac{d_0-V.T}{d_0})^2, P_{r0}.(\frac{d_0+V.T}{d_0})^2]$. Otherwise, the received message was sent by a Sybil node. In addition, the node A adds to the node B the label "suspect". To illustrate our approach, we use the following example where the average speed V is equal to $30m/sec$ and the period T is defined at $1sec$: each measured RSSI at $t = t_i$ should thus insure that:

$$Pr_{i-1}.\left(\frac{d_{i-1}}{d_{i-1}+30}\right)^2 \leq Pr_i \leq Pr_{i-1}.\left(\frac{d_{i-1}}{d_{i-1}-30}\right)^2$$

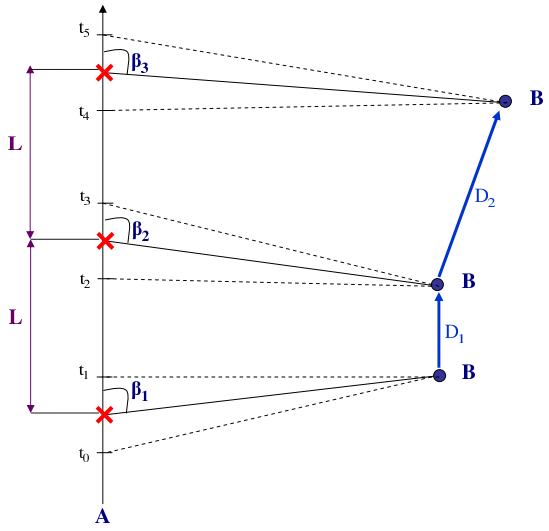

Figure 10: RSSI coherence verification

Mobile Verifier Node. Our objective is to allow a mobile node A to verify the coherence of the localizations of node B, evaluated via our localization technique presented in Section 4, according to its mean speed V. Figure 10 illustrates an example of this approach. To verify the coherence of node B, node A should verify each t_i that:

$$LOC_{B_i} - LOC_{B_{i-1}} <= V.(t_i - t_{i-1}).$$

In the example given in Figure 10, D_1 and D_2 should be smaller than $L.V$. If D_2 is much larger than $L.V$, the node A adds to the node B the label "suspect".

Our localization verification technique presented above (fixed or mobile verifier node A and verified node B), can thus have a positive or negative result. A positive result means that the entity with which the verifier node

communicates is authentic and corresponds to the same wireless node throughout the communication duration. Whereas a negative result of this technique can be the consequence of one of the following reasons:

- The node B creates Sybil nodes in the network, with the same identifier "B", making their neighbors believe that all messages are sent by the same entity.
- The node B is a Sybil one, created by a malicious entity to inject fault messages via the wireless network.
- Messages received by a verifier node A are sent by an intruder node which usurped the identity of the node B with which the verifier node communicated.

As a second step of our intrusion detection approach, and especially in the first two cases presented above, it would be judicious to detect and localize malicious and Sybil nodes among neighbors of the verifier node. Nodes labeled "suspect" as a result of the first technique should be particulary verified, by our second technique in order to detect if it is a Sybil node, and which malicious nodes create it. This technique is presented in the next section.

5.2 Distinguishability Degree Between Two Nodes

To define the distinguishability degree between two nodes X and Y, a verifier node stores their localizations, evaluated through RSSIs measurements variations. We assume that the verifier node receives enough messages from the two nodes, in order to evaluate their localizations each t_i.

We define M the balanced geometric mean of the differences between the localizations of the two nodes X and Y at each t_i, by adding to each difference between the localizations of the two nodes X and Y at each t_i a weight, corresponding to the number of measurements until t_i, as follows:

$$M_i(X <> Y) = \left(\prod_{k=1}^{i} (LOC_{X_k} - LOC_{Y_k})^k \right)^{1/\sum_{k=1}^{i} k}$$

We define D an evaluated distance which a node can traverse with the mean application speed within 2 seconds. We assume that two nodes localized at distance D at t_i are distinct. For example, in VANET we can choose $D = 60m$ (for a mean speed evaluated at 30m/sec).

