Expectation maximization algorithm for jamming detection in vehicular networks

Χρήση του αλγορίθμου μεγιστοποίησης προσδοκίας για ανίχνευση απειλών παρεμβολών σε δίκτυα αυτοκινήτων

Thesis Report

Author: Pavlos Evangelatos
Supervisor: Antonios Argyriou

October 2017

University of Thessaly
Department of Electrical and Computer
Engineering

Abstract

The scope of this thesis is the detection of interference in Vehicular ad-hoc networks (VANETs) just utilizing the received power of signals, stochastic processes and the machine learning algorithm Expectation Maximization.The objectives of the report are the simulation for different possible profiles and types of noise attacks (jamming) at the physical layer.VANETs and communication systems between vehicles including threats and hazards on them, is a groundbreaking field whose research is current and expected to be at the peak over the following years as the launching of intelligent driving systems, that limit human intervention and attribute higher safety effectiveness, will be prepared.The results and estimations of detection of malicious attacks using the algorithm are considered satisfactory and it is concluded that it is feasible to be applied.

ΠΕΡΙΛΗΨΗ

Η διπλωματική αυτή αποσκοπεί στην ανίχνευση παρεμβολής σε δίκτυα αυτοκινήτων .Χρησιμοποιώντας απλά την λαμβανόμενη ισχύ των σημάτων, στοχαστικές διεργασίες και τον machine learning αλγόριθμο Μεγιστοποίησης Προσδοκίας, επιχειρείται η προσομοίωση για διάφορα προφίλ και τύπους επιθέσεων θορύβου στο φυσικό επίπεδο.Τα ασύρματα δίκτυα επικοινωνίας μεταξύ κινούμενων οχημάτων και οι απειλές και κίνδυνοι που αντιμετωπίζουν είναι ένας πρωτοποριακός τομέας, η έρευνα πάνω στον οποίο βρίσκεται σε πλήρη εξέλιξη και αναμένεται να βρεθεί στην αιχμή τα επόμενα χρόνια καθώς θα προετοιμάζεται η έλευση ¨έξυπνων’ και αυτοματοποιημένων συστημάτων οδήγησης που θα ελαχιστοποιούν την ανθρώπινη παρέμβαση και θα έχουν καλύτερες επιδόσεις στην ασφάλεια.Τα αποτελέσματα και οι εκτιμήσεις της ανίχνευσης κακόβουλων επιθέσεων με τη χρήση του αλγορίθμου κρίνονται ικανοποιητικά και συμπεραίνεται ότι μπορεί να χρησιμοποιηθεί στην πράξη.

Preface-Acknowledgements

This report is submitted in fulfillment of the requirements for acquiring the degree in computer engineering. This research was conducted during the Spring semester and the summer of 2017 in the city of Thessaloniki, under the guidance of my supervisor, Assistant Professor Antonios Argyriou from the Department of Electrical and Computer Engineering.
Hereby, I would like to express my gratitude to my supervisor Dr. Argyriou for the knowledge that he provided and his willingness to help in every problem that I encountered with.His guidance allowed me to distinguish the targets of this research and acquire a solid understanding of the scientific concepts.
Furthermore, i would like to thank my parents, Effie and Vangelis, for doing everything they could in order for me to fulfill my dreams and my comrade Nøla who was always there.

Contents

List of Figures

List of Tables

1 Theoretical background

1.1 Platoon survey

According to the European Commission 25,500 people were killed on the roads in 2016. For every death on Europe's roads there are an estimated 4 permanently disabling injuries such as damage to the brain or spinal cord, 8 serious injuries and 50 minor injuries.Autonomous cars are evolving intelligent vehicles navigating without human input.They obtain formations in rows (Platoons) to circulate at high speeds through highways.Their occupants will delegate the driving tasks to a fully automated system that will use ways to share information like position, velocity and tension, in order to keep a small distance between cars and move fast in a very efficient way, leading to increase of the capacity of the roads.This concept is expected to attribute lots of benefits.The energy consumption (fuel combustion or electrical energy stored) and the consequent environmental impact will be lower due to speed fluctuation restraints.Human factors will not need to operate any manipulations during the platooning stage and manual driving will be happening into neighbourhoods and dense populated areas.Furthermore, circulating in a tight platoon implies reduced air resistance.If automation be applied in parking, a huge saving of space is going to be gained.Generally, as an outcome, a much more efficient use of the current highways with reduced traffic jams and congestion will occur.Each vehicle will be equipped with on board technology (ultrasonic sensors, Global Positioning System (GPS), cameras, and micro controllers) to measure the distance with its predecessor, the speed difference (Autonomous cruise control) and side intervals among objects in regard with the lanes of the motorway as long as the density of the circulation on the neighbouring lanes [1].The itinerary will be submitted in advance but alterations one the route are going to be feasible anytime and be handled with the optimal and safest way.As a consequence constant road particularities like junctions,interchanges and toll stations will be expected.Safe procedures for vehicles to join or leave the platoon will be anticipated.Vehicles will be up to date about critical data from the infrastructure relative with current road and road surface conditions (such as slippery road), forthcoming traffic lights,forward rules that demand speed modification, exact position of obstacles , priority assignments for emergency vehicles and temporary lane merging or detours due to roadworks.Roadside motion sensors and cameras installed along the highway, provide supervision and instant notification to the management application that controls the highway dynamically.Infrastructure of the motorway also gets shared data from vehicles, either forwarded messages by other nodes obtained by communication channels or measurements collected by their equipment.As an example, awareness of ice on the driveway may be achieved with the use of infrared radiation that is related with thermal radiation.All objects emit radiation proportional to the fourth power of the temperature of the object.The warmer the object, the more radiation emitted.The wavelength band in which the object releases radiation energy is dependent on the object's composition, e.g. radiation emitted by a gas is decided by thin spectral lines or band spectra.By measuring the emissivity, it becomes district weather the surface is asphalt or ice.Such data gathered by the vehicles could be shared.In case of a high-risk emergency situation like animal wandering, an irregular appearance of pedestrian or bicycle, a sudden deceleration of a vehicle caused by breakdown etc, a possible collapse of some of the system's component's and collisions despite the prevention policies, alert messages emerge.Based on radio tranceivers along road sections, a procedure recognizes hazards and broadcasts assistance to the involving with the incident vehicles, plus warnings and directions to the following ones.The development of algorithms that can resolve the problem of the intersections is necessary for the users, so they can share relevance data between different platoons or vehicles that approaching the intersection.

1.2 VANETs

To ensure the smallest possible space between vehicles, it is necessary a robust and stable communication between participants in the platoon.Vehicular ad-hoc networks (VANETs) have recently attracted the interest of researchers and industry due to their potential to improve road safety and capacity.There are two main protocol stacks for vehicular communication systems, one supporting exchange of data among vehicles (V2V communications) and the other one concerns between vehicles and the roadside infrastructure (V2I/I2V).These two families of standards correspond to the IEEE Wireless Access in Vehicular Environments (WAVE), adopted in the United States, and the ETSI ITS-G5 in Europe.At the physical and medium access control layers, both protocol stacks rely on the IEEE 802.11p standard, an amendment to the IEEE 802.11a Wi-Fi technology.In comparison with the typical wi-fi operation channel bandwidth was reduced from 20Mhz to 10Mhz in order to mitigate the effects of multipath propagation like intersymbol interference(ISI) and Doppler shift.As a consequence bit rate was modified to the half, from 3 up to 27 Mbps instead of 6 up to 54 Mbps.The IEEE 802.11p PHY is based on the Orthogonal Frequency Division Multiplexing (OFDM) policy.The IEEE 802.11p MAC sublayer is the IEEE 802.11 Distributed Coordination Function (DCF), which is based on the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) scheme, protocol where a transmitter attempts to detect the presence of a carrier signal from another node before attempting to transmit.If a carrier is sensed, the node waits for the transmission being in progress to end before initiating its own transmission of the intended message.The IEEE 802.11p MAC mechanism has been improved to reduce the process time of the algorithm.In addition, there are no authentication nor association in the MAC layer.When a vehicle want to join the network, It receives a WAVE non-IP message to configure the vehicle to join the Basic Service Set (BSS), avoiding the overhead introduced by the registration and authentication procedures, commonly present in wireless local area networks.In order to guarantee that vehicular communications will not suffer from any type of interference from unlicensed devices, the Federal Communications Commission (FCC) in the United States and the European Conference of Postal and Telecommunications Administrations (CEPT) in Europe, allocated a dedicated spectrum band at 5.9 GHz. In America, a bandwidth of 75 MHz was reserved, while in Europe only 50 MHz were assigned.This spectrum was divided into smaller 10 MHz wide channels in the American case to better cope with multipath fading and a 5 MHz guard band at the low end was also included. As a result, there are 7 different channels for IEEE WAVE operation and 5 for the case of ETSI ITS-G5.In Europe, 30 MHz (3 channels) are reserved for road safety in the (ITS-G5A) band and 20 MHz are assigned for general purpose (ITS-G5B) band.In fact, the European tolling systems operating in the 5.8 GHz frequency band and wireless communication systems operating near the 5.9 GHz frequency band cause serious problems to the performance of vehicular networks.Adjacent Channel Interference (ACI) can severely impair the integrity of the messages received by a radio unit, whenever simultaneous communications occur in the nearby channels.This phenomenon causes packet loss and if consolidated, results in large values of Packet Error Rate (PER).A realistic study proved that the IEEE 802.11p technology's supported communication range between vehicles or vehicles and infrastructure, can reach in a highway scenario up to 880 metres for the line of sight (LOS) and 58 to 230 metres in the none line of sight (NLOS) [2].A mobile peer receives signals either from the predecessor vehicles or the motorway static infrastructure.All received signal with strength below a certain threshold are considered ambient and generated from the presence and dissemination between other nodes so they are included in the noise level. Multihop broadcast protocols are essential to reduce collision due to congestion and waiting time of emergency messages.On the receiver's MAC sublayer, only packets sent by the sender can pass the Cyclic Redundancy Check (CRC).CRC is an error detecting code.the Packet Delivery Ratio (PDR) is measured by using the ratio of the number of packets passing the CRC check with respect to the number of total received packets (or preambles).The successfully received packets should be sent either by the sender or by the compromised vehicle[3].At the sender side, the PDR can be calculated by keeping track of how many acknowledgements it receives from the receiver.

