UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN MATHÉMATIQUES ET INFORMATIQUE
APPLIQUÉES

PAR
YOUSSEF LAHROUNI

DÉTECTION D'INTRUSIONS DANS LES RÉSEAUX VÉHICULAIRES SANS FIL

JUIN 2017

RÉSUMÉ

Le réseau VANET est une nouvelle technologie sur laquelle repose les systèmes de transport intelligents de l'avenir. Son but est de développer l'environnement véhiculaire et de le rendre plus confortable. En outre, il fournit plus de sûreté pour les véhicules et les conducteurs sur la route. Cependant, nous devons rendre cette technologie aussi sécurisée que possible contre de nombreuses menaces. Comme, le réseau VANET est une sous-classe du réseau MANET, il a donc hérité de nombreux problèmes de sécurité à cause de son architecture ouverte. Les attaques DOS font partie de ces problèmes.

Dans ce mémoire, nous nous sommes intéressés à l'étude des attaques DOS qui empêchent les utilisateurs de recevoir les bonnes informations au bon moment. Nous avons analysé le comportement et les effets des attaques DOS sur le réseau en utilisant différents modèles mathématiques afin de trouver une solution efficace.

Il ressort de ce travail que les deux méthodes basées sur la régression logistique et les réseaux de neurones permettent de mieux analyser les données et de mieux prédire ces attaques que les autres techniques mathématiques qui montrent des défauts d'efficacité.

ABSTRACT

VANET is a novel technology that will sweep the world in the future; its purpose is to develop the vehicular environment and to make it more comfortable. It also provides more safety on the road. Therefore, we have to make this technology as secured as possible against many threats. As VANET is a subclass of MANET, it has inherited many security problems with a different architecture and DOS attacks are one of them.

In this report, we focused on DOS attacks that prevent users to receive the right information at the right moment. We analyzed DOS attacks behavior and effects on the network using different mathematical models in order to find an efficient solution.

This work shows that the two methods of logistic regression and neural networks allow better analysis of the data and a better prediction of DOS attacks, whereas the other descriptive statistical techniques show less efficiency.

REMERCIEMENTS

Avant d'entamer mon présent mémoire, je remercie en premier lieu, Dieu et mes parents de m'avoir offert la possibilité de pouvoir poursuivre mes études dans de très bonnes conditions.

Cordialement je tiens à présenter mes sincères remerciements à ***Mr. Amar Boucif Bensaber,*** *mon Professeur et Directeur de recherche, d'avoir bien voulu m'encadrer durant toute la période de ma maitrise, ainsi que pour sa disponibilité, son aide financier et morale.*

Je ne saurai assez remercier mes professeurs ***Mr. Mhamed Mesfioui, Mr. Ismaïl Biskri*** *pour leur disponibilité, leur soutien et toutes leurs interventions pertinentes durant mon projet.*

Je tiens également à présenter mes sincères remerciements aux professeurs ***Mr. Mourad Badri, Mme. Linda Badri*** *et à tous les professeurs et personnels du département de Mathématiques et informatique appliquées, qui m'ont apporté l'aide nécessaire, je tiens à leurs exprimer mes remerciements pour la sympathie qu'ils m'ont adressé durant ma maitrise, ainsi que leurs conseils pertinents.*

Enfin, je témoigne ma profonde gratitude à tous les étudiants du laboratoire LAMIA qui m'ont aidé pour mener à bien ma recherche, et pour le temps qu'ils m'ont consacré malgré leurs diverses occupations, qu'ils soient assurés de toute ma gratitude.

TABLE DES MATIÈRES

LISTE DES FIGURES

LISTE DES ABREVIATIONS

AES: Advanced Encryption Standard.

ASTM: American Society for Testing and Materials.

CA: Central Authority.

CSMA/CA: Carrier Sense Multiple Access / Collision Avoidance.

D.E.S: Data Encryptions Standard.

DSRC: Dedicated Short Range Communication.

EDCA: Enhanced Distributed Channel Access.

GPS: Global Positioning System.

IEEE: Institute of Electrical and Electronics Engineers.

MAC: Medium Access Control.

MANET: Mobile Ad hoc Network.

OBU: On Board Unit.

RCP: Resource Command Processor.

RM: Resource Manager.

RMA: Resource Manager Application.

RSU: Road Side Unit.

STI: Systèmes de Transport Intelligents.

SUMO: Simulation of Urban Mobility.

TPD: Tamper Proof Device.

TCP: Transport Control Protocol.

VANET: Vehicular Ad hoc Network.

V2V: Vehicular-to-Vehicular.

V2I: Vehicular-to-Infrastructure.

WAVE: Wireless Access for the Vehicular Environment.

WiFi: Wireless Fidelity.

WLAN: Wireless Local Area Network.

WMAN: Wireless Metropolitan Area Network.

WSMP: WAVE Short Message Protocol.

WSP: Wave Short Message.

WPAN: Wireless Personal Area Network.

WWAN: Wireless Wide Area Network.

RMS : Root Mean Square.

MAV : Mean Absolute Value.

ER : Erreur Relative.

IVC : Inter Vehicular Communication.

DOS : Denial Of Service.

DDOS : Distributed Denial Of Service.

ORT : On board Radio Transponder.

RSRT : Road Side Radio Transductor.

RRDA : Request Response Detection Algorithm.

APDA : Attack Packet Detection Algorithm.

UDP : User Datagram Protocol.

IP : Internet Protocol.

RBSP : Reference Broadcast Synchronization Protocol.

MCF : Master Chock Filter.

SAN : Stochastic Automata Networks.

FHSS : Frequency Hopping Spread Spectrum.

DSDV : Destination Sequenced-Distance Vector.

SEAD : Secure and Efficient Ad hoc Distance Vector.

VA : Véhicule Attaqué.

VNA : Véhicule Non Attaqué.

INTRODUCTION

Les systèmes de transport intelligents (STI) sont des applications qui visent à fournir des services novateurs dans différents modes de transport et de gestion du trafic. Ils permettent aux utilisateurs d'être mieux informés sur l'état de la route, et rendent l'utilisation des réseaux de transport plus sûre, plus coordonnée et plus intelligente [26].

Les réseaux véhiculaires sans fil (VANET) sont un élément clé des systèmes de transport intelligents. Ils permettent aux utilisateurs d'avoir des informations sur leur itinéraire, sur la météorologie, l'état de la route, l'accès à Internet, ...tout en temps réel. Ces réseaux donnent accès également aux différentes applications de détente et de divertissement comme transporter des données multimédia, avoir des informations touristiques, etc.

Le réseau VANET donne accès aux informations en temps réel et en privé. Ce fait a un grand impact sur le comportement de l'utilisateur, sur sa vie et sur celle des passagers. Ce réseau assure également la sécurité des utilisateurs en préservant l'anonymat et en sécurisant leurs données personnelles.

Dans ce mémoire, nous nous intéressons à l'étude de la détection de l'attaque DOS (déni de service) sur le réseau VANET. Cette attaque est considérée parmi les plus dangereuses attaques sur les réseaux. Elle empêche l'utilisateur de recevoir l'information demandée au bon moment. Nous proposons ici, l'étude du comportement du réseau véhiculaire soumis à une attaque DOS simulée.

Le premier chapitre présente la définition des réseaux sans fil existant (Ad-hoc et VANET), ainsi que l'architecture du réseau VANET, ses différentes caractéristiques et normes de communication, et ses mécanismes de sécurité.

Le deuxième chapitre est consacré à la revue de la littérature présentant les travaux récents qui étudient l'attaque DOS sur le réseau VANET.

Dans le chapitre 3, nous présentons notre papier scientifique qui a été soumis à « The 15th ACM International Symposium on Mobility Management and Wireless Access », ainsi qu'un résumé du papier en français expliquant son objectif, sa démarche, et sa méthodologie.

Le quatrième et dernier chapitre présente les résultats de la simulation de l'attaque DOS sur le réseau VANET et des méthodes utilisées afin de prédire cette attaque. Nous terminons ce mémoire par une conclusion et nous proposons des pistes de solutions pour les travaux futurs.

CHAPITRE 1

RESEAUX VEHICULAIRES SANS FIL : ARCHITECTURE – CARACTERISTIQUES - SECURITE

1- Introduction

Les réseaux véhiculaires sans fil (VANET) sont un nouveau type d'application des réseaux Mobile Ad-hoc network (MANET). Le rôle du réseau VANET est de fournir et de garantir la communication entre les véhicules, et de permettre l'accès à l'information en temps réel. Le but de ce réseau est d'améliorer la sécurité routière, de réduire au maximum possible le risque des collisions, et de minimiser les dégâts en cas d'accident.

Cette technologie offre des services aux conducteurs permettant d'améliorer les conditions de la conduite. Parmi ces services on trouve : des informations sur l'état de la route et des informations météorologiques, la prévention des accidents et des congestions, l'accès à Internet,... Les passagers peuvent également en profiter en utilisant les applications de confort et de détente (partage de musiques et de vidéos) pour rendre la route plus agréable.

Ce premier chapitre présente tout d'abord la définition des réseaux sans fil : le réseau Ad-hoc et le réseau VANET. On y aborde ensuite l'architecture de ce réseau VANET (c'est-à-dire les différentes entités et modes de communication, les types de messages et d'applications, et l'environnement de déploiement) et ses différentes caractéristiques ainsi que les problèmes et les mécanismes de sécurité. Ce chapitre est terminé par la présentation des techniques d'accès et des différentes normes standards de communication.

2- Définition des réseaux

2.1-Réseau Ad-hoc

Les réseaux Ad-hoc sont des réseaux sans-fil capables de s'organiser spontanément et de manière autonome dans leur environnement. Les nœuds proches effectuent une communication directe entre eux, tandis que les autres nœuds jouent le rôle d'intermédiaire pour transporter le message jusqu'à sa destination finale. Dans le cas du réseau VANET, les véhicules sur la route jouent le rôle des nœuds qui transmettent le message dans les protocoles de routage.

Dans un réseau Ad-hoc on rencontre plusieurs problèmes comme l'absence d'infrastructure, une bande passante limité, beaucoup de perte de données, des erreurs de transmission, et plusieurs problèmes de sécurité.

2.2-Réseau VANET

Dans le réseau VANET, les nœuds sont les véhicules roulant sur la route. Ce réseau a pour but d'établir la communication entre ces nœuds et les unités situées aux bords de la route. Le réseau VANET est caractérisé par une forte mobilité des nœuds. Cette mobilité dépend de la différence de vitesses des véhicules sur la route ce qui rend la topologie du réseau très dynamique et variable en permanence, et donc le lien entre les véhicules est instable.

Ce type de réseaux a nécessairement besoin de certains équipements électroniques dans le véhicule, ainsi que de plusieurs technologies pour établir la communication véhiculaire tel que : le réseau WIFI 802.11, Bluetooth, et le standard IEEE1609.

3- Architecture du réseau VANET

3.1-Entités de communication

3.1.1- Road Side Unit

Les Road Side Unit (RSU) sont des entités situées et installées au bord de la route. Ces entités présentent des points d'accès au réseau et sont déployées tout au long de la route. Chaque RSU a pour objectif de transmettre des messages aux véhicules qui se trouvent dans sa zone radio. Ces messages contiennent des informations sur les conditions météorologiques, ainsi que sur l'état de la route (vitesse maximale, autorisation de dépassement, etc.).

3.1.2- Autorité centrale

L'Autorité Centrale (CA) est un serveur de stockage et de transaction qui a la confiance de toutes les entités du réseau. Elle fournit des services et des applications à tous les utilisateurs, ainsi que les certificats, les clés ou pseudonymes de communication des véhicules [1].

3.1.3- On Board Unit

L'On Board Unit (OBU) est une unité embarquée dans les véhicules intelligents. Son rôle est de permettre aux véhicules de se localiser, calculer, enregistrer et envoyer des messages sur une interface réseau à l'aide d'un ensemble de programmes.

Dans le réseau VANET, le conducteur ou l'utilisateur peut voir les pseudonymes des véhicules à proximité dans son OBU à l'aide des messages beacon. Ainsi, l'utilisateur peut choisir le véhicule avec lequel il veut communiquer.

3.2-Modes de communication

La communication dans le réseau VANET s'effectue entre les véhicules d'une part, et entre les véhicules et les entités fixes (RSU et CA) d'autre part (figure 1).

3.2.1- Communication en mode Ad-hoc (V2V)

Connue aussi sous l'appellation de communication Véhicule-à-Véhicule (V2V), la communication en mode Ad-hoc est la communication directe entre les véhicules. Chaque véhicule représente un nœud et établie la communication à l'aide de son OBU. Cette communication ne nécessite aucune infrastructure et sert aussi à diffuser l'information dans le réseau ou à la transporter d'un nœud vers un autre.

3.2.2- Communication en mode infrastructure (V2I)

La communication en mode infrastructure nommée aussi Véhicule-à-infrastructure (V2I), est la communication réalisée entre les nœuds (Véhicules en utilisant l'OBU) et les entités fixes (RSU et CA). Ceci permet aux véhicules d'accéder aux différents types d'applications (Sécurité, confort, gestion…) et aux différents types d'informations (Etat du trafic, météo …).