The distinguishability degree $DD(X <> Y)$ is thus evaluated as the percentage that nodes X and Y are distinct. $DD(X <> Y)$ is computed as follows:

if $(M_i < D)$

$$DD(X <> Y) = \left(100.\frac{M_i(X<>Y)}{D}\right)\%$$

else

$$DD(X <> Y) = 100\%$$

end if

If $DD(X <> Y) \rightarrow 0$, the verifier node can determine that the nodes X and Y are malicious ; one of them is a Sybil node. Otherwise, the two nodes are considered distinct at $DD(X <> Y)$ %. Depending of the concerned wireless application and according to its established security policy, each node in the network can evaluate the distinguishability degree between two particular nodes in its neighborhood (labeled "suspect" for example), or evaluate a higher triangular matrix of distinguishability of all its neighbors, at each time t_i, that we call $M_{Distinguish}$. For a node I, with c neighbors, this matrix is evaluated as follows:

$$M^i_{Distinguish} = \begin{pmatrix} 0 & DD(0 <> 1) & .... & DD(0 <> c) \\ 0 & 0 & .... & DD(1 <> c) \\ .... & .... & 0 & .... \\ 0 & .... & .... & 0 \end{pmatrix}$$

Note that $DD(i <> j)$ is stored at (row i, column j), $DD(i <> i) = 0$ and $DD(i <> j) = DD(j <> i)$. According to this matrix and at each time t_i, each node in the wireless network can evaluate the distinguishability of its neighbors and determine suspect ones. However, it can not determine precisely which entities are Sybil nodes and who creates them.

Determination of Sybil and Malicious Nodes through Geometrical Analysis. At this step of our Sybil detection approach, a verifier node can isolate suspect entities within its neighborhood. To determine precisely Sybil and malicious nodes, it should proceed as follows. For each pair of nodes (S, F) such that $DD(S <> F) \sim 0$, the verifier node can check if it is in a Sybil attack success area depending on the position of suspect nodes (S, F) (cf. Section 3.3).

Figures 11 and 12 shows how a node can identify the malicious node and the Sybil one in the pair (S, F). If there was no inconsistencies in the received signal strength (the Sybil nodes is placed in the success area), the verifier can combine the distinguishability degree of a pair of nodes and the computation of the successful area of the Sybil attack to determine which node is the Sybil one and which is the malicious one. Having the verified position of the nodes S and F, a node V, the verifier can compute the Sybil attack success area considering the node S as the malicious node (Figure 11) and then considering the node F as the malicious node (Figure 12). In Figure 11, the Sybil node is inside the success area so this is consistent with the signal measure. In Figure 12, the Sybil node is outside the success area so il should have been detected. This is inconsistent with the signal measure. The verifier node V can conclude that S is the malicious node and F is the Sybil one.

These combined methods allow to determine precisely the role of each node during the Sybil attack. However, we note that it remains an area where the verifier node cannot conclude about the role of the couple of suspect nodes (S, F). This occurs when the verifier node is placed inside the intersection of the two hatched zones.

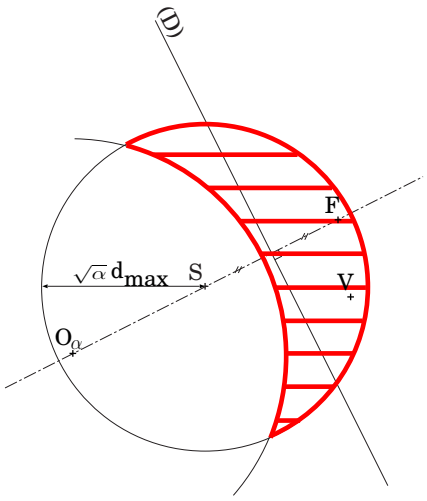

Figure 11: Success area of Sybil attack with S as malicious node and F as Sybil node