1.3 Jamming

The greatest challenge about VANETs is the security issue.Trustworthiness is vital when it comes to messages critical to life that is not permitted to be dropped or modified by a threat.Denial of service (DoS) attacks to VANET can aim at different layers with variations of three basic techniques: buffer overflow at Network Layer and above, protocol violation at Medium Access (MAC) sublayer (data link layer) and signal inference by a radio emitter, called jammer, at Physical (PHY) layer [4].If too many packets are buffered in the MAC layer, the newly arrived packets will be dropped.It is also possible that a packet stays in the MAC layer for too long, resulting in a timeout and packets being discarded.A jammer transmits electromagnetic energy to clash or block legitimate communications on the wireless medium.Moreover, given that 802.11 operates on relatively few frequency bands, multiple jamming devices operating on different channels can significantly hurt the performance in spite of using frequency hopping.The dominant way to defend against a jamming attack is retreat strategy: Channel Surfing to switch on another channel when the current frequency is blocked and Spatial Retreat to move on another location if the area involves interference.Jamming cannot be avoided by regular security mechanisms such as authentication, digital certificates, or encryption, because the jammer is often disregarding higher layers,focusing on the lower ones[5].Jamming attack models can be used by adversaries to defuse the operation and reliability of a wireless network[6].There are several jammer profiles and strategies:

- flat/constant jamming: jammers characterized by brute force considered the most disruptive as they indiscriminately affect all ongoing communication.All blocks are jammed with equal power.The power spectral density of the signal broadcast is flat across the entire bandwidth.This is achieved by transmitting continuously radio signal that represents random bits and does not follow any MAC procedure.
- deceptive jamming: jammers transmit semi-valid packets.In this case the packet header is valid but the payload is useless.Instead of sending out random bits, the deceptive jammer constantly injects regular packets to the channel without any time interval between subsequent packet transmissions.As a result, a legitimate peer will be mocked, forced into repeated backoffs believing there is a valid transmittion in progress.
- random jamming: jammers that interchange periodically between trying to impede the channel and sleeping mode.During the first jamming period, the jammer emits for a random period of time (it can behave either like a constant jammer or a deceptive jammer) while in the sleeping mode, the jammer is offline for random amount of time.
- smart/reactive jamming: protocol aware jammers able to target specific data or control packets.In this type of attack, the enemy detects the resource blocks actively used by the targeted user and allocates its power accordingly.In such case, the jammer stays quiet and senses.It decides to react only when a transmission occurs and overlaps with it, with raised power.This sort of jammer attempts to corrupt and modify the contents of the packet, invalidate its checksum so it will not be considered by the receiver.

Attackers usually prefer the rush hours when speeds are much lower and the network traffic is higher.
The figure 1.1 is taken by [7] shows the performance of a reference VANET receiver under the influence of RF jamming in the form of PDR vs SNIR curves obtained from measurements in an anechoic chamber.These curves show that a Constant Jammer is more devastating than a Reactive Jammer.While 80% PDR is achieved at 5dB power without interference, the same PDR is only achieved at 10 dB when a Reactive Jammer is active and at 22 dB when a Constant Jammer is present:

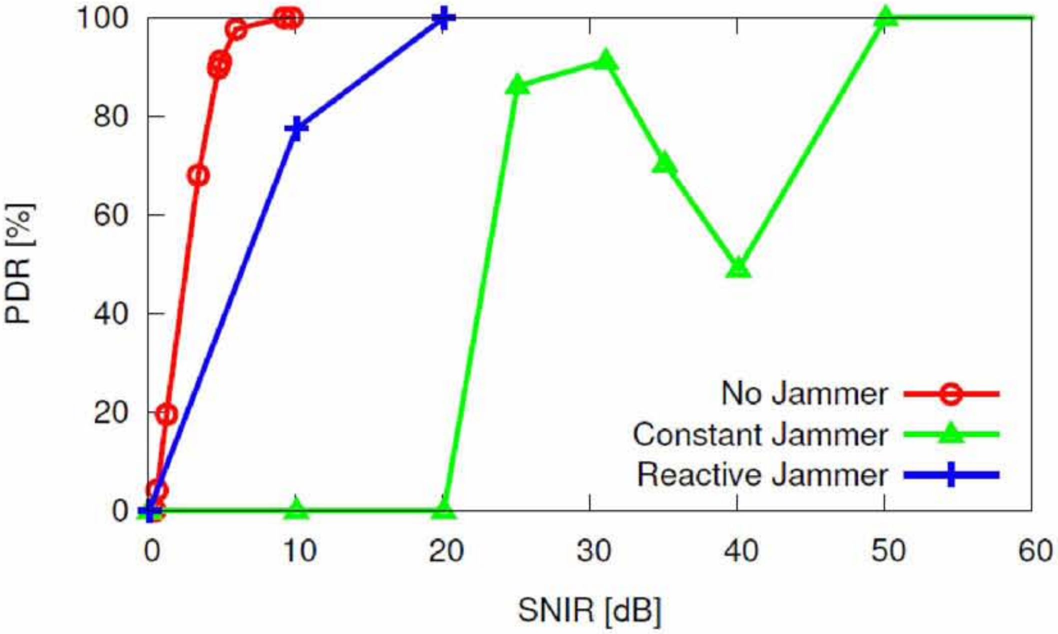

Figure 1.1: diagram of measurements in an anechoic chamber

2 RPMs

Most widely used stochastic Radio Propagation Models (e.g., Free Space, TwoRay-Ground Reflection, Rayleigh Fading model, Ricean Fading model, Shadowing model, Log-distance Path Loss and mix ones) rely on the overall statistical properties of the environment.Free Space model is not considering obstacles in city environments and the received signal power is based on three factors: the sender-receiver distance, antenna gains, and the transmitted power.Two-Ray Ground model demonstrates better performance than the Free Space model.Received signal strength can be predicted from long distance.However, two-ray is not concerned with the height and width of the nodes.It only assumes the received energy, which is the sum of direct LOS path and reflected path from the ground.In fact, there are different sizes of vehicles (i.e., cars, trucks, buses, and vans) on the roads.In Rayleigh Fading mode there is one dominant path and multiple indirect signals.The random multipath components are added to the LOS and can be seen as a Direct Current (DC) component in the random multipath in Rayleigh distribution.Similarly, the Ricean model considers indirect paths between the sender and the receiver.It only focuses on a single exact path and multiple reflected signals.Rayleigh and Ricean Fading are considering fast fading and caused by scattering while slow fading follows log-normal shadowing and occurs due to reflections of hill, building and obstacle.In Shadowing model a Gaussian random variable is added to the path loss to account for environmental influences.The radio signals are set to some particular values, which make it not suitable for real urban environments.Radio Propagation Model with Obstacles models obstacles, RPMO but when there are no obStacles, RPMO behaves like Two-ray Ground, so distance attenuation is not taken into account.Mahajan model behaves like Two-ray Ground, adding the influence of obstacles and the distance attenuation,but it is designed for the 802.11b environment.The LOS between two communicating vehicles (cars) runs a high risk of interruption due to the presence of large vehicles such as buses and trucks.To develop a realistic radio propagation model for highly dynamic VANET is computationally challenging, and moving obstacles makes them more complicated.