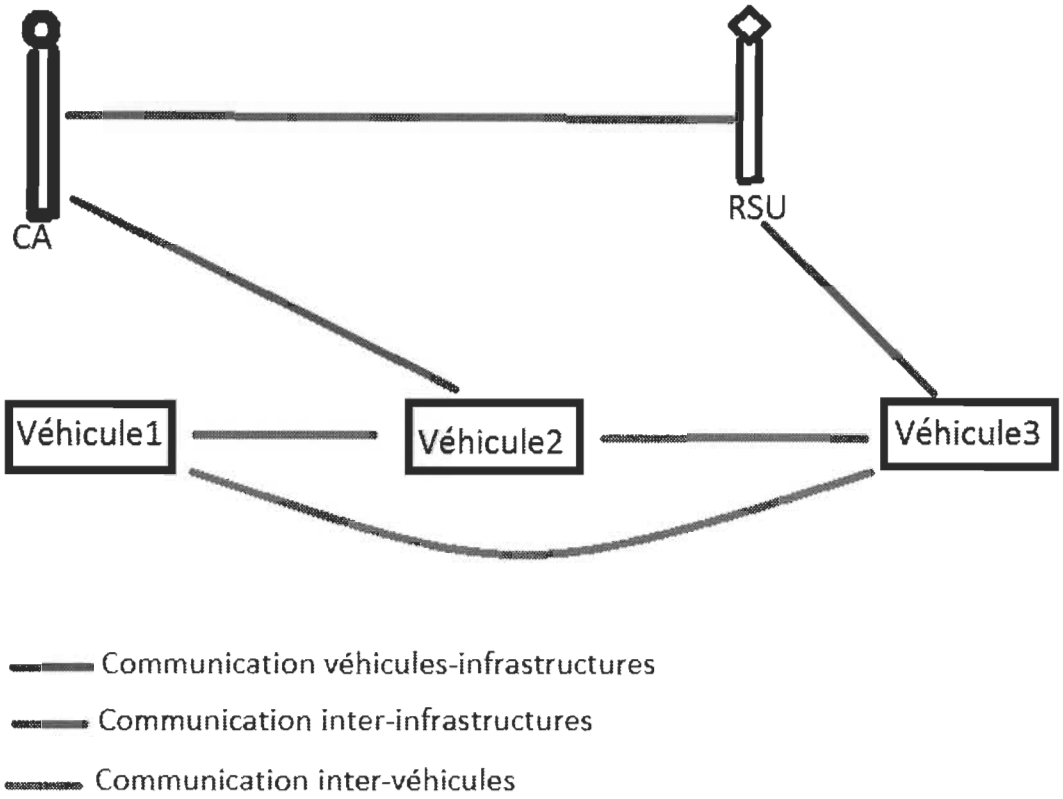

Figure 1 Les différents modes de communication dans le réseau VANET

3.3-Types de messages

3.3.1- Messages de contrôle

Les messages de contrôle sont les messages « Beacon ». Chaque véhicule émet un message de contrôle toutes les 100 ms pour s'afficher aux autres véhicules voisins. Ce message contient des informations sur le véhicule tel que sa position, sa vitesse, sa direction etc.

3.3.2- Messages d'alerte

Les messages d'alerte ou les messages de sécurité sont envoyés lors d'une situation dangereuse ou lors d'un événement méritant l'attention du conducteur. Dans le cas d'un accident par exemple (fig.2), le message d'alerte sert à prévenir les véhicules qui se dirigent vers la zone de l'accident ou de congestion de ce fait. Ces messages sont urgents et importants et leurs tailles sont petites afin de garantir leur transmission plus rapidement. Les véhicules désignés par la retransmission des messages d'alerte doivent les transmettre dès réception. Les messages d'alerte contiennent les coordonnées du lieu en question.

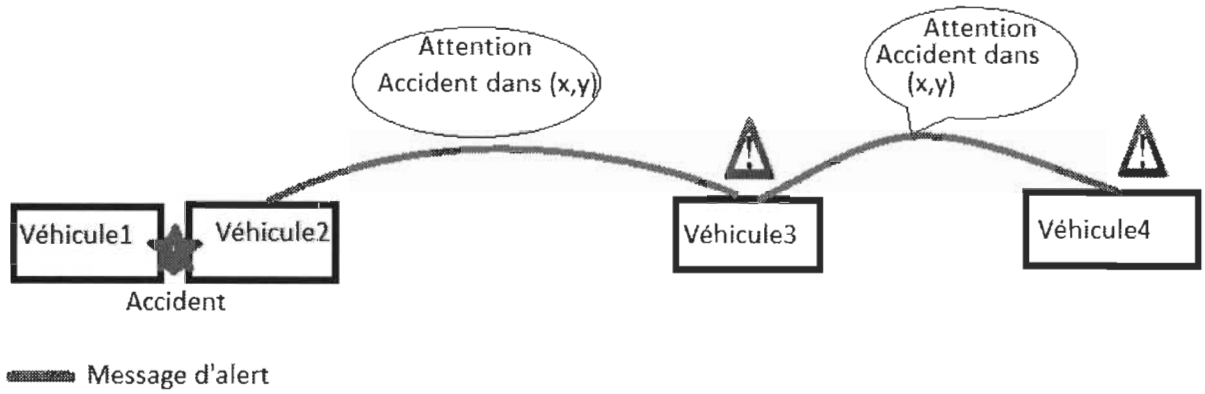

Figure 2 Exemple de message d'alerte lors d'un accident avec les coordonnées du lieu

D'autres messages existent qui ne sont ni message « Beacon » ni message sécurité, et ne sont émis qu'une seule fois, ils peuvent être des messages de courrier électronique ou des messages d'une application.

3.4-Application

On distingue trois types d'application dans les réseaux véhiculaires sans fil : des applications sur la gestion du trafic routier, des applications de confort et ceux du trafic routier.

3.4.1- Applications de gestion du trafic routier

Les applications de gestion du trafic routier permettent d'éclaircir au conducteur l'état de la route. Grâce aux messages échangés entre les véhicules, ces derniers deviennent des capteurs de trafic. En fournissant des informations sur l'état de la route, les véhicules collaborent afin de céder le passage à l'ambulance, d'éviter les congestions et les embouteillages (fig.3), et de proposer d'autres itinéraires aux véhicules qui se dirigent au même endroit.

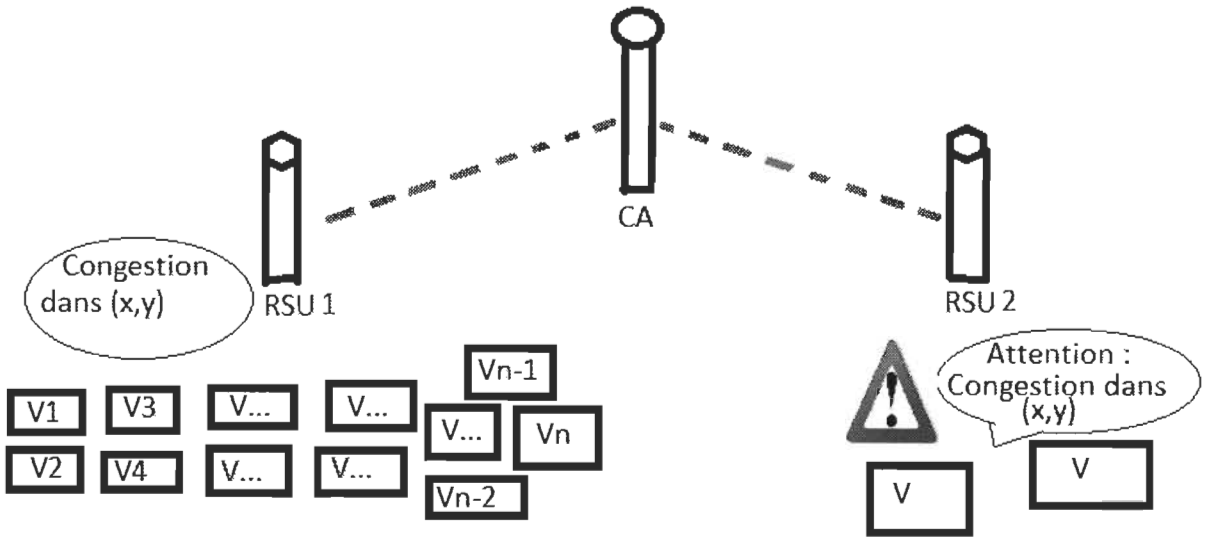

Figure 3 Prévention d'une congestion

3.4.2- Applications de confort

Les applications de confort sont des applications qui offrent le confort au conducteur tout au long de la route, surtout pendant les longs trajets (fig.4). Grâce à ces applications et à l'accès à internet, le conducteur et les passagers ont la possibilité d'écouter de la musique, de voir des vidéos, de jouer à des jeux en ligne, de localiser les restaurants et les stations à proximités, d'avoir des informations touristiques, ainsi que de vérifier à distance les papiers du conducteur et d'effectuer le paiement à distance sur l'autoroute.

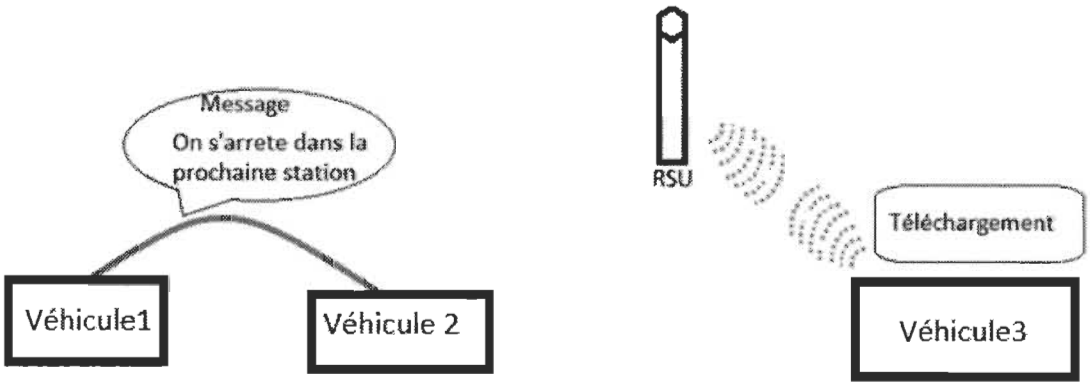

Figure 4 Exemple d'application de confort

3.4.3- Application de sécurité du trafic routier

De nombreux facteurs causent les accidents de la route, par exemple, l'inattention du conducteur, les conditions météorologiques, les problèmes techniques internes du véhicule, ou un problème venant de l'extérieur comme le cas d'un véhicule ne respectant pas le code de la route. Les applications de sécurité du trafic routier jouent un rôle très important pour éviter les accidents ou pour diminuer le risque d'en avoir par

la prévention des accidents à l'avance, la détection du problème et l'information du conducteur par un message d'alerte de sécurité.

4- Normes et standards de communication

Les standards de communication sont indispensables pour mettre en œuvre la communication entre les différentes entités dans les réseaux véhiculaires sans fil. L'IEEE a étendu sa famille de protocoles 802.11 en ajoutant le 802.11p [IEE 10], et en s'inspirant du standard ASTM E2213-03 [AST 07] qui est basé sur le 802.11a [IEE 99]. Ce protocole modifie la couche physique et la couche MAC pour s'adapter aux réseaux de véhicules en conformité avec la bande DSRC2. En complément, l'IEEE a défini la famille de protocoles 1609, dite WAVE, pour l'accès sans fil dans les réseaux de véhicules [IEE 10]. Ce standard structuré en quatre composantes (1609.1 à 1609.4), définit l'architecture, le modèle de communication, la structure de gestion, la sûreté et l'accès physique [5].

a. Le DSRC

Les DSRC (Dedicated Short Range Communications) sont des communications sans fil dédiées à courte ou moyenne portée et à sens unique ou double. Ce type de communications est spécialement conçu pour les systèmes de transport intelligent (STI), c'est-à-dire, pour la communication entre les véhicules (V2V) ou entre véhicules et infrastructure (V2I). Le sigle DSRC désigne aussi l'ensemble de protocoles et de normes mis en jeu dans ce type de communications. Des études ont montré que le DSRC assure le bon fonctionnement des applications de sécurité du trafic routier.

Le DSRC œuvre dans la bande de fréquence de 5.9 GHZ. Cette bande de fréquence est divisée en sept canaux de 10 MHZ chacun. L'ensemble de ces canaux se répartit fonctionnellement en un canal de contrôle et six canaux de service. Le canal de contrôle est réservé à la transmission des messages de gestion du réseau et des messages très importants tels que les messages liés à la sécurité routière. Les six autres canaux sont dédiés à la transmission de données de services annoncées sur le canal de contrôle [1].

b. IEEE 1609.1

Le standard IEEE 1609.1 se situe au niveau de la couche application, et définit les différents formats de messages et le mode de stockage des données utilisé par cette application.