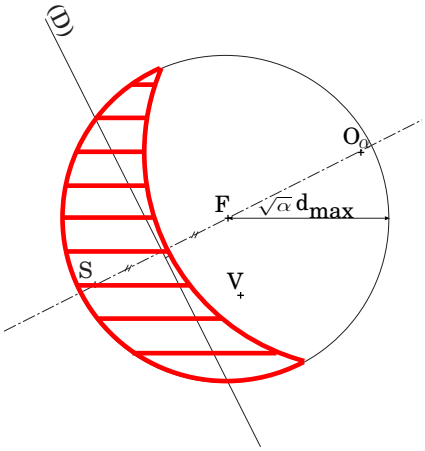

Figure 12: Success area of Sybil attack with F as malicious node and S as Sybil node

6 Analysis and Simulations

To validate the applicability of our contributions depicted above, we present in this section analysis and simulations we have done to evaluate the error margin of our localization technique and validate our distinguishability degree evaluation technique.

6.1 Simulation Results

In this section, we use the network simulator NS-2 to simulate our localization mechanism, described in Section 4.

Our simulation parameters under NS-2 are as follow. The propagation model is Free Space, the MAC protocol is 802.11, the antenna model is omni-directional, the number of nodes is 2 and finally the simulated traffic is CBR (Node 1 sends a CBR traffic to Node 0, which receives these packets, evaluate their RSSIs and compute the distance to reach Node 1).

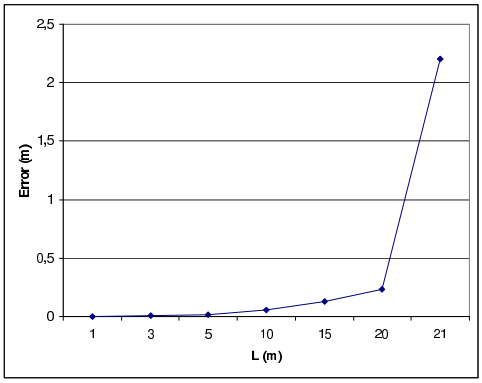

Figure 13: Localization error by L ($d = 180.27m$)

We carried out simulations to evaluate the localization error of our approach. In a first step, we evaluate the localization error according to the distance L between the two positions P_1 and P_2, the distance d is fixed to 180.27 (cf. Figure 13). Then, we evaluate the localization error according to the distance d, while fixing the L parameter ($L = 3m$) (cf. Figure 14).

We show in Figure 13 that the localization error is strongly dependant on the parameter L. The smallest the parameter L is, the smallest is the localization error. For $L = 3m$, the localization error is equal to 0.0043m. However, we cannot decrease indefinitely the parameter L in order not to distort the RSSI measurements.

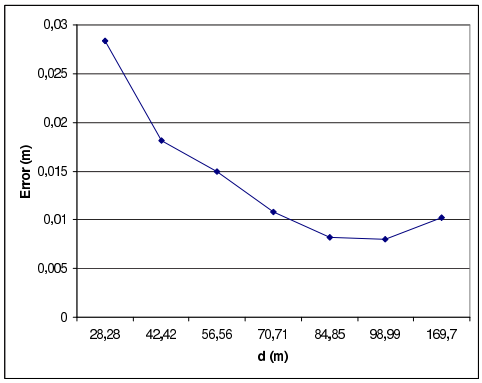

Figure 14: Localization error by the distance d ($L = 3m$)

Figure 14 shows that the localization error is dependant on the distance between the sender and the receiver, with a fixed L. We conclude that to have a small localization error, the parameter L should be much smaller then the distance d. In addition, we have used this hypothesis to compute the angle β between the sender and the receiver (cf. Section 4).

6.2 Validation by Real Measurements and Results

To validate our localization technique, we carried out RSSI measurements between two Nodes A and B, using the wireless tool **iwspy**. Our test framework is composed of a non isotropic source placed in free space as a transmitter antenna with P_T watts and a directivity gain G_T. At an arbitrary large distance d ($d >> \lambda$, where $\lambda = cT = c/f$ is a wavelength) from the source, the radiated power is uniformly distributed over a surface area of a sphere of radius d. If P_R is the power at the receiving antenna, which is located at distance d from the transmitter antenna and has a directivity gain G_R, then the path loss in decibels is given by the following equation (where $\gamma = 2$):

$$L_F = 34.44 + 20\, log_{10} f - 10\, (log_{10} G_T + log_{10} G_R - \gamma\, log_{10} d)$$

The distance d (km) is thus computed as follows:
$d = 10^{(L_F - 34.44 - 20\ log_{10} f + 10\ log_{10} G_T + 10\ log_{10} G_R)/10\ \gamma}$

Our testbed is composed of two laptops, connected via an ad hoc network, equipped each by a Holux GPS and a wireless card (Avaya and Buffalo wireless cards) with external antennas ($G_T = 3dbm$ *and* $G_R = 3dbm$). The frequency of the used channel is $f = 2.457Ghz$.