The radio signals in VANETs can potentially be obstructed by different radio obstacles.The wavelength of the radio signals in 5.9 GHz frequency band is approximately 5 cm; therefore, they have relatively less penetrating power in comparison to technologies such as GSM that typically operates in 1800 MHz frequency band.In other words, radio signals in the 5.9 GHz frequency band are obstructed by static objects (e.g. buildings, dense vegetation, and advertising boards) and moving objects (e.g. large buses, trailers and delivery trucks that impede radio signals) present in the VANETs environment.Thus, vehicular communication built on IEEE 802.11p standards suffers from a relatively small effective coverage communication area, and potential disruption that results in signal attenuation (e.g. due to radio obstacles) The tunnel geometry, electromagnetic properties of the tunnel's material, antenna characteristics and radio obstacles also affect the radio propagation in tunnels.

2.1 Free Space Propagation Model

A large scale propagation model assumes that there is only the LOS path between the transmitter and the receiver.The received power Pr at a distance d from the transmitter is given by Friis Equation:

$$Pr(d) = P_t G_t G_r \frac{\lambda^2}{4\pi^2 d^2 L} \tag{2.1}$$

where P_t is the transmitted power, G_t and G$_r$ are the transmitter and receiver antenna gains, L is the system loss and λ is the wave length in meters.

2.2 Shadowing Model

The shadowing model consists of two parts.The first part is the path loss component which is used to predict the received power at distance d from a known reference power at distance do.The second part is the log-normal shadowing which reflects the variations of the received power at certain distance d from the transmitter.It is a log-normal distribution or Gaussian distribution if measured in dB.Therefore the overall shadowing model is represented as:

$$\begin{aligned}
P_{L(dB)} &= 10log_{10}\frac{P_t}{P_r} \\
\text{or}\ \ P_{r(dB)} &= 10log_{10}P_r = 10log_{10}P_t - P_{L(dB)} \\
\text{where}\ \ P_L d0 \rightarrow d_{(dB)} &= P_L d0_{(dB)} + 10nlog_{10}\frac{d}{d0} + \chi_\sigma \\
\text{so}\ \ P_{r(dB)} &= 10log_{10}P_t - P_L d0_{(dB)} - 10nlog_{10}\frac{d}{d0} - \chi_\sigma \qquad (1) \\
\text{but}\ \ P_L d0_{(dB)} &= 10log_{10}P_t - 10log_{10}P_r(d0) \qquad (2) \\
\text{from (1),(2):}\ \ P_{r(dB)} &= 10log_{10}P_r = 10log_{10}P_r(d0) - 10nlog_{10}\frac{d}{d0} - \chi_\sigma
\end{aligned} \tag{2.2}$$

we can obtain the recieved power at reference distance d0 ($P_r d0$) by the Friis equation. n is the path loss exponent witch fluctutates from 2.7 to 3.5 for urban areas. χ_σ is the PDF of the Normal(Gaussian) distribution with zero mean and σ_{dB} value from 4 to 12 dB for outdoor environment

$$\chi_\sigma = N(y_i; 0, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y_i^2}{2\sigma^2}}$$

The Free Space and Two-Ray models are deterministic radio propagation models.They assume a successful reception of the signal if the received signal strength (RSS) is greater than a threshold.This means that their communication range is an ideal circle and they always determine the same RSS for the same distance.While in reality and because of the multipath propagation effects, the RSS is a random variable makes the successful detection of the signal not certain.The shadowing, Rayleigh, Ricean and Nakagami are probabilistic propagation models and their successful reception of the signal is a decreasing function of the distance.

3 EM (estimation-maximization)

Because of high complexity of Maximum-Likelihood (ML) estimate, the EM iterative algorithm that provides an attractive alternative to computing ML is proposed.In this case where the information of a priori estimated parameters is not given, and the observe data is incomplete, an easy-to-use iterative algorithm to compute the parameters of ML is proposed, and this algorithm guarantees convergence to get the maximum estimate.Expectation-maximization(EM) algorithm is an effective machine learning method when probability distributions of data samples are modeled by a Gaussian mixture model (GMM).EM algorithm is used for clustering and probability density estimation of a given data set.Recent applications of clustering and density estimation often require low-power consumption such as sensor networks and mobile terminals.The EM algorithm has been used in a lot of applications including system identification, array processing, medical imaging, and time series analysis.It works with a complete data specification, and iterates between estimating the log likelihood of the complete data using the observed (incomplete) data and the current parameter estimates(E-step), and maximizing the estimated log-likelihood function to obtain the updated parameter estimates (M-step).The algorithm converges, under certain regularity conditions, to a stationary point of the observed log likelihood function, where each iteration cycle increases the likelihood of the estimated parameters.The EM algorithm has disadvantage in the computation time due to repeat operations until convergence when the operations are adopted to a large data set, e.g. images or sounds.Although there are some fast computational methods of EM algorithm, e.g. parallel computing or data segmentation.

Step 1

Initialize mixture ratio α_n, mean vector μ_n, covariant matrix Σ_n.

Step 2 (E step)

Calculate the responsibility

$$\gamma(y_n, m) = \frac{\alpha_n N(x_m; \mu_n, \Sigma_n)}{\sum_{j=0}^{N-1} \alpha_j N(x_m; \mu_j, \Sigma_j)} \tag{3.1}$$

where n is an identifier of distribution, m is an identifier of input data.The conditional distributions within each class are Gaussian.

Step 3 (M step)

Update the parameters by

$$\begin{aligned}
\mu_n^{new} &= \frac{1}{M_n}\sum_{m=0}^{M-1}\gamma(y_n,m)x_m && \text{, for the means,}\\
\sum_n^{new} &= \frac{1}{M_n}\sum_{m=0}^{M-1}\gamma(y_n,m)(x_m-\mu_n^{new})(x_m-\mu_n^{new})^T && \text{, for the variances}\\
\alpha_n^{new} &= \frac{M_n}{M} && \text{, for the probabilities (ratios)}
\end{aligned} \tag{3.2}$$

where M is the number of samples and $M_n = \sum_{m=0}^{M-1}\gamma(y_n,m)$.

Step 4

Repeat Step 2 and Step 3 until the parameters convergence.The simulation of the jamming detection model regards two possible classes.These two labels are designated to have, different unknown means μ and variances although initial values are assumed.The mixture ratio terms are equivalent to probabilities.According to Bernoulli distribution if a label's α is q ,the opposite's is 1-q [8],[9].

4 Implementation

The implementation was elaborated, is a mathematical research of different jamming scenarios under the above RPM's.MATLAB was used for the process and computational results of the methods.The basic structure is set by a platoon of vehicles that keep constant distance between them and they support VANET's communication architecture and a range of applications.The usage of EM algorithm intended to achieve classification of the two specific cases.The first one that there is no occurrence of interference created by an external malicious factor and the second one that the referential existence affects the radio communication.The outer factor essentially attempts a jamming attack by generating noise.The selection of this algorithm is based to the manipulation of cases when the data are incomplete or even unknown.If hidden unobserved labels C_i were accessible, then the estimation for the parameters would be easy via Maximum-Likelihood.Getting all the points for which $C_i = N$ (noise class) and using those to estimate μ and σ^2.Moreover in our experimental environment there is no access to any legitimate information about packets and data stream transmitted or received correctly by the affiliated vehicles.Subsequently there is no metric or ratio of error detection, error correction, flow control like PDR and no outcome about the quality of the communication channel.