IEEE1609,1 décrit le gestionnaire de ressources RM (Resource Manager : assure les services qui permettent au RMA de contrôler les interfaces présentes dans l'OBU) qui spécifie la méthode d'accès sans fil dans un

environnement WAVE. Il permet au RSU de communiquer avec les OBU. Il permet aussi à l'RMA (resource manager application : entité distante qui utilise le RM pour communiquer avec le RCP) d'établir une connexion avec un RPC (resource command processor : exécute les commandes données par le RMA et fournit une réponse au RMA via le RM) dans l'OBU.
Lorsqu'une application (présente dans un OBU ou dans un RSU) désire envoyer une commande à un OBU, le composant RMA envoie un message aux RM. Le RM envoie la commande aux RCP qui commande les OBU connectés. Le RCP envoie un message de réponse au RM afin de délivrer le résultat. Le RM est donc le lien entre les applications d'un RSU (ou OBU) et les OBU des autres véhicules [5].

c. IEEE 1609.2

Le standard 1609.2 définit le format des messages sécurisés pour le système DRSC/WAVE. Il spécifie les algorithmes pour sécuriser les messages de gestion et d'application. Il décrit également les procédures pour assurer à chaque véhicule des services tels que l'authenticité, la confidentialité, l'intégrité et la non-répudiation des données. Toutes les applications ne requièrent pas ces services. Néanmoins, elles doivent être présentes en cas de nécessité. Le standard 1609.2 protège les entités du réseau contre les attaques telles que l'homme du milieu, l'usurpation d'identité, le rejoue de message, etc. [3].

d. IEEE 1609.3

Le standard 1609.3 définit le WAVE Short Message (WSM) et le protocole d'échange associé au WAVE Short Message Protocol (WSMP) afin d'assurer les fonctionnalités de la couche réseau et la couche transport des applications de sécurité routière. Le 1609.3 définit aussi le message WAVE Service Advertisement (WSA) qui est utilisé pour annoncer la disponibilité des services DSRC à une localisation donnée. Un WSA peut par exemple être envoyé pour annoncer la présence d'un service d'information trafic mis en place par un RSU [5].

e. IEEE 1609.4

Le standard IEEE 1609.4 définit l'organisation, l'ordonnancement et l'utilisation des différents canaux de DSRC [4,10]. Le but de ce standard est d'établir un mécanisme permettant à plusieurs équipements de se trouver (s'accorder sur le même canal au même moment afin de pouvoir communiquer). Le 1609.4 a une forte relation avec le mécanisme EDCA (Enhanced Distributed Channel Access) de la sous-couche MAC (Medium Access Control). L'EDCA est basé sur le CSMA/CA et est utilisé dans les réseaux Wifi supportant

le standard IEEE 802.11e. Le mécanisme EDCA permet d'attribuer des priorités à chaque type de message [1].

f. IEEE 802.11p

IEEE802.11pestunemodification approuvée à la normeIEEE802.11 pour ajouter un accès sans fil à des environnements de véhicules(WAVE). Il définit les améliorations du 802.11 (la base des produits commercialisés en tant que Wifi) nécessaires pour soutenir des systèmes de transport(ITS) intelligents. Ceci comprend l'échange de données dans la communication entre les véhicules (V2V) à grande vitesse et dans la communication entre les véhicules et l'infrastructure routière (V2I) dans une bande de fréquences de 5,9GHz (5,85 à 5,925 GHz). IEEE1609est une norme de couche supérieure sur la base de la norme IEEE802.11p.

5- Sécurité dans les réseaux véhiculaires sans fil

L'objectif des réseaux véhiculaires sans fil est d'améliorer le trafic routier et de préserver la sécurité des conducteurs et des passagers. Pour ces raisons, la sécurité dans ces réseaux présente un défi très important aux chercheurs. De nombreux travaux de recherche s'intéressent au sujet de la sécurité des réseaux véhiculaires sans fil afin de lutter et de protéger ces réseaux des menaces et des attaques.
Dans le même but, notre étude s'intéresse à la sécurité du réseau VANET, aux différents types d'attaques que peut subir ce réseau, aux moyens de prévention de ces attaques et aux solutions.

5.1-Les attaques dans le réseau VANET

a. Déni de service

Le déni de service est une attaque qui se produit quand l'attaquant prend le contrôle des ressources d'un véhicule et bloque le canal de communication utilisé par le Réseau des véhicules, de sorte qu'il empêche les informations importantes et urgentes d'arriver à leur destination. Cette attaque amplifie le risque si le conducteur doit dépendre de l'information.
La figure 5illustre une attaque par déni de service conduisant à une collision, l'attaquant C empêche l'échange de messages urgents et importants entre le véhicule accidenté B et le véhicule A.

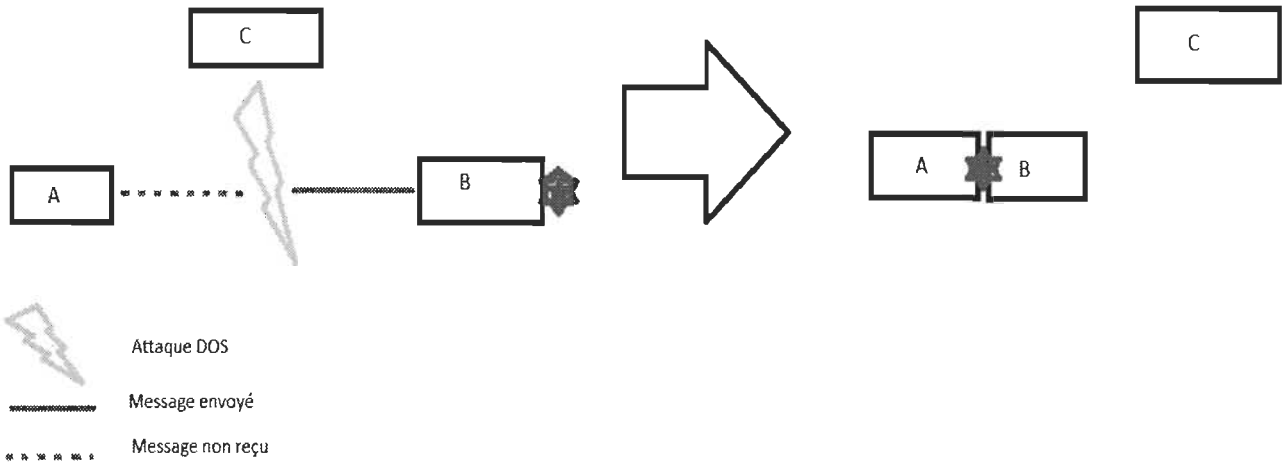

Figure 5 Déni de service

b. Attaque sur la cohérence de l'information

Dans l'attaque sur la cohérence de l'information, l'entité malveillante vise à injecter et à diffuser des informations erronées sur l'état du trafic routier dans le réseau, afin de modifier le comportement des autres entités comme provoquer le changement d'itinéraire des autres véhicules ou les guider vers une congestion.

c. Suppression de messages

Un attaquant peut supprimer des paquets du réseau. Ces paquets peuvent contenir des informations importantes pour le récepteur. L'attaquant supprime ces paquets et peut les utiliser à nouveau dans un autre temps. Le but d'un tel attaquant est d'empêcher les autorités d'immatriculation et d'assurance de savoir les collisions impliquant son véhicule ou d'éviter de remettre des rapports de collision aux points d'accès en bordure de route. Par exemple, un attaquant peut supprimer un avertissement de la congestion et l'utiliser dans un autre temps, pour que les autres véhicules ne reçoivent pas l'avertissement et se trouvent obliger d'attendre dans le trafic.

d. Altération des messages

L'altération des messages est une attaque qui se produit lorsqu'un attaquant modifie une des données existantes, ou modifie l'entrée effective des données transmises.

e. Sybil attaque

La Sybil attaque est une attaque qui se produit quand un attaquant crée un grand nombre de pseudonymes, et prétend ou agit comme s'il y a plus de véhicules sur la route pour informer les autres véhicules d'une fausse congestion et les forcer à changer d'itinéraire.

f. Attaque sur la vie privée

L'attaque sur la vie privée est une attaque qui se produit lorsqu'une entité malveillante essaye d'avoir l'identité et les informations personnelles d'une autre entité du réseau, afin de tracer les activités et les déplacements de cette dernière. Pour ce type d'attaque, les chercheurs proposent comme solution, un changement périodique des pseudonymes de communication afin d'éviter la traçabilité des véhicules [1].

g. Usurpation d'identité

Dans l'usurpation d'identité, l'entité malveillante utilise l'identité d'une autre entité pour être légitime dans le réseau et donc avoir les attributions de cette dernière.

5.2-Eléments et mécanismes de base de la sécurité dans le réseau VANET

Suite aux attaques citées ci-dessus, les réseaux véhiculaires sans fil doivent assurer des services de sécurité aux utilisateurs. Dans ce paragraphe, nous allons citer les différents éléments et mécanismes de base de la sécurité du réseau VANET.

5.2.1- Tamper-Proof Device (TPD)

L'objectif du TPD est de sécuriser les communications internes aux véhicules et externes en empêchant la récupération des données des capteurs. Le TPD est résistant au sabotage et permet de stocker les données confidentielles telles que les paires de clés et les certificats. Il permet même de signer les messages envoyés par le véhicule.

5.2.2- La cryptographie

La cryptographie est une discipline de la cryptologie. Elle permet de protéger les messages échangés contre les attaques et de garantir la confidentialité, l'authenticité et l'intégrité des informations en appliquant des

algorithmes de chiffrement. Elle transforme, à l'aide d'une clé, un message clair en un message chiffré et incompréhensible pour toute entité ne disposant pas de clé de déchiffrement.

5.2.3- Les fonctions de hachage

Les fonctions de hachage calculent, à partir d'une donnée fournit à l'entrée, une empreinte qui sert à identifier rapidement la donnée initiale. Ces fonctions sont utilisées aussi dans la cryptographie. Pour un ensemble de données de grande taille, la fonction de hachage peut renvoyer des résultats bien optimisés.

5.2.4- La signature numérique

La signature numérique est une donnée reliée aux messages et générée par la fonction de hachage et la clé privée de l'émetteur pour assurer l'authenticité, l'intégrité et la non-répudiation des messages.

5.2.5- Les certificats numériques

Les certificats numériques dans le réseau VANET sont des preuves d'identité du propriétaire d'une clé publique. Ces certificats sont délivrés par l'autorité de certification, ils renforcent la sécurité.

5.3- Concepts et mécanismes de sécurité

Les attaques menacent les usagers des réseaux véhiculaires sans fil. Afin de détecter et de prévenir ces attaques et d'améliorer la sécurité des systèmes, des services de sécurité existent. Parmi ces services on trouve la confidentialité, l'authentification, l'intégrité des données, la non-répudiation, la disponibilité, la gestion de la vie privée et le contrôle d'accès.

5.3.1- La confidentialité

La confidentialité protège les informations transmises dans le réseau contre les attaques des entités malveillantes ou des entités non autorisées à les avoir (l'écoute clandestine par exemple). On distingue deux niveaux de protection :

- Le service global : protège toutes les données transmises entre deux utilisateurs pendant une période donnée par l'établissement d'un circuit virtuel.
- Le service plus restreint : protège un message par l'ajout de champ(s) spécifique(s) à l'intérieur de ce message.

5.3.2- L'authentification

Pendant la communication, chaque message doit être authentifié. L'authentification est le concept qui permet de savoir le vrai auteur de chaque message dans le réseau. Elle aide à faire confiance aux messages diffusés et à s'assurer de leur origine surtout quand il s'agit d'un message de sécurité. L'authentification permet aussi d'empêcher l'attaque Sybil. L'authenticité est assurée par des mécanismes proactifs d'authentification. Il existe deux types d'authentification : l'authentification des messages et l'authentification des entités.

5.3.3- L'intégrité

L'intégrité protège les messages et empêche les attaquants de les altérer ou les modifier. L'intégrité physique est une fonction permettant d'assurer le matériel destiné aux opérations cryptographiques et l'envoi des messages contre les altérations.

Le service d'intégrité des messages assure la rapidité de réception des messages envoyés, sans duplication, insertion, modification, réorganisation ou répétition. A l'instar de la confidentialité, l'intégrité s'applique à un flux de messages, à un seul message, ou à certains champs à l'intérieur d'un message. La meilleure solution est la protection totale du flux.

5.3.4- La non-répudiation

La non-répudiation facilite la capacité d'identifier les attaquants et de retrouver la source des messages erronés même après que l'attaque soit produite. Ceci empêche les attaquants de diffuser les fausses informations dans le réseau. Toute information relative à la voiture (le voyage déroute, vitesse, temps, tout contrevenant) est stockée dans le TPD. Une autorisation officielle de maintien latéral peut récupérer ces données.

Généralement la signature numérique est utilisée pour garantir la non-répudiation des messages, des applications de sécurité et de gestion du trafic routier. Quant aux messages des applications de gestion de confort, la non-répudiation n'est pas si nécessaire sauf pour les messages impliquant des transactions financières [1].

5.3.5- La disponibilité

Le réseau véhiculaire doit être disponible tout le temps, donc on parle de l'accès permanent de plusieurs applications réseaux en temps réel. Ces applications doivent répondre rapidement à partir du réseau de capteurs ou du réseau Ad-hoc. Un délai en secondes pour les applications de sécurité du trafic peut rendre le message sans importance et peut être les conséquences seront graves.
Tenter de répondre à la demande en temps réel rend le système vulnérable à l'attaque DOS. Dans certains messages, un retard dans la milliseconde, rend le message vide de son sens.

5.3.6- La gestion de la vie privée

Il est très important de garder les informations personnelles des conducteurs (l'identité réelle, le chemin de voyage, la vitesse) loin des observateurs non autorisés. Pour préserver la vie privée des conducteurs, il faut mettre en place un protocole de gestion de l'anonymat à l'aide des pseudonymes de communication qui doivent être changés fréquemment.

5.3.7- Le contrôle d'accès

Le rôle du contrôle d'accès est de définir les entités qui sont autorisées à se connecter au réseau et de bloquer les utilisateurs non autorisés. Ce contrôle d'accès est très important pour que quelques applications puissent distinguer les différents niveaux d'accès en fonction de l'entité. Par exemple, permettre à la police et les secours d'échanger des informations avec les feux tricolores pour contrôler le trafic.

6- Conclusion

Ce chapitre définit le réseau véhiculaire sans fil, son architecture, ses différentes caractéristiques, ses domaines d'applications, ainsi que ses différents problèmes liés à la sécurité.
Le but principal du réseau VANET est de faciliter la conduite et de rendre la route plus agréable, plus confortable et plus amusante aux conducteurs et aux passagers. Il s'agit aussi d'éclaircir le trafic routier pour les conducteurs.
Les attaques du réseau VANET présentent un danger primordial, car elles menacent la vie privée des conducteurs et des passagers. Elles peuvent aussi causer des accidents et des congestions sur la route. Il est très important de mettre en œuvre les protocoles et les mécanismes de sécurité afin de contrôler les entités

du réseau, de préserver la sécurité des conducteurs, des passagers, et des véhicules, et permettre aussi aux usagers d'utiliser cette technologie en étant confiant.

Le prochain chapitre, nous présentons quelques travaux de recherche récents qui se sont intéressés à l'attaque DOS dans le réseau VANET. Cet état de l'art permettra d'avoir plus d'informations sur ce type d'attaque ainsi que sur les différentes solutions proposées.