The objective of our measurements is to verify through GPS that the distance evaluated using the received signal strength indicators is exact with an error margin to be determined.

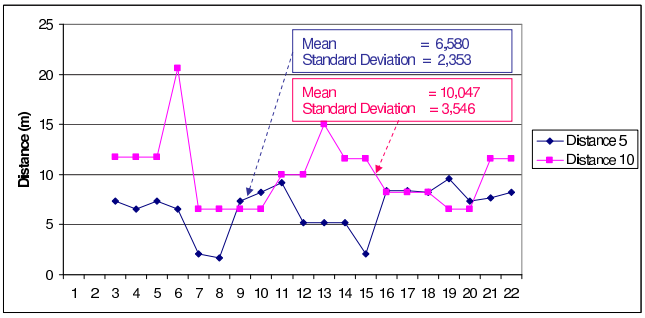

Figure 15: Real RSSIs measurements

Figure 15 shows the measurements we have done between two nodes at $5m$ and $10m$. Although there are peaks in the distances evaluation, the mean distance is equal to $6.58m$ in the case of $5m$, while it is equal to $10.047m$ when the real distance is $10m$. We note also that the standard deviations are not dependent of the measured distances.

The tests we did to validate our localization technique consists of evaluating the localization of a node B with co-ordinates $(0m, 10m)$, by a node A with co-ordinates $(0m, 0m)$. The parameter L of our technique is chosen equal to $3m$. The result of our measurements is the localization of the node B at the co-ordinates (3.29m,10.08m), which represents an error margin equal to $3.29m$. This result is very promising because it is not dependent of the real distance between the two nodes.

6.3 Validation of Our Distinguishability Degree Evaluation Technique

In this section, we validate through analysis the efficiency of our distinguishability degree metric. In our simulated example, our network is composed of four real Nodes $(1, 2, 3$ and $4)$ and a Sybil one (5 created by the Node 4). We study hereafter the behavior of the distinguishability matrix on the Node 1 throughout the simulation duration.

Figure 16 shows the balanced geometric mean M evaluated by the Node 1 between the Nodes $2, 3, 4$ and 5. We note that Nodes 2 and 3 are distinct, the distinguishability degrees between the node 2 and 4 and the Nodes 2 and 5 are almost identical, and the distinguishability degree between Nodes 4 and 5 tends towards 0% since the second measurement. The Node 1 can thus directly conclude that Nodes 4 and 5 are malicious and Sybil ones.

7 Conclusions and Future Work

The establishment of secure communications within wireless networks remain a key issue because of the vulnerabilities of such environment (mobility, dynamicity, wireless links, lack of infrastructure, ...) [2]. Indeed, wireless networks are subject to malicious attacks, such as Sybil node attack : a malicious node creates Sybil entities in the network, able to inject fault and malicious messages. Such attack is very compromising especially within VANET, where the number of nodes and the communication overhead are significant. We presented in this paper a Sybil detection approach based essentially on received signal strength variations. Our approach allows a node to verify the authenticity of nodes with which it is communicating, via two complementary techniques: the verification of their geographical localizations and the evaluation of their distinguishability degree. We demonstrate through geometrical analysis that verifier nodes can determine precisely which entities are Sybil within VANET, and which malicious nodes create them.

To validate our contributions, we carried out analysis, simulations and real tests. We showed that for our localization technique, the choice of the parameters is important to minimize the computed localization error. We showed also that our distinguishability degree metric is significant and efficient to detect Sybil nodes within VANET. Finally, the results of the real tests that we carried out are very promising and validate the real applicability of our contributions.