The received signal strength indicator (RSSI) is a measurement of the power present in a received radio signal.IEEE 802.11 devices often make the measurement available to users.In our simulation the power level is evaluated by consideration of geometry, velocity, distance, transmit power and RPM's as mathematical models that represent how the signals decay or attenuate.The signal to interference plus noise ratio (SINR) is defined as the power of a certain signal of interest divided by the sum of the interference power (from all the other interfering signals) and the power of some background noise.It is supposed that adjacent vehicles in platoon are network peers that have established sessions after having their authentication verified.The jammer does not have rights to this access mechanism but operates in order to disrupt.Therefore,the power received from the jammer is additive to the Gaussian noise w.During the simulation an adherence with the maintenance the transmit power in the same level with the power the jammer emits is observed.Every vehicle sets a session with the following vehicle.In a supposed system where every vehicle's emitted power is adjusted to 100mW (-10dB) in environment without path loss, and the interval between the vehicles is 5 metres, the received power from the vehicle after next (10 metres away) 5.69mW and from the next to it (15 metres away) is 2.5mW including with the thermal noise in w which is set as one by ten of the transmit power (10mW or -20dB).Of course in real environments this value fluctuates.The real aim of this research is to overcome a deterministic approach that classifies as jamming case every occasion the level of the noise surpasses a prefixed value.

The noise plus interference The SINR is the indication utilized by the EM algorithm.Random variable $SINR_i \subseteq R$ is observed as measurement for each district time of the experiments.It is assumed that there is an unobserved label $C_i \subseteq \{N, W\}$ referring to the class of existence or absence of noise.The mixing coefficients are the probabilities of the two cases π_N and π_W are always initialized as 0.5 and 1 - 0.5 = 0.5 .Both variances σ^2 for the two cases, varN (with noise) and varW (without noise) are initially equal to 1. The conditional distributions within each class are Gaussian:

$$pSINR_i|C_i(sinr_i|c_i) = \prod_c N(sinr_i;\mu_c,\sigma^2)^{(c_i=c)} \tag{4.1}$$

The posterior probability for $C_i = W$ is derived as

$$\gamma(y_n,m) = \frac{\pi_W N(sinr_i;\mu_W,\sigma^2)}{\pi_W N(sinr_i;\mu_W,\sigma^2) + \pi_N N(sinr_i;\mu_N,\sigma^2)} = qc_i(W) \tag{4.2}$$

and similarly is obtained the one for the noise case.
This is an unsupervised learning problem.It is not intended the observation the noise/without noise labels for the data, but the acknowledgement of parameters based on those labels.The expected values(mean) μ_W (without noise) and μ_N (with noise) are parameters initialized somehow and going to be estimated.

- μ_W, μ_N are given primary values to begin solving the hidden variables $(qc_i(C_i))$.This is the E step.
- the acquired hidden variables are now set and via the posterior distribution and optimize the parameters (means).This is the M step.

This process is repetitive and during the iterations, except from the estimation of means and variances for the noise an no noise classes, the optimization of the probabilities of the two cases happening is attempted.

5 Results

5.1 Static jammer that transmits for limited time period

In this scenario, transmitter and receiver vehicles are moving to a direction keeping constant distance 5 metres between them.A static jammer is set so as angle formed by line segment bounded by jammer and receiver at t=0.0 seconds and line of platoon direction is 50°.The initial distance between jammer and receiver is 75 metres.The jammer detects the velocity of the vehicles, therefore has the ability to compute the time needed for the receiver to reach the point where the above line segment becomes vertical to the line.Using this information, the jammer remains silent and starts radiating from 1.8s before this theoretical intersection time point until 2.2s after, and returns to silence (lasting for duration equal to the first one).
The velocity of vehicles is comes to be very significant and affects the results as 5.1 shows for means $\mu_W = 1.92797$, $\mu_N = 1.40497$:

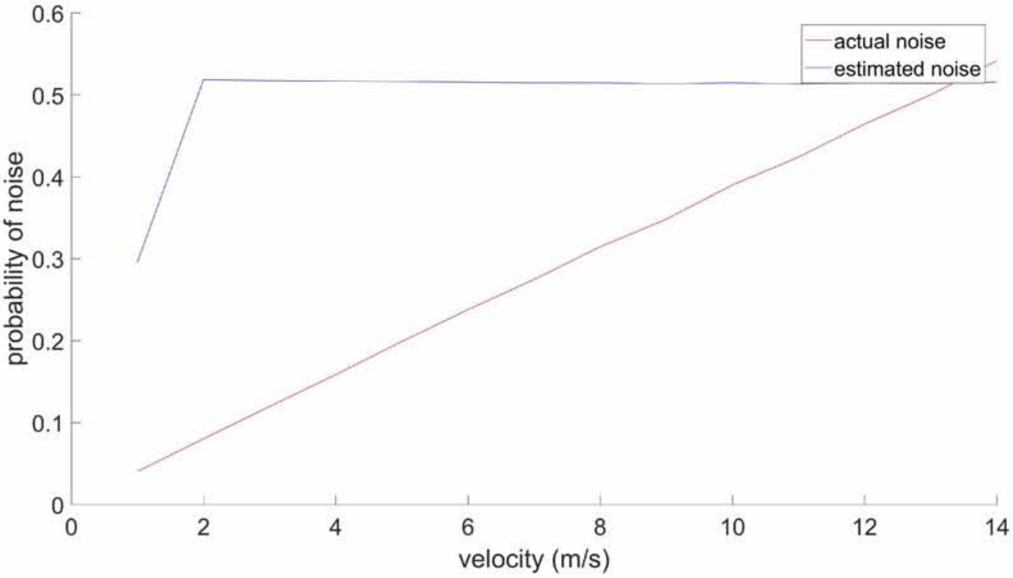

Figure 5.1: plot of static jammer estimation for variable velocity

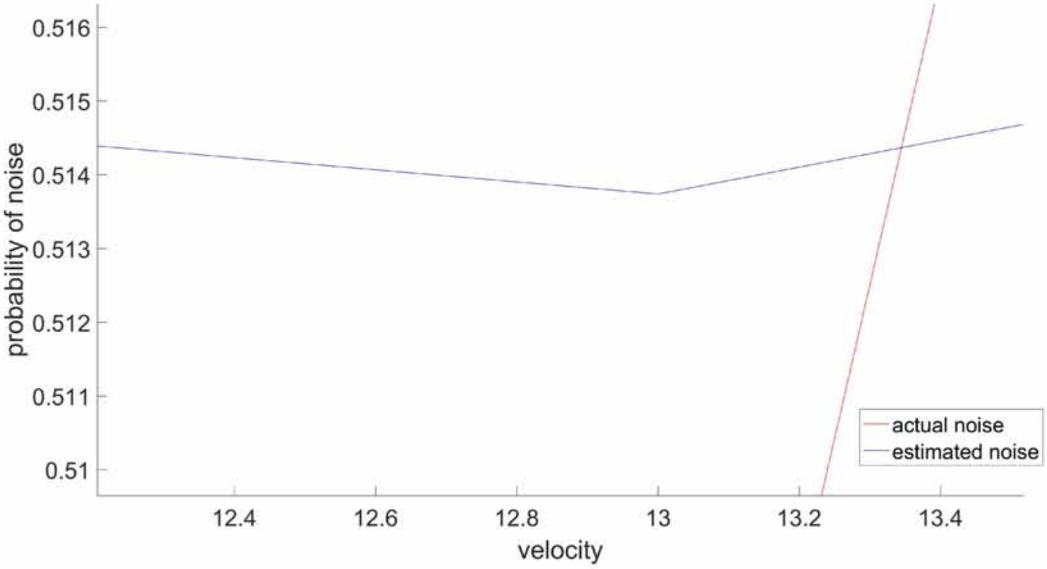

Figure 5.2: enlargement

The initial means μ_N and μ_W affect crucially too.Assuming Rx velocity equal to 10m/s, noise will really happen at 39% of counted time.Taking as input different means:

Table 5.1: indicative measurements of static jammer with respect to initial means

static jammer			
initial μ_W , μ_N	optimal μ_W , μ_N	optimal varW, varN	estimated π_N
2.50 , 1.24	2.1222 , 2.1221	3.722876 , 3.722874	51.7 %
2.34 , 1.72	2.1223 , 2.1218	3.72289 , 3.72287	50.5 %
2.18 , 1.48	2.1223 , 2.1210	3.7230 , 3.7228	50.04 %

The Mean Squared Error (MSE) of total the received power (signal strength on the receiver) as it deviates between an ideal channel to this noisy channel:

$$\frac{\sum_{i=1}^{n}(Pideal_i - Pnoisy_i)^2}{n} \simeq 0.00000174$$

.