CHAPITRE 2

ÉTAT DE L'ART

1- Introduction

Dans un réseau VANET, la sécurité du conducteur et des passagers dépend fortement du fonctionnement et de la sécurité de ce réseau, que ce soit pendant l'envoie des messages, la gestion de l'anonymat, ou tout autres actions (les transactions bancaires par exemple) nécessitant un niveau de sécurité cruciale.

Dans notre étude, le but est de protéger le réseau contre les attaques DOS afin de prévenir le dysfonctionnement du réseau et l'occupation de la bande passante. Ainsi le bon fonctionnement du réseau est garanti et les utilisateurs peuvent en profiter au maximum.

Dans ce chapitre, nous présentons une étude bibliographique portant sur certains scénarios d'attaques DOS et les solutions proposées par les auteurs pour lutter contre ces attaques DOS. Ces solutions ne sont pas très efficaces contre ce genre d'attaque, mais elles nous ont permis de mieux cerner le problème pour apporter une contribution à la solution.

2- Définition de l'attaque DOS

Dans l'étude [26], les auteurs définissent l'attaque DOS (Deni Of Service) comme étant le résultat de toutes actions qui empêchent la totalité ou une partie du réseau de fonctionner correctement.
On distingue deux types d'attaquants : le premier type est l'attaquant « Insider », c'est un utilisateur authentifié dans le réseau, il a une connaissance détaillée du réseau. Le deuxième type est l'attaquant « Outsider », il est considéré comme un intrus ayant une capacité d'attaque limitée, et il ne dispose d'aucun détail sur l'architecture ou le fonctionnement du réseau car il n'est pas authentifié. Quel que soit le type d'attaquant, les dégâts sont les même et peuvent être graves.
Les attaques DOS sont classées suivant la source de l'attaque, on a :

- Des attaques qui ne nécessitent pas de pénétrer le réseau, dans ce cas, L'attaquant peut lancer une attaque à distance sur le réseau ;
- Des attaques où l'attaquant exploite une certaine vulnérabilité connue pour pénétrer dans le réseau puis effectue des attaques de consommation des ressources ;

- L'attaque DOS distribué (DDOS) où les attaquants DDOS pénètrent ou compromettent de nombreux ordinateurs de tiers, appelés zombies et les utilisent pour lancer une attaque DOS contre le réseau cible.

Dans de nombreux cas, les propriétaires ne savent pas que leurs machines sont utilisées pour effectuer des attaques.

On peut trouver ce type d'attaques dans n'importe quel type de réseau, et aussi dans le réseau VANET. Les attaquants peuvent lancer une attaque directement sur les autres véhicules ou sur le RSU. Les services d'information liés à la circulation peuvent être perturbés et occupés lors d'une telle attaque. Ainsi, cette dernière produit des dommages sérieux aux utilisateurs humains, en les empêchant d'accéder aux ressources du réseau en temps réel.

3- Solutions et algorithmes de littérature

Pour lutter contre l'attaque DOS dans un réseau véhiculaire, plusieurs solutions et algorithmes ont été proposés et étudiés dans la littérature. Ces solutions présentent des idées permettant la détection et l'arrêt de cette attaque.

L'étude [14] propose l'algorithme RRDA (Request Response Detection Algorithm) pour détecter l'attaque DOS dans le réseau VANET. Les auteurs assument que tous les nœuds sont équipés par un ORT (On board Radio Transponder) pour pouvoir communiquer avec les autres nœuds dans le réseau. Ils forment un réseau de véhicules. Le réseau formé par les véhicules se compose de plusieurs nœuds et requiert une authentification. Le RSRT (Road Side Radio Transductor) décide et définit les véhicules qui peuvent former un réseau en fonction de leur portée de transmission. Ceci est considéré comme seuil. Les véhicules peuvent faire une demande au RSDT pour rejoindre le réseau créé en utilisant le mécanisme APDA (Attack Packet Detection Algorithm). Le RSDT vérifie les véhicules et crée une base de données. L'emplacement, l'horodatage, etc. sont fournis avec le paquet à l'RSDT. Le RSU utilise une autre base de données des demandes et des réponses et fournit des services uniquement à ces ORT qui sont déjà vérifiées, ce qui réduit l'attaque DOS.

Dans l'étude [15], les auteurs proposent l'algorithme APDA (Attacked Packet Detection Algorithme) pour détecter l'attaque DOS. Cet algorithme minimise le retard du traitement et améliore la sécurité dans le réseau VANET. Ce mécanisme est lié à chaque RSU. Les véhicules peuvent envoyer des messages au RSU à travers le mécanisme APDA afin de détecter les positions des autres véhicules. Après la détection de la position, les informations des autres véhicules sont stockées dans le RSU. Chaque véhicule a un OBU et un

Tamper Proof Device. Ces dispositifs stockent les informations détaillées sur les véhicules (la vitesse, la position, etc). Les positions des véhicules sont identifiées par la fréquence, la vitesse et l'utilisation de l'OBU.

Dans l'étude [16], les attaquants peuvent diffuser des messages forgés avec des signatures non valides pour forcer les véhicules récepteurs à effectuer de nombreuses vérifications inutiles des signatures. De cette façon les conducteurs ne peuvent pas vérifier l'authenticité des messages provenant d'autres véhicules. Le système proposé pour faire face à l'attaque DOS dispose d'un processus de pré-authentification avant le processus de vérification de signature. Ce processus de pré-authentification tire profit de la chaîne de hachage à sens unique et du schéma « Group rekeying scheme ».

L'étude [17] s'intéresse à « User Datagram Protocol (UDP) – based flooding », une forme courante de déni de service (DOS). Dans cette attaque, un nœud malveillant forge un grand nombre de fausses identités, par exemple « Internet Protocol (IP) spoofing adresses » pour perturber les fonctions propres du transfert équitable des données entre deux véhicules qui se déplacent rapidement. L'intégration de « IP spoofing » dans les attaques DOS rend la défense encore plus difficile. Les auteurs proposent une méthode efficace pour détecter et se défendre contre les attaques d'inondation « UDP » sous différents types d'usurpation d'IP. La méthode fait usage d'une structure de stockage de données et une méthode de « Bloom filter based IPCHOCKREFERENCE detection ». Cette approche rend relativement son implémentation facile à déployer et son coût est raisonnablement faible. Les principaux objectifs de ce travail sont :

- L'augmentation du rapport de distribution des messages,
- L'augmentation de la fiabilité et de la connectivité du système du véhicule (en diminuant le lien de rupture entre deux nœuds de communication),
- La réduction de la surcharge qui résulte de l'envoi des messages Beacon entre les nœuds en trouvant une nouvelle technique permettant d'organiser la diffusion de ces messages dans le réseau.

L'étude [18] propose une stratégie de refuge adaptée à l'antibrouillage dans le réseau VANET. Les véhicules choisissent entre deux voies de la route pour atteindre une destination particulière : route normale et route alternative. Lorsque la route normale est congestionnée en raison de la dynamique du transport, les véhicules sans réseau VANET se dirigent vers la route encombrée. Les véhicules équipés d'un OBU sont informés par le RSU ou des véhicules voisins à proximité pour prendre la route alternative. Si un brouilleur attaque la zone où les deux voies se séparent, les services du réseau VANET ne seront pas disponibles pour fournir les informations d'avertissement à temps réel pour empêcher les véhicules de prendre la route

encombrée. La nouvelle stratégie est perçue avec de nouvelles mesures de sécurité pour mesurer l'efficacité des brouilleurs et diriger la conception des mécanismes de défense. La façon dominante pour se défendre contre une attaque de brouillage est la stratégie de retraite : « channel surfing » pour basculer sur un autre canal lorsque la fréquence actuelle est bloquée et « spatial retreat » pour se déplacer sur un autre emplacement si la zone implique des interférences.

Dans l'étude [19], les auteurs présentent la méthode de "Bloom-filter-based detection method " qui prévoit la disponibilité d'un service pour les véhicules légitimes dans le réseau VANET, tel qu'il est utilisé, pour détecter l'attaque et se défendre contre l'usurpation d'adresses IP dans les attaques DOS. Ces attaques sont produites par des nœuds frauduleux et malveillants. La méthode proposée permet une communication sécurisée et permet également de libérer la bande passante. Cette approche nécessite moins de ressources et est facile à implémenter. Plus précisément, cette méthode fournit un temps de détection plus rapide et une capacité de stockage avec un coût de calcul amoindris. Le module « IP-Chock DoS attack detection » perçoit le trafic anormal en utilisant des algorithmes IP-cale. Supposons que beaucoup de véhicules circulent sur une route, chaque véhicule maintient sa vitesse dans un intervalle fixe, le processus se fait alors en trois principales phases qui sont : « detection engine phase 1 », « detection engine phase 2 », et « Bloom Filter phase ». La conséquence de ces procédés est basée sur la première étape pour détecter le changement par l'intermédiaire des capteurs montés sur le véhicule. Dans la deuxième étape, les valeurs de ces capteurs sont traitées afin de déterminer si les valeurs indiquent la possibilité d'influer sur le réseau. Une fois que la décision est prise, la troisième étape joue le rôle de la détection des attaques de déni de service dans l'infrastructure.

Dans l'étude [20], les auteurs ont proposé le modèle RBS protocol (Reference Broadcast Synchronization) pour la prévention de l'attaque DOS dans le réseau VANET. Ce modèle est basé sur le concept « Master Chock Filter » pour la filtration de paquets pendant que le trafic est intense ou pendant d'autres attaques. Le protocole a été évaluée par deux méthodes qui bloquent l'IP source origine de l'attaque DOS et contrôle la prévention de TCP / UDP flooding et attaque de IP Sniffing. L'évaluation du protocole est basée sur l'utilisation de la bande passante du nœud et l'interaction de sa mobilité. Le protocole RBS a montré que le taux de livraison de paquets, le débit, la temporisation sont améliorés par rapport au protocole IP-trackback.

Dans l'étude [21], le modèle d'automate stochastique SAN (Stochastic Automata Networks) a été développé pour décider si le réseau véhiculaire est sous l'attaque DOS ou non. Ce modèle a été mis en place en tant que solution pour les systèmes complexes avec un grand nombre d'états et de synchronisations complexes. Le modèle SAN est donné pour construire chaque sous-système par un automate et les interactions peuvent

être représentées par des bords dirigés. Ces bords peuvent être les taux de transition s'ils ne concernent que le sous-système, ou les synchronisations s'ils représentent les interactions entre les autres sous-systèmes. Deux types de transitions sont définis: la transition locale et la transition synchronisée. Une transition locale se produit uniquement dans l'automate alors que la transition synchronisée se produit dans plusieurs automates en même temps. Ainsi, un ensemble d'automate représente un SAN associé au système réel.

L'étude [22] discute plusieurs solutions cryptographiques pour plusieurs attaques possibles sur le réseau VANET dont deux solutions pour des attaques de « Jamming » et DOS sur la disponibilité du réseau (availability). Pour la première solution (citée en [23]), les auteurs proposent de changer le canal de transmission et d'utiliser la technique FHSS (Frequency Hopping Spread Spectrum) qui implique des algorithmes cryptographiques pour générer des nombres pseudo-aléatoires pour l'algorithme des sauts (FHSS). Cette proposition nécessite une modification de la norme utilisée. La deuxième solution est la même solution proposée dans l'étude [16]. La proposition est d'utiliser « Signature based authentication mechanisms » un mécanisme pour réduire l'effet de l'attaque DOS.

L'étude [24] discute aussi plusieurs solutions pour les attaques dans le réseau VANET. Parmi ces solutions, on trouve SEAD (Secure and Efficient Ad hoc Distance Vector). Les auteurs ont proposé un nouveau protocole de routage sécurisé qui protège contre plusieurs attaquants non coordonnés qui créent un routage incorrect dans un autre nœud. Ce protocole est basé sur le routage DSDV (Destination Sequenced-Distance Vector). SEAD prend en charge le nœud qui a une capacité de traitement du processeur limitée et le protège contre l'attaque DOS où les attaquants tentent d'utiliser la bande passante en excès. Il utilise la fonction de hachage à sens unique. L'autre solution est un protocole de routage qui empêche les attaques DOS dans le réseau. Cette approche utilise une cryptographie symétrique très efficace. L'émetteur et le récepteur se mettent d'accord sur deux clés: Ksr et Krs de l'émetteur au récepteur et du récepteur à l'expéditeur respectivement, en utilisant l'adresse MAC. Pour authentifier la demande de l'itinéraire, l'expéditeur envoie le message contenant des données uniques comme le « timestamp », calcule le MAC et l'envoie au récepteur en utilisant Ksr.

Dans l'étude [25], les auteurs proposent un modèle probabiliste innovant basée sur la régression logistique. Cette méthode permet d'estimer l'apparition d'une attaque. La méthode est basée sur une base de données qui estime les occurrences des attaques. Lorsque le modèle de la régression est validé, il est utilisé pour estimer la probabilité d'une attaque. Si cette probabilité dépasse le seuil défini à l'avance, l'attaque est donc confirmée.

4- Conclusion

Pour conclure, on remarque que l'attaque DOS est d'une grande importance dans le domaine de la recherche vu qu'elle abordée par une communauté importante des chercheurs. Les différents travaux s'intéressent à l'attaque sous différents angles, aux dégâts qu'elle peut causer et proposent des solutions.

Malgré les diverses solutions proposées, les chercheurs n'ont pas encore abouti à une solution qui arrête l'attaque DOS au moins de manière idéale. Dans le but de contribuer à trouver une solution efficace contre cette attaque, nous proposons dans le chapitre 3 notre papier scientifique, qui a été soumis à «The 15th ACM International Symposium on Mobility Management and Wireless Access ». La simulation réalisée pour l'environnement VANET, les différentes méthodes mathématiques utilisées, ainsi que les résultats de nos travaux à ce sujet seront présentés dans le chapitre 4. Ce travail permet d'une part de mieux comprendre les effets de cette attaque, et d'autre part de réaliser une analyse comparative des différentes méthodes de lutte contre ces attaques.