As future work, we plan to elaborate a distributed trust management architecture within VANET, allowing network nodes to cooperate and collaborate in order to ensure secure communications between them, while detecting Sybil and malicious entities.

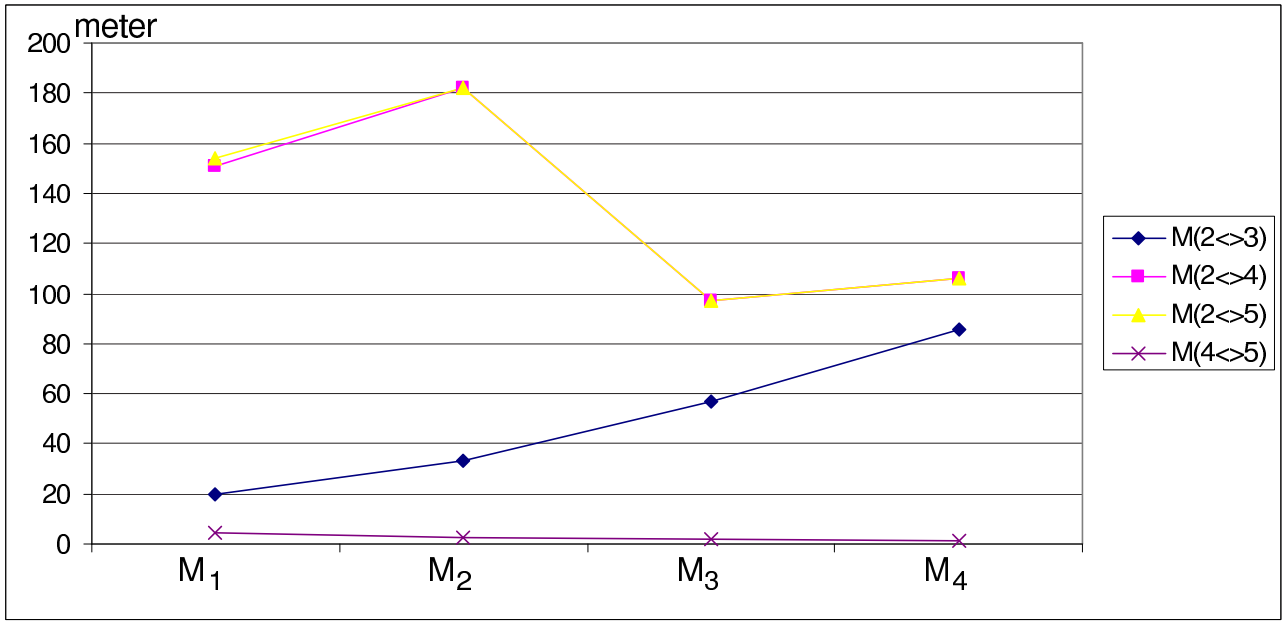

Figure 16: Distinguishability degree illustration

References

[1] J. Blum, and A. Eskandarian, "The threat of intelligent collisions," *IT Professional*, vol. 6, no. 1, pp. 24-29, Jan. Feb. 2004.

[2] M. S. Bouassida, I. Chrisment, and O. Festor, "Group key management within ad hoc networks," *International Journal of Network Security*, vol. 6, no. 1, pp. 67-79, 2008.

[3] M. S. Bouassida, and M. Shawky, "Relative nodes localization in wireless networks using received strength signal variations," *International Conference on Wireless Information Networks and Systems*, Accepted for publication, 2008.

[4] S. Brands, and D. Chaum, "Distance-bounding protocols," *Workshop on the Theory and Application of Cryptographic Techniques on Advances in Cryptology*, pp. 344-359, 1994.

[5] S. Capkun, and JP. Hubaux, "Secure positioning of wireless devices with application to sensor networks," *IEEE INFOCOM 2005*, vol. 3, pp. 1917-1928, 2005.

[6] J. Douceur, "The sybil attack," *First International Workshop on Peer-to-Peer Systems*, pp. 251-260, Mar. 2002.