Trying to estimate the Bit Error Rate (BER) utilizing the Error Function , in the case of QPSK modulation and considering the channel as AWGN and not a fading one,since the MSE is too low.

$$BER = 0.5 * erfc(\sqrt{E_b/N_o})$$

Considering a bitrate 15Mbps and Bandwidth of the channel 10MHz (Hz = 1/s), we get

$$E_b/N_o = Bw * SINR/Bitrate = 2/3 * SINR$$

as it is showed in5.3.

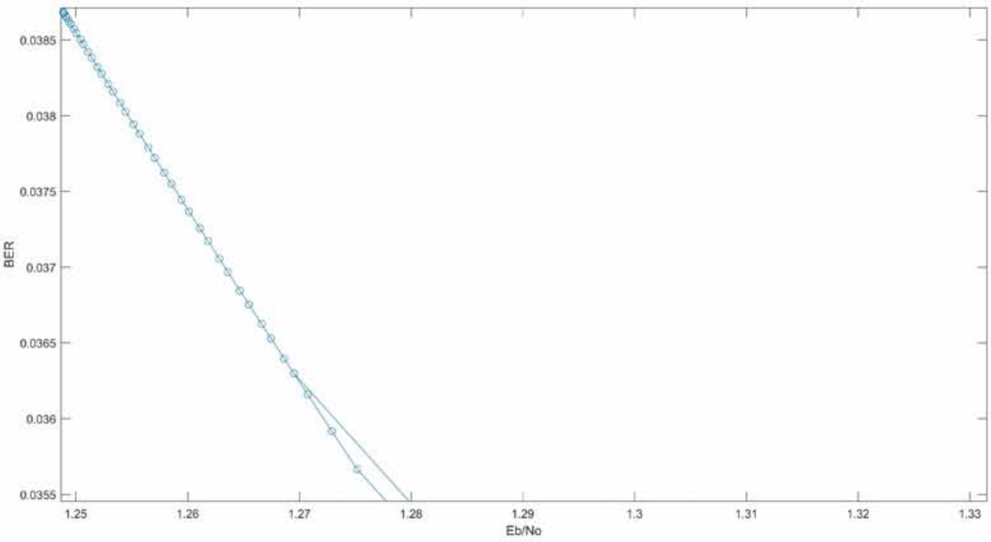

Figure 5.3: hypothetical plot of BER for static jammer

5.2 Constant jammer

in this scenario the jammer follows the receiver from behind or sideways having the same velocity (independently of the magnitude of it) as the platoon.It is assumed that the distance between the receiver and the legit transmitter is 5 metres.The jammer has only two states, originally remaining inactive and invisible until the attack mode is activated.The receiver has to distinguish sets, which include as objects ,variables that have the same common value.The distance the jammer keeps from the receiver is very critical here.

For the a case similar with D below (but with noise at 61% of its length) 5.4 and means $\mu_W = 1.92797$, $\mu_N = 1.40497$:

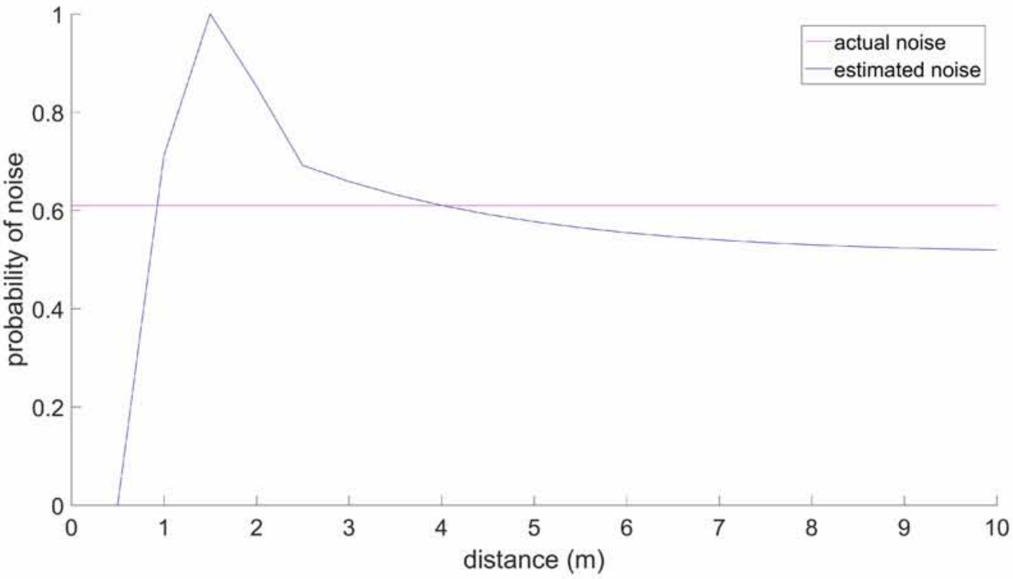

Figure 5.4: plot of constant jammer estimation for variable distance

Initial expected values also play key role again.Executing the following experiments according to free space propagation model for different μ_W and μ_N ,for jammer at 5 metres range too:

- for intermediate jamming interference that starts and ends covering a percentage 30% of time,among absention of the jammer:

Table 5.2: indicative measurements of constant jammer with respect to initial means,case A

constant jammer			
initial μ_W , μ_N	optimal μ_W , μ_N	optimal varW, varN	estimated π_N
2.34 , 1.24	1.8028 , 1.8026	0.528322 , 0.528322	54.1%
2.18 , 1.40	1.8028 , 1.8024	0.528323 , 0.528322	52.1%
2.02 , 1.72	1.8030 , 1.8019	0.528323 , 0.528323	50.7%
1.86 , 1.48	1.8070 , 1.7907	0.528467 , 0.528341	50.1%

- for inactivity at the beginning for 28% of testing time before jammer's awakening and continuous attack (82% noise circumstance):

Table 5.3: indicative measurements of constant jammer with respect to initial means,case B

constant jammer			
initial μ_W , μ_N	optimal μ_W , μ_N	optimal varW, varN	estimated π_N
2.5 , 1.24	2.18 , 0.29	1.1664 , 3.431	93.0%
1.38 , 1.0	0.6374 , 0.6373	1.049023 , 1.049023	81.5%
1.38 , 1.24	2.17 , 0.256	1.193 , 3.419	90.5%
1.22 , 1.0	0.63739 , 0.63731	1.049023 , 1.049023	74.3%

- for jamming attack designated from the beginning of the experiment, lasting for the 66% of its duration:

Table 5.4: indicative measurements of constant jammer with respect to initial means,case C

constant jammer			
initial μ_W , μ_N	optimal μ_W , μ_N	optimal varW, varN	estimated π_N
2.50 , 1.72	1.23389 , 1.23384	5.6348 , 5.6348	72.9%
2.34 , 1.48	1.23389 , 1.23383	5.6348 , 5.6348	68.1%
2.18 , 1.96	1.2339 , 1.2338	5.6348 , 5.6348	63.6%
2.02 , 1.72	1.2339 , 1.2338	5.6348 , 5.6348	59.5%
1.86 , 1.72	1.2339 , 1.2337	5.6348 , 5.6348	56.1%

- for vicious intervention from the inception of timing and again before the end for a total of 52% combined with neutrality in the middle:

Table 5.5: indicative measurements of constant jammer with respect to initial means,case D

constant jammer			
initial μ_W , μ_N	optimal μ_W , μ_N	optimal varW, varN	estimated π_N
2.34 ,1.96	1.4557 , 1.4556	0.62675 , 0.62675	61.8%
2.18 , 1.24	1.4557 , 1.4556	0.62675 , 0.62675	58.1%
2.02 , 1.24	1.4558 , 1.4555	0.62675 , 0.62675	55.0%
1.86 , 1.48	1.4558 , 1.4554	0.62675 , 0.62675	52.6%
1.54 , 1.00	1.4558 , 1.4548	0.62675 , 0.62675	50.1%

- last, if there is no noise and the only impairment at the channel is the Additive White Gaussian Noise.Some consecutive estimated probabilities (%)for different means are showed below :

Table 5.6: indicative measurements of constant jammer.case of a zero noise appearance

99.9	82.1	18.5	39.5	97.8
16.2	34.3	42.6	76.9	57.3
43.6	69.8	54.7	56.8	17.4
87.5	35.8	17.5	100.0	NaN

It is clear that a receiver could easily identify the case.