CHAPITRE 3
PAPIER SCIENTIFIQUE

1- Résumé du papier

1.1.Introduction

Notre papier scientifique "Using Mathematical Methods against Denial of Service (DoS) Attacks in VANET", aborde le sujet de VANET d'un point de vue sécurité et plus précisément la détection d'intrusions.

L'attaque DoS est toujours l'un des problèmes majeurs dans les réseaux informatiques, ce qui nécessite des recherches plus profondes pour mieux comprendre ses effets ainsi que le comportement du réseau sous ces effets.

La différence entre l'attaque DoS sur Internet et l'attaque DoS dans le VANET, c'est que sur un réseau internet, les nœuds sont des machines et dans le VANET, les nœuds sont des véhicules qui circulent dans la route à haute vitesse avec une topologie du réseau VANET qui change continuellement. Une telle attaque peut changer le comportement du conducteur, ce qui met en danger la vie des conducteurs ainsi que celle des passagers.

Notre papier est organisé somme suit : Dans la section I, on présente l'introduction, dans la section II, on discute l'état de l'art des dernières recherches reliées à l'attaque DoS et la section III, on fait une introduction des modèles mathématiques utilisés. Dans la section IV, on présente la description et l'analyse de données. Dans la section V, on présente les paramètres des simulations et les algorithmes utilisés. Dans la section VI, on discute les résultats et enfin une conclusion est présentée dans la section VII.

1.2.Objectif

L'objectif du papier est de fournir une étude globale et profonde sur l'attaque DoS dans le réseau VANET. On veut comprendre comment le réseau et ses entités réagissent au moment de l'attaque, à quel point une telle attaque peut perturber le fonctionnement du réseau et affecter la transmission des données, vu que dans le réseau VANET, il y a des types de données qui sont critiques et qui doivent être transmises en temps réel.

Notre étude se concentre sur les véhicules attaqués dans le réseau, plusieurs paramètres sont pris en considération afin de comprendre le comportement des nœuds avant et pendant l'attaque, ainsi que d'utiliser ces paramètres dans des méthodes mathématiques pour faire des prédictions et voir si on a suffisamment de données pour prédire si un nœud est sous attaque ou non.

Une telle étude va nous permettre dans des prochaines études, de se concentrer sur des différents aspects comme par exemple détecter la source de l'attaque, la stopper le plus tôt possible, etc.

1.3.Démarche et méthodologie

Le travail présenté s'est déroulé comme suit :

1- Simulation

2- Préparation des modèles et algorithmes de prédiction

3- Analyse des résultats

1.3.1. Simulation

Deux simulations ont été mis en place afin de les étudier et faire des comparaisons, la première s'est déroulé dans un environnement normal de VANET où les véhicules circulent et échangent les informations dans le réseau, la deuxième simulation s'est performée dans un environnement VANET sous l'attaque DoS afin d'obtenir des informations sur le réseau sous cette attaque.

Les simulations sont programmées avec le Framework Omnet++, les données sont extraites à la fin de chaque simulation afin de les utiliser dans nos analyses.

1.3.2. Préparation des modèles et algorithmes de prédiction

Plusieurs méthodes mathématiques ont été utilisées pour faire l'analyse de données, ainsi que de construire un modèle de prédiction pour prédire les véhicules attaqués et non attaqués.

L'analyse de données en utilisant la régression logistique et les statistiques descriptives ont été effectués à l'aide de XLSTAT, afin de comprendre comment les données sont dispersés, ainsi que la corrélation des paramètres avec le statut de chaque nœud (attaqué ou non attaqué). Trois algorithmes ont été implémentés sur MATLAB pour les méthodes MAV, MSE et RMS afin de les utiliser pour faire des prédictions. Aussi la régression logistique sur XLSTAT et le réseau de neurone de classification et reconnaissance implémenté sur MATLAB ont été utilisés comme des méthodes prédéfinies pour prédire et comparer les résultats.

1.3.3. Analyse des résultats

En ce qui concerne les résultats des simulations, on a pu extraire un total de 200 échantillons (nœuds) : 150 véhicules non attaqués, et 50 véhicules attaqués. Aussi, pour chaque nœud, on a les informations suivantes: Vehicle (identificateur du véhicule), Attacked (variable booléenne), Received Broadcasts (nombre de paquets reçus en mode broadcast), RXTXLostPackets (nombre de paquets perdus durant la transmission), SlotsBackoff (Nombre de slots dû au Backoffs), SNIRLostPackets (Nombre de paquets perdus à cause de la collision ou des erreurs), TimesIntoBackoff (Temps passé en mode Backoffs), TotalBusyTime (durée pendant laquelle le canal était occupé), TotalLostPackets (RXTXLostPackets + SNIRLostPackets)

Des analyses descriptives et de corrélations entre les paramètres ont été effectuées, les prédictions avec la méthode RMS donnent une incertitude de 0% pour les véhicules attaqués et 12.04 % pour les véhicules non attaqués. La méthode MAV montre une plus grande incertitude pour les véhicules non attaqués égale à 34.33% et 0% pour les véhicules attaqués. La méthode MSE est la plus exacte parmi les trois avec 0% pour les véhicules non attaquées et 2.38% pour les véhicules attaqués.

Les modèles de la régression logistique et le réseau de neurone pour la classification et reconnaissance montrent une très haute exactitude dans les résultats de prédiction, avec 100% d'exactitude pour les véhicules attaqués et non attaqués.

1.4.Conclusion

Le réseau VANET est un nouveau domaine de recherche en plein essor. Notre papier aborde un problème majeur qui est l'attaque DoS. L'environnement des VANETs est différent de celui d'Internet qu'on connait déjà, ainsi on a besoin de faire des études de différents aspects pour ce genre d'attaques, tout en prenant en considérations les différences entre les deux milieux étudiés.

Dans notre cas, nos études ont été réalisées d'un point de vue mathématique afin d'analyser le comportement du réseau et prédire parmi les véhicules, ceux qui sont sous attaque et ceux qui ne le sont pas. Les différentes méthodes utilisées nous ont permis d'avoir des résultats différents. Aussi, nous les avons comparées afin de choisir les meilleurs d'entre-elles.

Using Mathematical Methods against Denial of Service (DoS) Attacks in VANET

Youssef Lahrouni, Caroly Pereira, Amar Bensaber Boucif
Laboratoire de Mathématiques et Informatique Appliquées (LAMIA)
Departement of Mathematics and Computer Science
Université du Québec à Trois-Rivières
Trois-Rivières, QC, Canada
Youssef.Lahrouni@uqtr.ca, Caroly.Pereira@uqtr.ca, Boucif.Amar.Bensaber@uqtr.ca

***Abstract*: VANET is a novel technology that will sweep the world in the future; its purpose is to develop the vehicular environment and make it more comfortable. Also, it provides more safety on the road. Therefore, we have to make this technology as secured as possible against many threats. As VANET is a subclass of MANET, it has inherited many security problems with a different architecture and DOS attacks are one of them. In this paper, we focused on DOS attacks that prevent users to receive the right information at the right moment. We analyzed DOS attacks behavior and effects on the network using different mathematical models in order to find an efficient solution.**

Key words: VANET, security, DOS attack, mathematical models.

I- INTRODUCTION

Intelligent Transportation Systems (ITS) are applications that aim to provide innovative services for different modes of transport and traffic management. They allow users to be more informed about the road conditions and to drive safety, so drivers have more coordinated and more intelligent use of transport networks [1].

Vehicular Ad hoc Networks (VANET) are a key element of intelligent transport systems. They provide users with a lot of information they need during their journey such as meteorological information, road condition, Internet access or game in real time. Passengers can also enjoy it by using different applications of relaxation and entertainment.

VANET inherited several security issues, and so that users can enjoy it in the right way, in real time, and in private, they must offer the best possible security to safeguard identity and users' data.

One of the key VANET networks' utilities, is that users should have access to information in real-time, otherwise, it is useless because there is a lot of information that the user needs in real-time, and that have an important impact on his behavior. As a result, it influences his life and that of the passengers.

In this paper, we will focus on one of the most dangerous attacks on networks; the DOS attack (Denial of Service). Some attacker nodes generates mass of useless requests on the network to overload then block the transmission channel, this can prevent vehicles to receive critical messages from security applications, which prevents users to receive the right information at the right moment. What makes VANET DOS special, as compared to Internet DOS, is that in VANET, the vehicles are considered as nodes; therefore, the architecture can be different in any moment, especially because the nodes are moving continually in high speed. In addition, attacks in VANET can directly affect human life. That is why we should worry more about this kind of issues; also, the main purpose of VANET is to make the road condition more easy and comfortable for both drivers and passengers. Therefore, we will firstly simulate the attack on a Vehicular Ad hoc Network. Then, we will study the behavior of the network under the attack. Next, we will compare it with a normal network without attack. Finally, we will use several mathematical methods to predict if a node is under an attack or not. This will ease the task of offering solutions in future studies.

The use of mathematical methods is necessary to get some new insights of the problem; this could help us to minimize our field of study by predicting on which nodes the attack is performed. This also can help us in future studies to analyze from where or from whom the attack is done.

The remainder of this paper in organized as follow: In section II, we discuss the state of art on Dos attacks. In section III, we introduce Mathematical models used. In section IV, we present the description and data analysis. Section V presents the parameters of simulation and the algorithms used. In section VI, we discuss the results and we will conclude in section VII.

II- RELATED WORKS

In [2], the authors propose an algorithm called RRDA (Request Response Detection Algorithm) for detecting DOS attacks in VANET. They assume that all nodes are equipped with an ORT (On board Radio Transponder) to communicate with other nodes in the network. The RSRT (Road Side Radio Transductor) defines vehicles that can form a network based on their transmission range; this is considered as a threshold. The vehicle sends a request to the RSRT to join the created network using the APDA mechanism (Packet Attack Detection Algorithm). The RSRT checks each vehicle and creates a database. Location, time stamp, etc. ... are provided with the packet to RSRT. The roadside device uses another database of requests and responses then provides services only to that ORT which is verified in order to reduce DOS attacks.

In [3], the authors propose APDA algorithm (Attacked Packet Detection Algorithm) which is used to detect the DOS attack before the verification time. This minimizes the delay of treatment and improves safety in the VANET. This mechanism is attached to each RSU. Vehicles can send messages to the RSU through the APDA mechanism to detect certain vehicle positions. Each vehicle contains an OBU and Tamper Proof Device (TPD). These devices store detailed information about the vehicles such as speed, position etc. The proposed algorithm is based on position changing requirements. Attacked packets are identified by the frequency (number of broadcast packets per second), velocity, α (determined by the road characteristics and Xmax (maximum speed).

$$f = \alpha * |V - \frac{V_{max}}{2}|$$

f and V are high because the position will change quickly, f and V are low because the position will not change much. The proposed algorithm based on the change of position, frequency, and the velocity.

In [4], the authors say that in such an attack (DoS), attackers can spread forged messages with invalid signatures to force the receiver's vehicles to execute many unnecessary signatures' checks and so these vehicles cannot check messages from other legitimate vehicles. To deal with this kind of attack, the proposed system performs a process of pre-authentication before the signature verification process starts.

In [5], the authors discuss "User Datagram Protocol (UDP) - based flooding", which is a common form of denial of service (DoS) attack, in which a malicious node forges a large number of false identities, for example "Internet Protocol (IP) address spoofing" to disrupt the proper functions of the equitable transfer of data between two vehicles moving rapidly. The integration of "IP spoofing" in DoS attacks makes it even more difficult to defend against such an attack. An effective method is proposed to detect and defend against flood attacks under various types of IP spoofing. The method makes use of an efficient storage data structure and a method of "Bloom filter based IPCHOCKREFERENCE detection". This lightweight approach makes its relatively easy implementation to deploy as it costs reasonably low.

In [6], the authors propose a refuge strategy, suitable for anti-jamming in VANET. Vehicles choose between two lanes of the road to reach a particular destination: Normal Route and Alternative Route. When the normal route is congested because of the dynamics of transport, vehicles without VANET head to the congested road because it is the shortest but vehicles equipped with an OBU will be informed by the RSU or neighbors nearby vehicles to take the alternative route. If a jammer attacks the area where the two lanes divided, VANET services will not be available to provide warning information in real time to prevent vehicles to take the congested road. The new strategy is implemented with new security measures as packet send ratio (ratio of the number of packets that a node actually transmits to the number of total packets that it intends to send), segregation (quantifies how nodes spread across different channels) to assess the effectiveness of jammers, directing the design of defense mechanisms. The dominant way to defend against a jamming attack is the exit strategy, "channel surfing" to switch to another channel when the current frequency is blocked and "spatial retreat" to move to another location if the area involves interference.

In [7], the authors present the method of "Bloom-filter-based detection method" which provides the availability of a service for legitimate vehicles in VANET, as used to detect and defend against spoofing IP addresses in DoS attacks. The proposed method provides secure communication and also frees bandwidth. The module "IP-Chock DoS attack detection" perceives abnormal traffic. Every vehicle on the road keeps its speed in a fixed interval. There are three main phases of the process: 1/detection engine Phase, 2/detection engine and 3/Bloom Filter phase. The consequence of these processes is based on the first step to detect the change through sensors mounted on the vehicle. In the second step, treating the values of these sensors to determine whether the values indicate the possibility of influencing the network. Once the decision is made, the third stage serves as the detection of infrastructure denial of service attacks.

In [8], the authors proposed a new model called RBS protocol (Reference Broadcast Synchronization) for the prevention of DOS attacks in VANET. The proposed model is based on a concept called "Master Chock Filter" for packet filtering when the traffic is intense. The protocol was also evaluated by two other methods, which block the source IP of the original DOS attack and preventing control of TCP / UDP flooding attack and IP Sniffing. The evaluation of the protocol is based on the use of the bandwidth of the node and the interaction of its mobility. RBS protocol shows that the packet delivery rate, the throughput and the delay are improved compared to IP-protocol trackback.