[7] H. T. Friis, "A note on a simple transmission formula," *Proceedings of the IRE*, vol. 34, pp. 254- 256, 1946.

[8] P. Golle, D. Greene, and J. Staddon, "Detecting and correcting malicious data in VANETS," *Proceedings of the 1st ACM International Workshop on Vehicular Ad Hoc Networks*, pp. 29-37, Oct. 2004.

[9] R. C. Merkle, "Secure communications over insecure channels," *Communications of the ACM*, vol. 21, no. 4, pp. 294-299, Apr. 1978.

[10] J. Newsome, E. Shi, D. Song, and A. Perrig, "The sybil attack in sensor networks: Analysis and defenses," *Proceedings of the Third International Symposium on Information Processing in Sensor Networks*, pp. 259-268, 2004.

[11] W. Pires, T. D. P. Figueiredo, H. C. Wong, and A. Loureiro, "Malicious node detection in wireless sensor networks," *Proceedings of the 8th IEEE International Parallel and Distributed Processing Symposium*, pp. 24, 2004.

[12] M. Raya and J. P. Hubaux, "The security of vehicular ad hoc networks," *Proceedings of the 3rd ACM Workshop on Security of Ad Hoc and Sensor Networks*, pp. 11-21, 2005.

[13] M. Raya and J. P. Hubaux, "Securing vehicular ad hoc networks," *Journal of Computer Security, Special Issue on Security of Ad Hoc and Sensor Networks*, vol. 15, no. 1, pp. 39-68, 2007.

[14] M. Raya, P. Papadimitratos, and J. P. Hubaux, "Securing vehicular communications," *IEEE Wireless Communications Magazine, Special Issue on Inter-Vehicular Communications*, vol. 13, no. 5, pp. 8-15, 2006.

[15] T. Suen, and A. Yasinsac, "Ad hoc network security: Peer identication and authentication using signal properties," *Systems, Man and Cybernetics (SMC) Information Assurance Workshop*, pp. 432-433, June 2005.

[16] B. Xiao, B. Yu, and C. Gao, "Detection and localization of sybil nodes in VANETs," *ACM/SIGMOBILE Workshop on Dependability Issues in Wireless Ad Hoc Networks and Sensor Networks*, pp. 1-8, 2006.

Bertrand Ducourthial received the Ph.D. degree in Computer Science from Paris Sud University in 1999. He

then joined the University of Technology of Compiègne (UTC), and the Heudiasyc lab (UMR CNRS 6599), France, were he is in charge of the computer systems and networks teaching specialty. His research work deals with highly dynamic ad hoc networks (such as inter-vehicles networks): networking, distributed algorithms, security, software architecture.... Dr. Ducourthial was advisor for several masters and PhD theses. He received the *Habilitation à diriger des recherches* in 2005. He is regularly involved in program committees of conferences. His work is funded by industrial, regional, national and European grants.

Gilles Guette received its M.Sc. in computer science and Ph.D. from the University of Rennes with honors in 2002 and 2005 respectively. After a post doctoral position at the University of Compiègne, working on the security of vehicular ad hoc network, he is currently associate professor at the University of Rennes. His current research interests include network security, key management and mobility.

Mohamed Salah Bouassida is a CNRS researcher at HEUDIASYC laboratory in France. He has a Ph.D. and master degree from Henry Poincaré University, Nancy France, within the MADYNES research team in the LORIA laboratory (in 2006 and 2003 respectively). His main research interests are localization within wireless networks, security services of group communications in the context of ad hoc networks, establishment of group key management protocols within MANETs and congestion control within wireless networks.

Mohamed Shawky received his MSc and PhD respectively in 1989 and 1992 on distributed computing from University de Technologie de Compine, France. He is currently associate professor, HDR, at the Computer Science Department, with Heudiasyc Laboratory. His research interests include real-time embedded computing for highly mobile networks. He is responsible of several European, industrial, national and regional research projects. He is co-ordinating the workgroup "embedded diagnostic" with the research cluster System@tic and the workgroup "Robust Design" of Eicose European Institute. He is also on the list of European Union experts for INCO programs.