5.3 Random jammer

Supposing a condition, there is a jammer attempting to mislead the receiver.It also holds a stable velocity as the platoon remains holds distance from Receiver, equal with the transmitter-receiver's one.On the contrary to constant jammer case, this kind of

attacker awakes from the sleeping mode and and randomly interchanges radio link time slots with intermission ones.This policy aims to deception of technics detecting a vicious conjugation.
5.5 compares the level of the desired signal of the preceding vehicle of the receiver to the level of background noise.

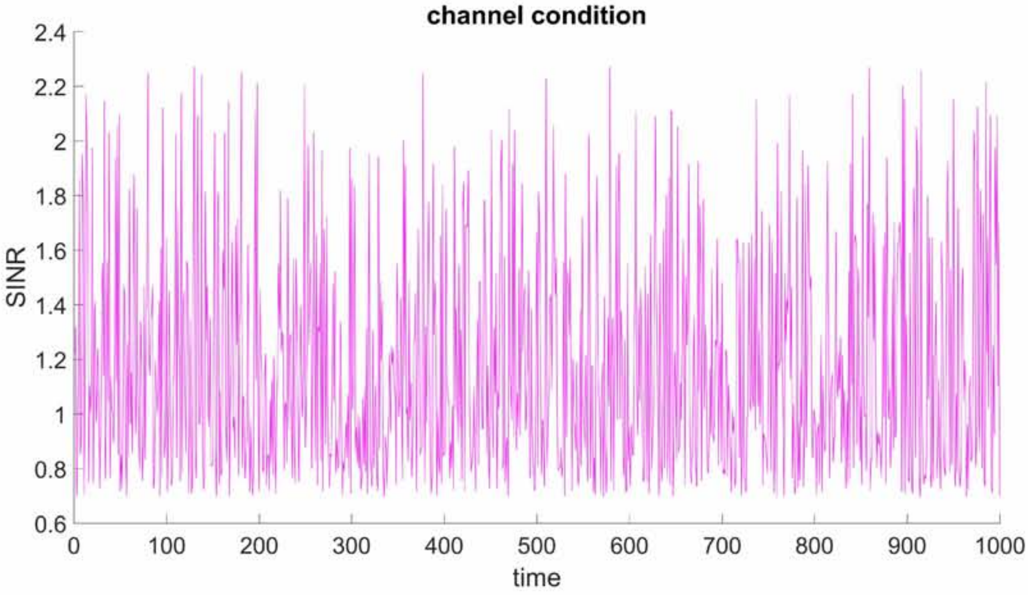

Figure 5.5: SINR affected by random jammer

For the total sum of assault moments reaching 48.6% of experiment's period and under use of the free space model and the jammer fixed 5 metres away, the behaviour of the algorithm is again depended just to the initialization of the means μ_W and μ_N according is displayed in the following table:

Table 5.7: indicative measurements of random jammer with respect to initial means

random jammer			
initial μ_W , μ_N	optimal μ_W , μ_N	optimal varW, varN	estimated π_N
2.18 , 1.24	1.16865 , 1.16863	1.626968 , 1.626968	63.0%
2.02 , 1.48	1.16865 , 1.16862	1.626968 , 1.626968	59.3%
1.86 , 1.0	1.16865 , 1.16861	1.626968 , 1.626968	56.2%
1.70 , 1.48	1.16860 , 1.16857	1.626968 , 1.626968	53.6%
1.54 , 1.24	1.1686 , 1.1684	1.626968 , 1.626968	51.8%
1.38 , 1.24	1.1687 ,1.1681	1.626968 , 1.626968	50.5%

For assault moments reaching 51.6% of the experiment, 5.6 shows how distance between jammer and reveiver affects estimation.

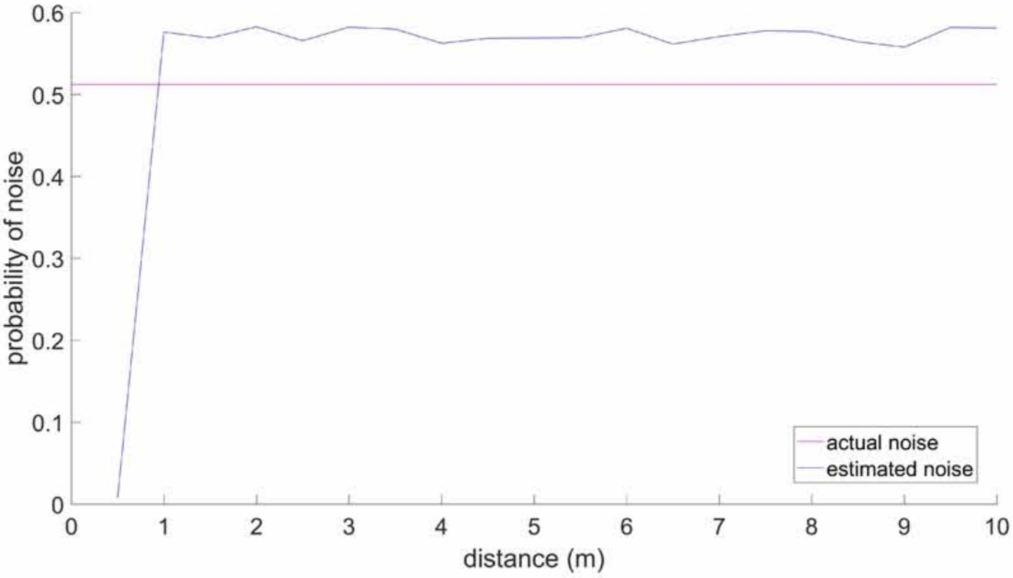

Figure 5.6: plot of random jammer estimation for variable Rx-J distance

GeThe Mean Squared Error (MSE) of total the received power (signal strength on the receiver) as it deviates between an ideal channel to this noisy channel:

$$\frac{\sum_{i=1}^{n}(Pideal_i - Pnoisy_i)^2}{n} \simeq 0.0005197$$

. Since MSE is again low the channel can be considered to be AWGN.The 5.7 displays the relation between BER and SINR per bit.

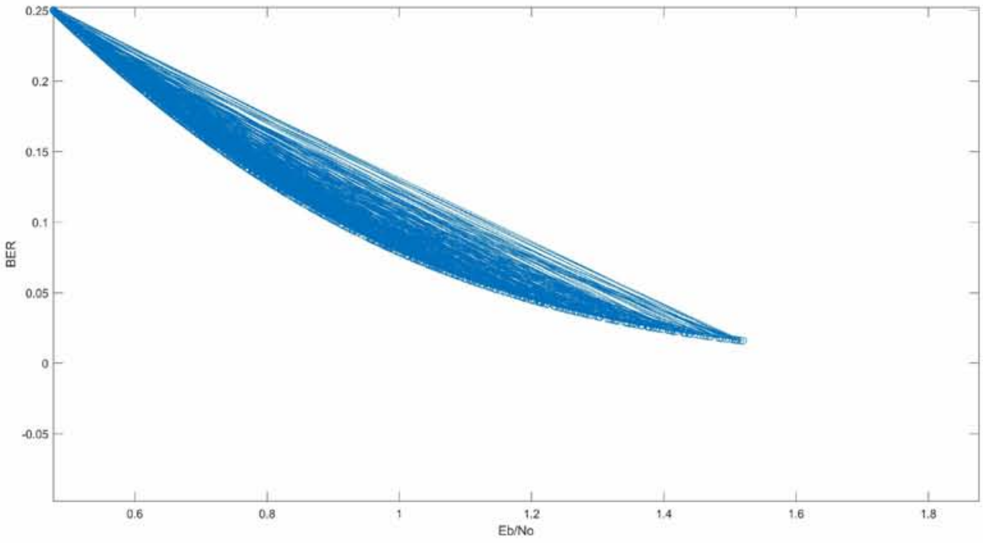

Figure 5.7: hypothetical plot of BER for random jammer

5.4 Jammer closes with the receiver

In this set-up, the jammer is originally waiting in a spot.When the target shows up, the device is capable of tracking receiver's velocity and the angle θ, formed between platoon's direction line and the one passes through receiver and jammer on a moment after a time that processes and configures the right settings.Furthermore, it chooses the angle ϕ, formed between the path it will move approaching the motorway (how long it will be) and the platoon's direction line.Then it adjusts its acceleration to obtain the velocity needed in order to meet the receiver at the point on the vertex.It stops on the limit before collision. There are two different cases depending both angles exposed in 5.8 and 5.9:

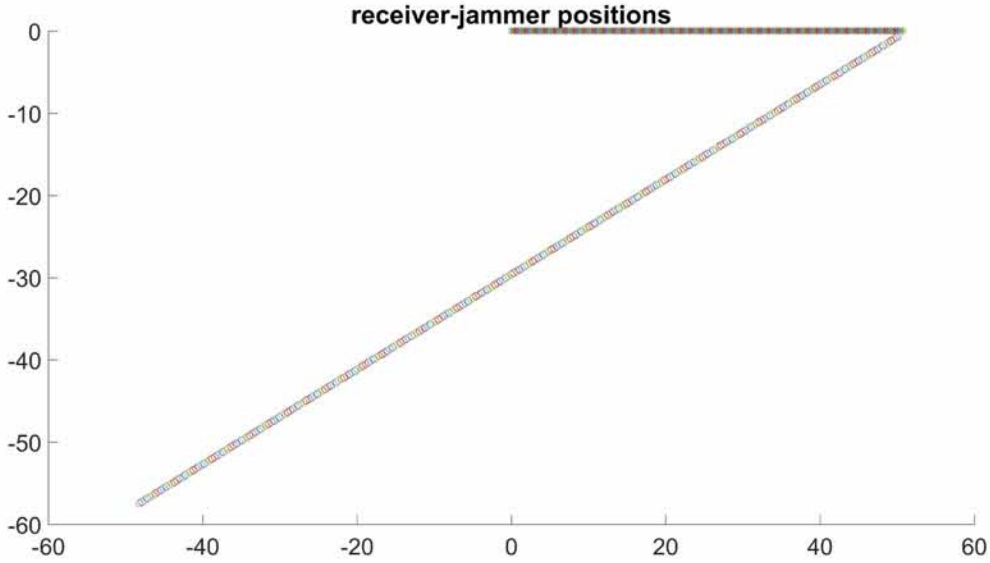

Figure 5.8: $\theta - \phi < 90°.(Vj > Vrx)$

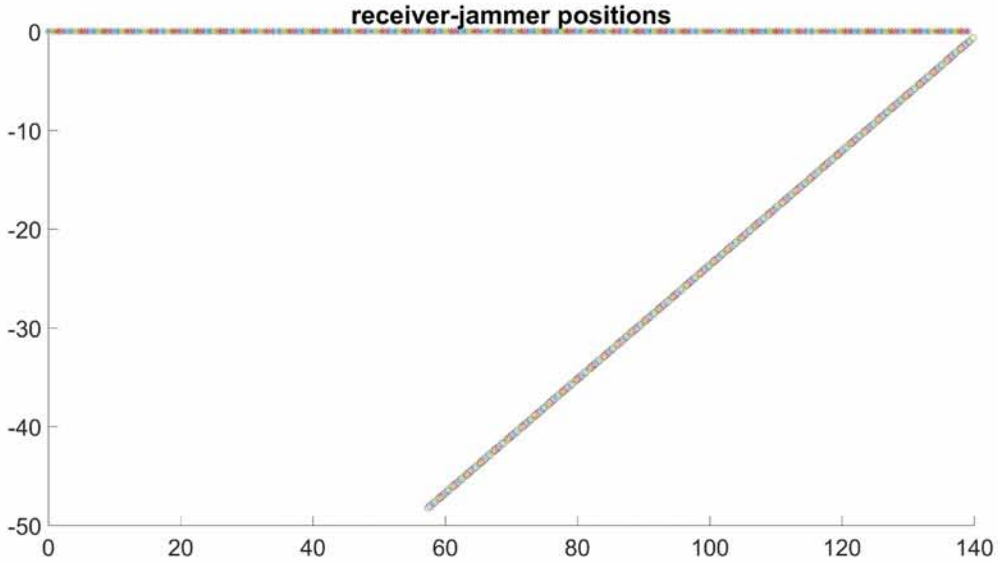

Figure 5.9: $\theta - \phi > 90°.(Vj < Vrx)$

Utilizing the Shadowing model for path loss exponent n = 3 and $\sigma_{dB} = 7$ and considering $\theta = 50°$ and $\phi = 20°$, for absent jammer at the beginning, noise at 55.5% of time as velocity defined at 5 m/s :

Table 5.8: indicative measurements of rush jammer with respect to initial means

rush jammer			
initial μ_W , μ_N	optimal μ_W , μ_N	optimal varW, varN	estimated π_N
2.50 , 2.44	2.2258 , 2.2255	10.805 , 10.805	50.9%
2.34 , 1.0	2.2258 , 2.2236	10.805 , 10.805	50.1%
2.34 , 1.96	2.2258 , 2.2236	10.805 , 10.805	50.1%

5.10 displays how velocity affects estimated ratio for means $\mu_W = 1.92797$, $\mu_N = 1.40497$
.

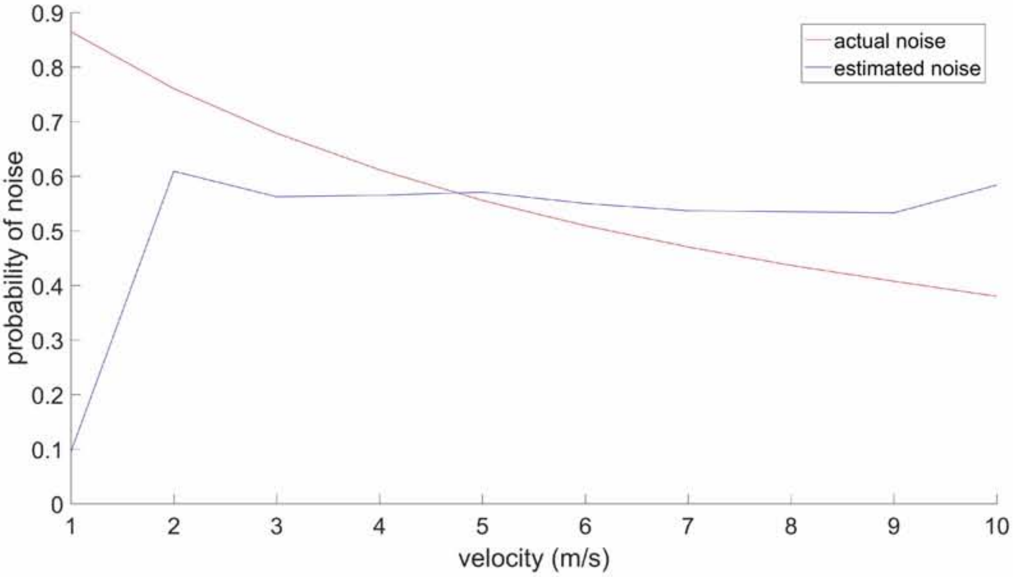

Figure 5.10: plot of jammer that rushes estimation for variable velocity

MSE = 0.0000441 and 5.11 dislays BER - SINR graph

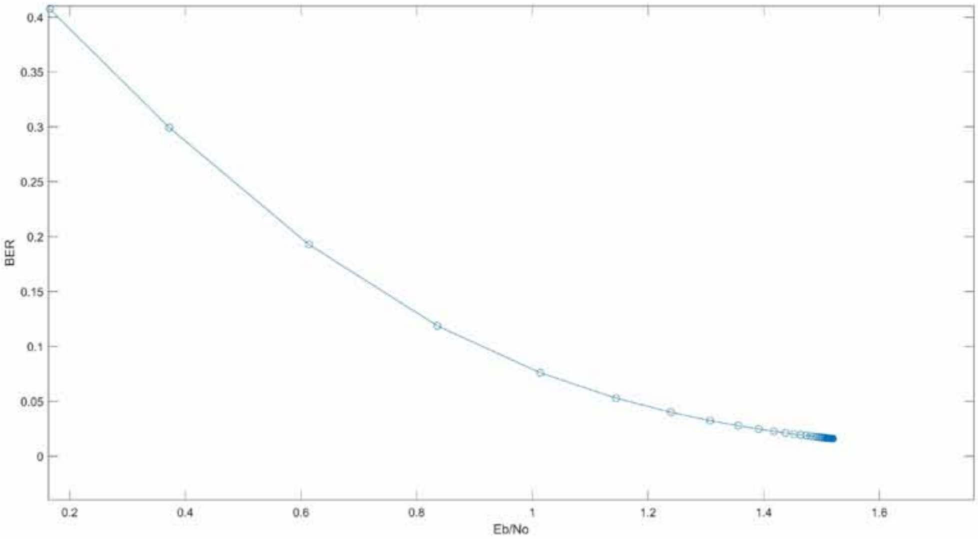

Figure 5.11: hypothetical plot of BER for rush jammer

5.5 Oncoming Jammer from the opposite direction

In this scenario , the jammer moves to the platoon's opposite direction on an adjacent lane.It is capable of tracking victim's exact velocity and adjusts its own velocity equal with it.It just has also knowledge about the distance between the two lanes(vertical distance when antennas aligned) and the angle between the connective line with the receiver and the vertical distance θ, the time the assault starts, before vehicles meet each other.For V = 1.8 m/s, and silence just after the beginning (preparation of jammer) and before the completion of the experiment when the vehicles are away and out of range, with noise at 49.3% of counted time, the selection of the initial means is again determined factor.Estimation is also depended on velocity variations.Numerical results were produced by the use of Shadowing model for path loss exponent n = 3 and $\sigma_{dB} = 7$ and considering $\theta = 82°$ and the velocity of receiver adopted by jammer is 1.8 m/s.