In [9], the authors developed a stochastic model SAN (Stochastic Automata Networks) to decide whether the vehicular network is under DOS attack or not. SAN was developed as a solution for complex systems with a large number of states and complex synchronizations. SAN yield to build each subsystem by an automaton; interactions are represented by directed edges. These edges can be transition rates if they are for the subsystem or synchronizations if they represent the interactions between other subsystems. Two types of transitions are defined: local transition and synchronized transition. A local transition occurs only in the automata, while the synchronized transition occurs in multiple automata simultaneously. Thus, a set of automaton represents a SAN associated to the reel system.

In [10], the authors discuss several cryptographic solutions for several possible attacks on VANET; two of them are for the attacks of "Jamming" and DOS on network availability. For the first solution, the authors propose for such attack changing the transmission channel and to use FHSS (Frequency Hopping Spread Spectrum), which involves cryptographic algorithms to generate pseudo-random numbers for the algorithm hopping (FHSS). This proposal requires an amendment of the used standard. The second solution is the same in [4] used by authors Li He Wen and Tao Zhu. They propose to use "signature based authentication mechanisms". The mechanism reduces the effect of the DOS attack.

In [11], the authors discuss several solutions for attacks in VANET, one of these solutions is SEAD (Secure and Efficient Ad hoc Distance Vector). The authors proposed a new secure routing protocol that protects against several uncoordinated attackers who create incorrect routing to another node. It's based on the routing protocol DSDV (Destination Sequenced Distance-Vector). SEAD supports node that has a limited processing capacity and protects it against DOS attack where attackers attempt to use bandwidth in excess. The alternative is an approach that uses a very efficient symmetric cryptography. The sender and receiver agree on two keys saying Ksr and Krs using the MAC address. To authenticate the request of the route, the sender sends the message containing unique data such as timestamp, calculates the MAC and sends it to the receiver using Ksr.

To improve the method of corroboration attacks, the authors in [12] propose an innovative probabilistic model based on logistic regression. This method estimates the occurrence of an event called an attack. The method is based on a knowledge base that considers the occurrences of attacks. When the model of the regression is validated, it will be used to estimate the probability of an attack and if it exceeds the threshold set in advance, the attack is confirmed.

We can notice that the DOS attack is addressed in several researches, either to detect it, to stop it or to limit it, but it remains without a definitive solution. In the following, we will focus our study on this attack in order to explore it and to better understand its effects. Also, we will present analysis with several mathematical models and we will compare them. Unlike the studies cited on the related works, mathematical analysis will help us to manipulate those kinds of problems from a different angle of view, and handling any operation that seems to overload the network. Our main problematic reflected on how to detect and stop the DOS attack. Therefore, we divided our work on two main sections: the first section is to simulate a vehicular network environment under the DOS attack and after we will collect the output data to create a database. The second section includes analysis and comprehension of this data and makes conclusion to understand the reactions and network behavior in response to this attack by applying formulas and mathematical models. Therefore, for the second section, we will use analysis and prediction methods and formulas as root mean square, mean absolute value, mean squared error, logistic regression and neural networks.

III- MATHEMATICAL MODELS

In this section, we will present different methods we used to analyze our data. First, we will start with the description of the data using XLSTAT, and then we will make predictions using formulas of mean squared error, logic regression and neural networks.

A. MEAN SQUARED ERROR METHOD:

In statistics, the mean squared error (MSE) or mean squared deviation (MSD) of an estimator measures the average of the squares of the errors or deviations: that is, the difference between the estimator and what is estimated. MSE is a risk function, corresponding to the expected value of the squared error loss or quadratic loss. The MSE is a measure of the quality of an estimator: it is always non-negative, and values closer to zero are better [30]. The lower the means squared error, the closer you are getting to find the best fit line. It depends on the data.

$$\mathrm{MSE} = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(x_j - \bar{x})^2}$$

B. ROOT MEAN SQUARE AND MEANS ABSOLUTE VALUE METHODS:

For the RMS and MAV methods, we will use its formula as described in [15].

$$RMS = \sqrt{\frac{1}{N}\sum_{n=1}^{N} x_n^2}$$

$$MAV = \frac{1}{N}\sum_{n=1}^{N} |x|$$

For the mean squared error algorithm, the main idea is that if the value of the mean squared error between two vehicles is very low, then both vehicles represent the same information and vice versa. Therefore, each data row will be considered as a vector and the mean squared error formula will be applied between each row of the training data of not attacked vehicles and all vehicles in the database and will be saved. After, we will look for the minimum of these values for each vehicle and save it. Afterwards, we will do the same treatment between the training data of the attacked vehicles and all the vehicles in the database. Thus, for each vehicle in the database, we have its minimum mean squared error with the attacked vehicles and with the not attacked ones. Finally, one loop to traverse all the values, and to compare the two mean squared errors of all the vehicles, to decide to which category belongs the selected node.

For the RMS and MAV methods, we will use the same algorithm. The RMS is calculated for each vehicle in the database, also, for the attacked and the unattacked training samples and the minimum value is used to decide to which group belongs each node. Figure 1 shows the process from simulation to data separation.

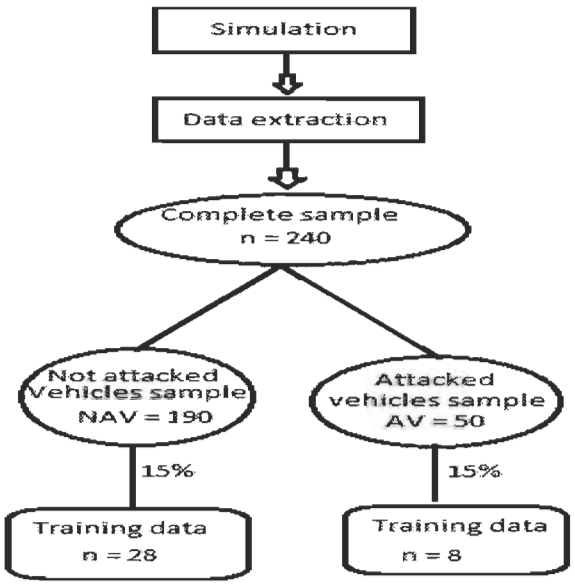

Figure 1: Extracting data scheme.

C. LOGISTIC REGRESSION MODEL:

Logistic regression is the convenient regression analysis when we have a binary dependent variable. The logistic regression is a predictive analysis; it is also used to describe data and to explain the relationship between one dependent binary variable and one or more metric independent variables.

We will use XLSTAT to generate results for analysis and estimations.

D. NEURAL NETWORK MODEL:

For the neural network, we will use the pattern of recognition and classification implemented on Matlab.

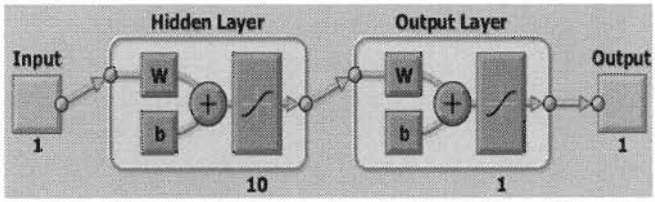

Figure 2: Network used.

For our network, we used 15% of information as training data, 15% for validation and 10 hidden layers. (fig.2)

IV- SIMULATION AND ALGORITHM

Regarding the simulation part, we have two simulations using Omnet ++, both have a duration of 90 seconds. The difference is that the first simulation is performed in a normal vehicular environment (network without attack), while the second simulation was performed in a network under the DOS attack. This attack was carried out between vehicles, e.g. an attacker node can attack vehicles within its covered area, so the attacked nodes cannot handle the mass of received requests. The same parameters were applied in both simulations; so both environments are identical. We were able to extract a set of 200 samples (nodes), 150 unattacked nodes, and 50 attacked nodes (fig.1).

V- ANALYSIS AND DATA DESCRIPTION

Each variable used in the simulation is presented below in table 1.

TABLE 1: Parameters definitions.

Variable	Definition
Vehicles	Vehicle ID.
Attacked	Binary variable to indicate whether the node was attacked or not.
ReceivedBroadcasts	Number of Received Broadcasts.
RXTXLostPackets	Number of lost packets during transmission.
SlotsBackoff	Number of time slots due to backoffs.
SNIRLostPackets	Lost packets per vehicle, i.e. packet collisions and packets not received due to bit errors.
TimesIntoBackoff	Number of times a vehicle was backoff during the simulation.
TotalBusyTime	Collected by the MAC layer and indicates the total time the wireless channel was busy.
TotalLostPackets	Total number of lost packets

	during the simulation.

A. DESCRIPTIVE STATISTICS

Variable	Modalités	Effectifs	%
Attacked	0	190	79,167
	1	50	20,833

Figure 3: descriptive statistics.

Statistique	ReceivedBroadcasts	RXTXLostPackets	SlotsBackoff	SNIRLostPackets	TimesIntoBackoff	totalBusyTime	TotalLostPackets
Nb. d'observ	240	240	240	240	240	240	240
Minimum	2,720	0,000	0,000	0,000	1,000	0,001	0,000
Maximum	80,000	47,840	27,710	27,370	8,340	0,017	71,410
Eff. du minin	1	163	12	42	131	1	40
Eff. du maxir	3	1	1	1	1	1	1
1er Quartile	11,000	0,000	3,000	2,000	1,000	0,002	2,000
Médiane	15,000	0,000	6,000	6,000	1,000	0,002	6,000
3ème Quarti	19,000	1,000	9,250	10,000	2,000	0,003	13,438
Moyenne	16,174	5,253	7,372	6,667	1,970	0,003	11,920
Variance (n)	113,875	118,352	38,608	33,254	2,282	0,000	234,759
Ecart-type (n	10,671	10,879	6,214	5,767	1,511	0,002	15,322

Figure 4: Descriptive statistics.

Variables	Attacked	ReceivedBroadcasts	RXTXLostPackets	SlotsBackoff	SNIRLostPackets	TimesIntoBackoff	totalBusyTime	TotalLostPackets
Attacked	1	-0,687	0,849	0,691	0,543	0,756	0,001	0,701
ReceivedBro	-0,687	1	-0,540	-0,494	-0,242	-0,499	0,557	-0,376
RXTXLostPac	0,849	-0,540	1	0,605	0,579	0,652	0,057	0,719
SlotsBackoff	0,691	-0,494	0,605	1	0,465	0,722	0,002	0,567
SNIRLostPacl	0,543	-0,242	0,579	0,465	1	0,488	0,335	0,965
TimesIntoBa	0,756	-0,499	0,652	0,722	0,488	1	0,052	0,590
totalBusyTin	0,001	0,557	0,057	0,002	0,335	0,052	1	0,297
TotalLostPac	0,701	-0,376	0,719	0,567	0,965	0,590	0,297	1

Figure 5: Correlation matrix.

In the descriptive statistics (Fig. 3 and Fig. 4), for all the variables, we have the same number of observations and in this analysis, we can see the minimum and the maximum of each variable for each vehicle. First quartile separates the bottom 25% of the data, the median cut the data set into two equal parts, so it is the midpoint of all, and third quartile separates the top 25% of the data. To do this, we sort the data in ascending order. After, we calculate the average of each variable. For the variance, a low variance means that the values are close to each other while a high variance means that they are widely spaced. The standard deviation is the square root of the variance, to measure the degree of dispersion of a set of data. In our case, the biggest variance values and standard deviations are found in "ReceivedBroadcasts", "RXTXLostPackets" and "TotalLostPackets", so the data in these variables are very dispersed over other. It means that values are widely distributed, as the smallest values are in "SlotsBackoff", "SNIRLostPackets" and "totalBusyTime". When the standard deviation approaches zero, all values in the data set are close.

Correlation matrix (Fig. 5) shows correlations between simulation variables; we will study the intensity of the connection that may exist between these variables. In this case, we want to study the correlation between the variable "Attacked" and other variables, to see the effect of the attack on the behavior of the attacked vehicles. The strong correlation between variables shows that the two variables behave in the same way (in the case of a strong and positive correlation). As we can see, there is a

strong positive correlation between the variable "Attacked" and the variables "RXTXLostPackets", "SlotsBackoff", "SNIRLostPackets", "TimesIntoBackoff", "TotalLostPackets". So, when the variable "Attacked" is set to 1, the other variables increase and vice versa. There is also a strong but negative correlation between the variable "Attacked" and the variable "ReceivedBroadcasts"s, and "totalBusyTime". Therefore, when the variable "Attacked" takes the value 1, the others decrease and vice versa.

B. LOGISTIC REGRESSION ANALYSIS

Statistique	Indépendant	Complet
Observation	240	240
Somme des	240,000	240,000
DDL	239	233
-2 Log(Vraise	245,635	0,000
R^2(McFadder	0,000	1,000
R^2(Cox and S	0,000	0,641
R^2(Nagelkerl	0,000	1,000
AIC	247,635	14,000
SBC	251,116	38,364
Itérations	0	20

Figure 6: Goodness of fit statistics.

The adjustment coefficient (Goodness of fit statistics fig.6) measures the level of representativeness of the model compared to baseline. Does the model represent enough information contained in the original data? More the model represents enough data, the more the model is more accurate; the more R^2 (the proportion of variance) is high, the more the variance is explanatory. We try to evaluate if the variables bring significant information to explain the variability of the binary variable (Attacked).
In our case, we have R^2 (McFadden) = 1, R (Cox and Snell) = 0.641, and R^2 (Nagelkerke's) = 1. So we can conclude that we have a good model.

Statistique	DDL	Khi²	Pr > Khi²
-2 Log(Vraise	6	245,635	< 0,0001
Score	6	212,387	< 0,0001
Wald	6	0,000	1,000

Figure 7: Hypothesis test.

The null hypothesis test (fig. 7) is used if it is true or not. It is assumed that the initial model we built is not better than the model based on chance and we must show that it is wrong. The more the probability is low, the model is significant.
In our case, -2 Log (Vraisemblance) <0.0001 and Score <0.0001, while the Wald method gives a value of 1. It was noted that in all the results of the analysis, the method of Wald gives different results from other methods or it displays errors, perhaps the Wald method is not consistent with our model.

Source	DDL	Khi² (Wald)	Pr > Wald	Khi² (LR)	Pr > LR
ReceivedBro	1	##########	1,000	0,000	0,995
RXTXLostPac	1	##########	1,000	1965,814	< 0,0001
SlotsBackoff	1	##########	1,000	3604,365	< 0,0001
SNIRLostPacl	1	##########	1,000	3604,365	< 0,0001
TimesIntoBa	1	##########	1,000	3604,365	< 0,0001
totalBusyTin	1	##########	1,000	3604,365	< 0,0001

Figure 8: Analysis type III.

For the analysis of type III (fig.8), for each variable, we assume that it is not significant for the model. We remove it and we made the test. It creates the same model without this variable: if the probability Pr> LR (likelihood ratio) is lower than 0.0001, so the result is false, so the variable is important for the model. Here, all PR <LR (variable probabilities) are <0.0001 except for the variable "ReceivedBroadcasts" which is equal to 0.995.

Source	Valeur	Ecart-type	Khi² de Wald	Pr > Khi²	Wald Borne inf. (95%	ald Borne sup. (95%
ReceivedBro	-6,070	0,000				
RXTXLostPac	24,389	0,000				
SlotsBackoff	0,472	0,000				
SNIRLostPacl	-1,357	0,000				
TimesIntoBa	-5,646	0,000				
totalBusyTin	1,514	0,000				
TotalLostPac	0,000	0,000				

Figure 9: Standardized coefficients.

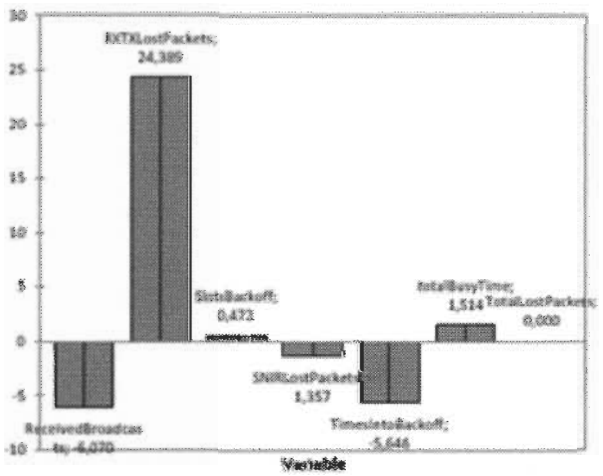

Figure10: Standardized coefficients diagram.

In the table of standardized coefficients (fig.9), we can see the importance or the relative weight of each variable compared to others. The more the absolute value of the variable is high, the more the variable is more significant and its weight is important. In our case, the variable "RXTXLostPackets" is the most important compared to the others.

VI- PREDICTION RESULTS

In the following, we will present the predictions results for each used method and then we will compare them to see which ones give the best results.

- ROOT MEAN SQUARE:

For Root Mean Square method (Fig. 11), we were able to predict the attacked vehicles with a 0% error rate. On the other hand, for not attacked vehicles, we have an error rate of 12.048%.

	Wrong estimated (%)	Well estimated (%)
Attacked vehicles	0.00	100
Non attacked vehicles	12.04	87.96

Figure 11: RMS results

- MEAN ABSOLUTE VALUE:

Using the MAV method (Fig. 12), we are able to predict the attacked vehicles with a 0% error rate, but for not attacked vehicles, we have a high error rate, which is equal to 34.337%, as can be seen in Picture below.

	Wrong estimated (%)	Well estimated (%)
Attacked vehicles	0.00	100
Non attacked vehicles	34.33	65.67

Figure 12: MAV result.

- MEAN SQUARED ERROR:

As we can see in fig. 13, all not attacked vehicles were properly classified (Error percentage = 0), whilst attacked vehicles were not perfectly classified, we got a small percentage error equals to 2.4%.

	Wrong estimated (%)	Well estimated (%)
Attacked vehicles	2.38	97.62
Non attacked vehicles	100	0.00

Figure 13: Mean squared error results.

- LOGISTIC REGRESSION RESULTS:

de \ Vers	0	1	Total	% correct
0	190	0	190	100,00%
1	0	50	50	100,00%
Total	190	50	240	100,00%

Figure14: Logic regression estimation.

As we can see in figure 14, we have 100% of correctness for the classification; all attacked and not attacked vehicles were perfectly classified.

- NEURAL NETWORK RESULTS:

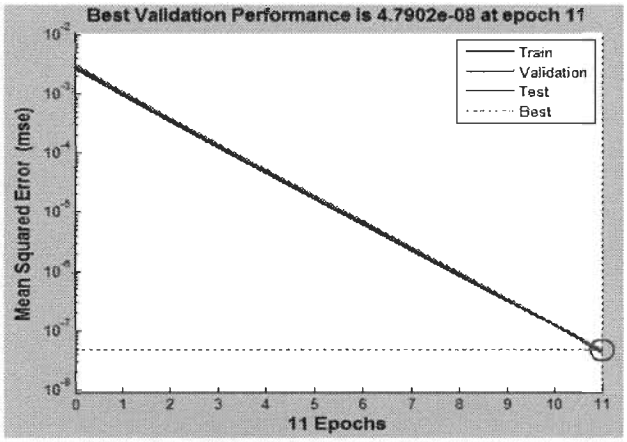

Figure 15: Performance plot.

The performance plot (fig. 15) shows how the network converged in a low air solution. This figure shows the various types of errors that occurred for the final trained network. In our case, it does not indicate any problem with the training; the plot shows a perfect training. The validation and test curves are similar. No overfitting occurred.

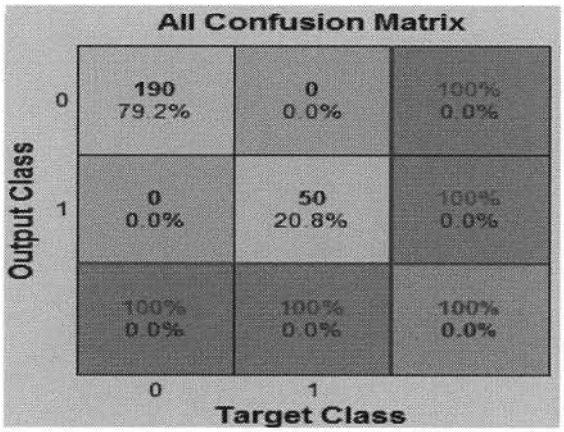

Figure 16: Confusion matrix.

The confusion matrix (fig.16) shows the number of vehicles (attacked and not attacked) which were properly classified in green squares, and it shows how many vehicles were improperly classified in red squares. The total percentages of correct and incorrect classifications are shown in the blue square.

As we can see, we have 100% of correct classifications and 0% of incorrect classification; this result is due to the significant effect of the DOS attack on the network, which extremely changed the behavior of many vehicles, which facilitates the classification.

To conclude, as shown from the results, we got good prediction results: neural network and logistic regression are more fitted for this kind of studies than the other methods.

VII- Conclusion and perspectives

Attacks in VANET are a paramount danger, as they threaten the privacy of drivers and passengers and they can cause accidents and congestion on the road. It is therefore very important to implement protocols and security mechanisms in order to control the entities of the network and to preserve the safety of drivers and passengers, as well as vehicles. DOS attack is approached in several researches, but it remains without a definitive solution.

In this work, we have studied this attack on VANET environment to understand its effects as well as the behavior of the network under such problems.

As can be seen from the results, this attack has a great effect on the network and especially that the vehicles lose a large number of packets during the transmission. It can prevent the vehicles from receiving important messages.

Several mathematics methods were used in this work. We have observed that logistic regression and neural network are better than MSE, RMS and MAV to analyze data and to predict attacks. Each parameter from the simulation was studied using the logistic regression in order to understand the effects in details. The prediction results by neural network and logistic regression show a high precision to detect the attack in the network.

In our future work, we will use this study to distinguish the attacked and non-attacked vehicles, as well as the deployment of the attacked vehicles can help to reduce our field of study to detect the source of the attack. Predictive methods can also be used to predict attacker vehicles.

REFERENCES

[1]https://en.wikipedia.org/wiki/Intelligent_transportation_system (25/07/2016)

[2] Usha Devi Gandhi, R.V.S.M Keerthana "Request Response Detection Algorithm for Detecting DoS Attack in VANET",International Conference on Reliability, Optimization and Information Technology, 2014, MRIU, India, Feb 6-8 2014. Electronic ISBN: 978-1-4799-2995-5.

[3] S.RoselinMary, M.Maheshwari and M.Thamaraiselvan. « Early Detection of DOS Attacks in VANET Using Attacked Packet Detection Algorithm (APDA) ». Information Communication and Embedded Systems (ICICES), 2013 International Conference, 21-22 Feb. 2013, Electronic ISBN: 978-1-4673-5788-3.

[4] Li He, Wen Tao Zhu « Mitigating DoS Attacks against Signature-Based Authentication in VANETs », IEEE International Conference on Computer Science and Automation Engineering (CSAE), 2012. Electronic ISBN: 978-1-4673-0089-6

[5] Karan Verma, Halabi Hasbullah, Ashok Kumar «An Efficient Defense Method against UDP Spoofed Flooding Traffic of Denial of Service (DoS) Attacks in VANET », 978-1-4673-4529-3/$31.00. Advance Computing Conference (IACC), 22-23 Feb. 2013 IEEE 3rd International. Electronic ISBN: 978-1-4673-4529-3.

[6] Ikechukwu K. Azogu, Michael T. Ferreira, Jonathan A. Larcom, Hong Liu « A New Anti-Jamming Strategy for VANET », Globcom Workshop – Vehicular Network Evolution. 2013. Electronic ISBN: 978-1-4799-2851-4

[7] Karan Verma, Halabi Hasbullah « IP-CHOCK (filter)-Based Detection Scheme for Denial of Service (DoS) attacks in VANET »,. International Conference on Computer and Information Sciences (ICCOINS), 2014. Electronic ISBN: 978-1-4799-4390-6

[8] Karan Verma, Halabi Hasbullah, Hemant Kumar Saini «Reference Broadcast Synchronization-Based Prevention to DoS attacks in VANET », Seventh International Conference on Contemporary Computing (IC3), 7-9 Aug. 2014. Electronic ISBN: 978-1-4799-5173-4

[9] Jalel Ben-Othman, Lynda Mokdad « Modeling and Verification Tools for jamming attacks in VANETS », Globcom 2014 – Wireless Networking Symposium. 2014. Electronic ISBN: 978-1-4799-3512-3

[10] Mohamed Nidhal Mejri, Jalel Ben-Othman, Mohamed Hamdi « Survey on VANET security challenges and possible cryptographic solutions », Vehicular communication 1 (2014) 53-66. 2014. Volume 1, Issue 2, April 2014, Pages 53–66.

[11] Adil Mudasir Malla, IndiaRavi Kant Sahu « Security attacks with an effective solution for DOS attacks in VANET ». International Journal of Computer Applications (0975 – 8887) Volume 66– No.22, March 2013.

[12] Ram Shringar Raw, Manish Kumar, Nanhay Singh « Security challenges, issues and their solutions for VANET » International journal of network security and its applications (IJNSA) 5(5):95-105. September 2013.

[13] Laroussi Karim, Amar Bensaber Boucif, Mesfioui Mhamed, Biskri Ismail « A probabilistic model to corroborate three attacks in Vehicular Ad hoc Networks », 2015 IEEE Symposium on Computers and Communication (ISCC), 6-9 July 2015. Electronic ISBN: 978-1-4673-7194-0.

[14] https://en.wikipedia.org/wiki/Mean_squared_error (01/11/2016)

[15] Angkoon Phinyomark, Chusak Limsakul, and Pornchai Phukpattaranont" A Novel Feature Extraction for Robust EMG Pattern Recognition" Journal of computing, Volume 1, ISSUE 1, pp 71-80 December 2009, ISSN: 2151-9617

CHAPITRE 4

SIMULATION ET RESULTATS

1- Introduction

Ce chapitre comporte trois sections principales. La première section consiste en l'exécution de la simulation de l'attaque DOS sur Omnet++ dans un environnement véhiculaire (réseau VANET). Les informations recueillies en sortie (résultats) créent une base de données. Cette base de données est analysée afin de comprendre les réactions et le comportement du réseau VANET sous l'effet de cette attaque. Les deux dernières sections consistent en l'application de certaines méthodes de prédiction de l'attaque DOS dans le réseau VANET et à la comparaison du degré de prédiction de ces méthodes. Les méthodes utilisées sont les méthodes d'analyse statistiques (erreur relative, RMS (Root Mean Square) et MAV (Mean Absolute Value), et la méthode des réseaux de neurones.

2- Simulation de l'attaque DOS

Deux simulations de 90 secondes sur Omnet++ sont réalisées. La différence entre ces deux simulations est que la première simulation est effectuée dans un environnement véhiculaire normal (réseau sans attaque), tandis que la deuxième simulation est effectuée dans un réseau sous l'attaque DOS. Cette attaque est effectuée entre les véhicules. Les deux simulations ont été réalisées en utilisant les mêmes paramètres. Un ensemble de 200 échantillons (nœuds) a été extrait : 150 nœuds non attaqués, et 50 nœuds attaqués.