Table 5.9: indicative measurements of oncoming jammer

oncoming jammer			
initial μ_W , μ_N	optimal μ_W , μ_N	optimal varW, varN	estimated π_N
2.50 , 1.72	1.9813 , 1.9812	46.595 , 46.595	53.3%
2.34 , 1.24	1.9813 , 1.9810	46.595 , 46.595	51.7%
2.18 , 1.24	1.9814 , 1.9806	46.595 , 46.595	50.5%
2.02 , 1.48	1.9825 , 1.9740	46.603 , 46.595	50.0%

Figure 5.12 shows how velocity affects for means $\mu_W = 1.92797$, $\mu_N = 1.40497$.

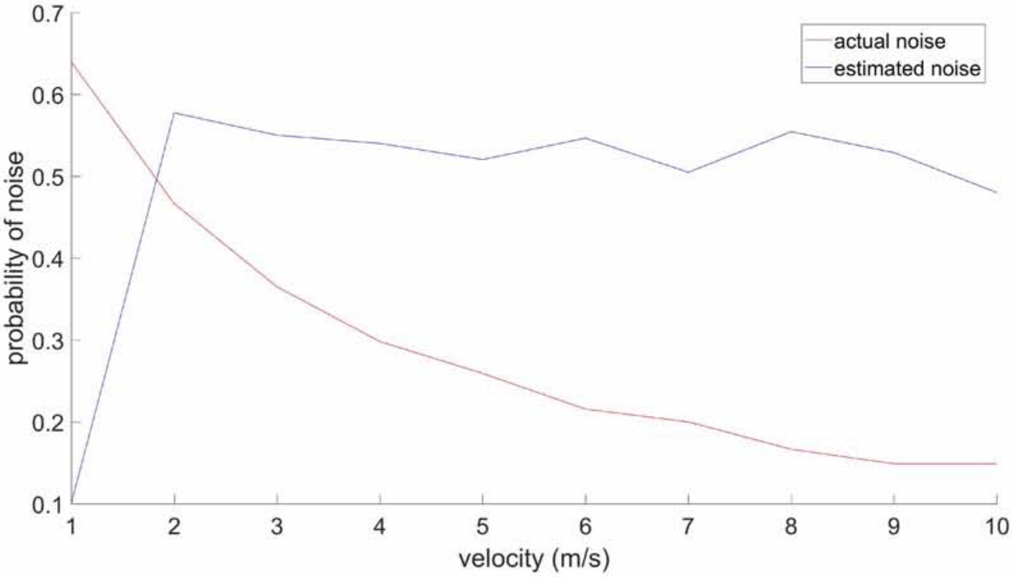

Figure 5.12: plot of oncoming jammer estimation for variable velocity

The MSE of the power in the channel is evaluated approximately: 0.000189 so graph 5.13 shows BER as function of SINR.

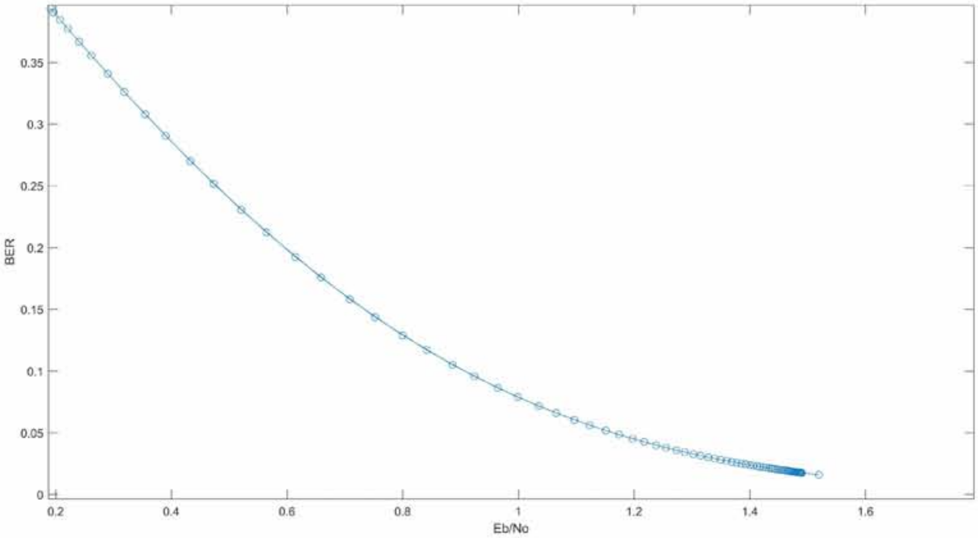

Figure 5.13: hypothetical plot of BER for oncoming jammer

6 Discussion

As a review of our results, the behaviour of the algorithm is satisfying.It always detects the jamming when occurs.Suitable initializations are important for the first means of the two probable cases.When the jamming percentage is lower than half, the estimated probability of jammng fluctuates around 50% but when it surpasses it, the algorithm is much more accurate and returns a much more precise estimation.The velocity of the moving vehicles affects in some kinds of attacks.

7 Conclusion

In this report the adequacy of Expectation Maximization algorithm for detection of jamming interference for vehicles moving in platoons was investigated.The outcomes of the algorithm for different profiles of jamming cases were computed and its judgement of what is happening in the last period of some seconds was estimated.
It is concluded that it is feasible EM algorithm to be used in vehicular networks, when critical messages, crucial for life are delivered and possibility of them to be modified, or dropped is urgent to be verified rapidly hence defense strategies will be activated immediately.

References

[1] S. R. Santana, J. J. Sanchez-Medina, and E. Rubio-Royo, “Platoon driving intelligence. a survey,” *Lecture Notes in Computer Science*, vol. 9520, pp. 765–772, 2015.

[2] J. Almeida, M. Alam, J. Ferreira, and A. S. Oliveira, “Mitigating adjacent channel interference in vehicular communication systems,” *Digital Communications and Networks*, vol. 2, no. 2, pp. 57–64, 2016.

[3] L. Mokdad, J. Ben-Othman, and A. Nguyen, “Djavan: Detecting jamming attacks in vehicle ad hoc networks,” *Performance Evaluation*, vol. 87, pp. 47–59, 2015.

[4] I. K. Azogu, M. T. Ferreira, J. A. Larcom, and H. Liu, “A new anti-jamming strategy for vanet,” *Ieee Globecom Workshops*, pp. 1344–1349, 2013.

[5] Y. O. Basciftci, F. Chen, J. Weston, R. Burton, and C. E. Koksal, “How vulnerable is vehicular communication to physical layer jamming attacks?,” *2015 Ieee 82nd Vehicular Technology Conference, Vtc Fall 2015 - Proceedings*, p. 7390968, 2015.

[6] D. Kosmanos, N. Prodromou, A. Argyriou, L. A. Maglaras, and H. Janicke, “Mimo techniques for jamming threat suppression in vehicular networks,” 2016.

[7] C. Pereira and A. Aguiar, “A realistic rf jamming model for vehicular networks: Design and validation,” *Ieee International Symposium on Personal, Indoor and Mobile Radio Communications, Pimrc*, pp. 6666447, 1868–1872, 2013.

[8] T. Levy and I. Lapidot, “Selective gmm em for telephone diarization,” *2016 Ieee International Conference on the Science of Electrical Engineering (icsee)*, 2016.

[9] H. Watanabe, S. Muramatsu, and H. Kikuchi, “Interval calculation of em algorithm for gmm parameter estimation,” *Ieee International Symposium on Circuits and Systems (iscas)*, pp. 2686–2689, 2010.