2.1-Logiciels utilisés

2.2.1 - VEINS

Veins est un Framework open source pour exécuter des simulations de réseau de véhicules. Il est basé sur deux simulateurs:

- OMNeT ++ : c'est un simulateur de réseau basé sur les événements,

- SUMO : c'est un simulateur de trafic routier qui propose plusieurs modèles pour la simulation IVC (Inter Vehicular Communication).

2.2.2 - OMNET++

OMNeT ++ est un Framework discret, modulaire, orienté objet, permet de simuler des évènements réseau. Il a une architecture générique, de sorte qu'il peut être utilisé dans plusieurs domaines comme : la modélisation des réseaux de communication filaires et sans fil, la modélisation des protocoles, la modélisation de multiprocesseurs et d'autres systèmes de matériel distribués, la validation des architectures matérielles, etc. Dans notre cas, nous l'utilisons pour simuler un réseau véhiculaire sans fil (VANET).

2.2.3 - SUMO

SUMO (**S**imulation of **U**rban **M**obility) permet de générer la mobilité des véhicules et le réseau routier, ainsi que l'extraction d'une carte, la création du fichier du réseau routier, la création des fichiers de configuration, la création d'obstacles, etc.

2.2-Paramètres de simulation

Les paramètres utilisés pendant la simulation sont :

- **cmdenv-autoflush** : Paramètre de type booléen, par défaut FALSE, affecte à la fois le mode express et normal. Le fait de mettre cette variable TRUE peut engendrer des effets négatifs sur la performance, mais il peut être utile avec le débogage « printf » pour traquer les plantages du programme.
- **cmdenv-express-mode** : Paramètre de type booléen, par défaut TRUE, utilisé pour activer le mode express dans la simulation. Dans ce mode, le simulateur fournit le minimum des informations sur l'état de la simulation.
- **cmdenv-status-frequency** : Paramètre de type DOUBLE, son unité est la seconde, par défaut 2s, lorsque le paramètre cmdenv-express-mode est TRUE, l'état de la mise à jour est affiché chaque n secondes.

- **ned-path** : Paramètre globale, il s'applique sur tous les essais de simulation. Il est le chemin des répertoires considérés comme des racines du paquet NED. Cette option est normalement vide, mais pour les simulations qui commencent à l'extérieur de l'IDE, il est plus commode de la spécifier via une option de ligne de commande ou de variable d'environnement NEDPATH.
- **Network** : Paramètre de type String. Le nom du réseau à simuler. Le nom du package peut-être omis si le fichier « .ini » est dans le même répertoire que le fichier NED qui contient le réseau.
- **output-scalar-file** : Nom du fichier de type « .sca » qui contient les résultats de type scalaire de la simulation.
- **output-scalar-file-append** : Paramètre de type Booléen, par défaut False, permet de décider quoi faire lorsque le fichier scalaire de résultat existe déjà: soit fusionner les deux fichiers, ou le supprimer et commencer un nouveau fichier (c'est le choix par défaut).
- **debug-on-errors** : Paramètre de type Booléen. Si une erreur est détectée pendant l'exécution du programme, un point d'arrêt est généré de sorte que la vérification de l'emplacement et du contexte du problème dans le débogueur est possible.
- **sim-time-limit** : limite le temps de simulation en secondes.
- **print-undisposed** : Paramètre de type Booléen, par défaut True, utilisé pour signaler des objets laissés, non détruits par le destructeur après le nettoyage du réseau.
- **scalar-recording** : Paramètre de type Booléen, par défaut True, permet d'enregistrer les scalaires de sortie correspondants et les objets statistiques.
- **vector-recording** : Paramètre de type Booléen, par défaut True, permet d'enregistrer les données écrites dans le vecteur de sortie.

3- Analyse des résultats

Dans cette partie, nous analysons les données de plusieurs paramètres (variables). Ces données sont les résultats issus de la simulation de l'attaque DOS sur un réseau véhiculaire et de la simulation sans attaque. Nous appliquons sur ces données trois types d'analyses : l'analyse descriptive, la corrélation entre les variables et la régression logistique Multi-variée. Pour chaque analyse nous interprétons et discutons les résultats afin de comprendre le

comportement du réseau sous l'effet de cette attaque. Pour se faire, nous avons utilisé logiciel XLSTAT.

Le tableau 1 ci-dessous rassemble les variables de notre base de données :

Tableau1

Définition des variables

Variables	Définition
Vehicles	ID du véhicule.
Attacked	Variable binaire permet de mentionner si le nœud a été attaqué ou non.
ReceivedBroadcasts	Les broadcasts reçus.
RXTXLostPackets	Nombres de paquets perdus pendant la transmission.
SlotsBackoff	le nombre de slots de temps à cause des backoffs.
SNIRLostPackets	Lost packets per vehicle, i.e. packet collisions and packets not received due to bit errors.
TimesIntoBackoff	Le nombre de fois qu'un véhicule était en backoff pendant la simulation.
TotalBusyTime	Collected by the MAC layer and indicates the total time the wireless channel was busy.
TotalLostPackets	Le nombre total de paquets perdu durant la simulation.

3.1-Analyse descriptive des données

Les résultats de cette analyse sont présentés dans les deux figures ci-dessous (figure 6 et 7).

- Statistiques descriptives (Données quantitatives):

Statistique	ReceivedBroadcasts	RXTXLostPackets	SlotsBackoff	SNIRLostPackets	TimesIntoBackoff	totalBusyTime	TotalLostPackets
Nb. d'observ	240	240	240	240	240	240	240
Minimum	2,720	0,000	0,000	0,000	1,000	0,001	0,000
Maximum	80,000	47,840	27,710	27,370	8,340	0,017	71,410
Eff. du minin	1	163	12	42	131	1	40
Eff. du maxi	3	1	1	1	1	1	1
1er Quartile	11,000	0,000	3,000	2,000	1,000	0,002	2,000
Médiane	15,000	0,000	6,000	6,000	1,000	0,002	6,000
3ème Quarti	19,000	1,000	9,250	10,000	2,000	0,003	13,438
Moyenne	16,174	5,253	7,372	6,667	1,970	0,003	11,920
Variance (n)	113,875	118,352	38,608	33,254	2,282	0,000	234,759
Ecart-type (n	10,671	10,879	6,214	5,767	1,511	0,002	15,322

Figure 6 Statistiques descriptives (Données quantitatives)

Dans l'analyse descriptive, pour toutes les variables, on a le même nombre d'observations ainsi que le minimum et le maximum. Le 1er quartile est la donnée qui sépare les 25%

inférieurs des données. La médiane permet de couper l'ensemble des valeurs en deux parties égales, c'est le point milieu de l'ensemble. Le 3ème quartile est la donnée qui sépare les 25% supérieurs des données. On classe les données par ordre croissant. Ensuite, on calcule la moyenne de chaque variable. La variance et l'écart type mesurent toutes les deux la dispersion des valeurs autour de la moyenne. Pour la variance, une petite variance signifie que les valeurs sont proches les unes par rapport aux autres, alors qu'une variance élevée signifie que les valeurs sont très écartées. L'écart type est la racine carrée de la variance, donc si les valeurs possèdent une unité alors l'écart type s'exprime dans la même unité. Quand l'écart-type est proche de zéro, toutes les valeurs de l'ensemble de données sont proches (faible dispersion).

Dans notre cas, les valeurs de la variance, ainsi que celles de l'écart type, sont jugées relativement grandes dans les variables : « ReceivedBroadcasts », « RXTXLostPackets » et « TotalLostPackets ». Les données dans ces variables sont donc très dispersées les unes par rapport aux autres, c'est-à-dire, les valeurs sont largement distribuées. Les faibles valeurs de la variance sont obtenues dans les variables « SlotsBackoff », « SNIRLostPackets » et « totalBusyTime ».

- Corrélations entre les variables: Matrice de corrélation (Spearman) (figure 7):

Variables	Attacked	ReceivedBroadcasts	RXTXLostPackets	SlotsBackoff	SNIRLostPackets	TimesIntoBackoff	totalBusyTime	TotalLostPackets
Attacked	**1**	**-0,687**	**0,849**	**0,691**	**0,543**	**0,756**	0,001	**0,701**
ReceivedBro	**-0,687**	**1**	**-0,540**	**-0,494**	**-0,242**	**-0,499**	**0,557**	**-0,376**
RXTXLostPac	**0,849**	**-0,540**	**1**	**0,605**	**0,579**	**0,652**	0,057	**0,719**
SlotsBackoff	**0,691**	**-0,494**	**0,605**	**1**	**0,465**	**0,722**	0,002	**0,567**
SNIRLostPacl	**0,543**	**-0,242**	**0,579**	**0,465**	**1**	**0,488**	**0,335**	**0,965**
TimesIntoBa	**0,756**	**-0,499**	**0,652**	**0,722**	**0,488**	**1**	0,052	**0,590**
totalBusyTin	0,001	**0,557**	0,057	0,002	**0,335**	0,052	**1**	**0,297**
TotalLostPac	**0,701**	**-0,376**	**0,719**	**0,567**	**0,965**	**0,590**	**0,297**	**1**

Les valeurs en gras sont différentes de 0 à un niveau de signification alpha=0,05

Figure 7 Matrice de corrélation

L'intensité de la liaison qui peut exister entre les variables de la simulation est mesurée par l'étude de la corrélation (fig.7). La corrélation entre la variable « Attacked » et les autres variables, permet de voir l'effet de l'attaque sur le comportement des véhicules attaqués. La corrélation forte et positive entre deux variables montre que les deux variables se comportent de la même façon.

Dans notre cas, il y a une forte corrélation positive entre la variable « Attacked » et les variables «RXTXLostPackets», «SlotsBackoff», «SNIRLostPackets», «TimesIntoBackoff», «TotalLostPackets». C'est à dire lorsque la variable « Attacked » prend la valeur 1, les autres variables augmentent et vice versa.On a aussi obtenu une forte corrélation mais négative entre la variable « Attacked » et la variable « ReceivedBroadcasts », ce qui signifie que lorsque la variable «Attacked » prend la valeur 1, l'autre variable diminue et vice versa. Le dernier résultat tiré de l'analyse donne l'existence d'une faible liaison positive entre « Attacked » et « totalBusyTime ».

3.2-Analyse basée sur la régression logistique multi-variée

- Résultats et interprétations (figure 8):

Variable	Modalités	Effectifs	%
Attacked	0	190	79,167
	1	50	20,833

Figure 8 Statistiques descriptives

- Correspondance entre les modalités de la variable réponse et les probabilités (Variable « Attacked ») (figure 9):

Modalités	Probabilités
0	0
1	1

Figure 9 Correspondance entre les modalités de la variable réponse et les probabilités

La figure8 montre qu'on a un échantillon de 190 véhicules non attaqués, qui représente 79.167% de l'échantillon total, et 50 véhicules attaqués, soit 20.833% de l'échantillon total.

Dans la figure 9, la probabilité 0 signifie que la variable «Attacked» (dans fig.8) prend une valeur de 0, tandis que si la probabilité est 1 alors la variable «Attacked» prend la valeur 1.

Le coefficient d'ajustement (fig.10) permet de mesurer le niveau de représentativité (l'adéquation) du modèle par rapport aux données de départ. Est-ce que le modèle construit représente le maximum d'informations contenues dans les données de départ ? C'est aussi une mesure de la précision du modèle. Plus R^2 (coefficient de détermination, la proportion de variance) est proche de 1 (grande), plus la variance est explicative. L'évaluation des variables

permet d'avoir des informations significatives permettant d'expliquer la variabilité de la variable binaire (Attacked).

Dans notre cas, on a R^2(McFadden)=1, R^2 (Cox and Snell) = 0.641, R^2(Nagelkerke)=1. Donc, on peut conclure que notre modèle est bien adapté pour décrire la distribution des données de départ.

Coefficients d'ajustement (Variable Attacked		
Statistique	Indépendant	Complet
Observation	240	240
Somme des	240,000	240,000
DDL	239	233
-2 Log(Vraise	245,635	0,000
R^2(McFadder	0,000	1,000
R^2(Cox and S	0,000	0,641
R^2(Nagelkerl	0,000	1,000
AIC	247,635	14,000
SBC	251,116	38,364
Itérations	0	20

Figure 10 Coefficients d'ajustement

Le test d'hypothèse (fig.11) permet de tester si l'hypothèse nulle (H_0 : Y=0,208) est vraie ou non. C'est-à-dire, on teste si la performance du modèle de départ qu'on a construit est pratiquement similaire à celle du modèle basé sur le hasard. Ainsi, plus la probabilité calculée (Khi^2 de Pearsons) est plus petite que la valeur théorique de Khi^2, plus le modèle construit est plus précis.

Dans notre cas, les deux méthodes, -2 Log(Vraisemblance) et Score présentent des probabilités toutes les deux inférieures à 0,0001. Par contre la méthode de Wald donne une valeur de probabilité égale à 1. Également, on remarque que dans tous les résultats de l'analyse, les valeurs de Khi^2 relatives à la méthode de Wald sont très significativement différentes de celles des deux autres méthodes. Ceci permet de conclure que la méthode de Wald n'est peut-être pas compatible avec notre modèle proposé.

Test de l'hypothèse nulle H0 : Y=0,208 (Variable Attacked)			
Statistique	DDL	Khi²	Pr > Khi²
-2 Log(Vraise	6	245,635	< 0,0001
Score	6	212,387	< 0,0001
Wald	6	0,000	1,000

Figure 11 Test de l'hypothèse nulle

Dans l'analyse de type III (fig.12), on étudie l'importance des variables du modèle sur ses sorties. Pour cela, on enlève la variable concernée et on réalise le test d'hypothèse nulle (H0 : la variable n'a pas d'effet significatif pour le modèle). Dans chaque cas, si la probabilité calculée est plus petite que 0.0001 (Pr > LR), l'hypothèse nulle est rejetée, c'est-à-dire que, la variable est vraiment nécessaire pour le modèle et par conséquent on doit la remettre. Comme, on peut le voir dans le tableau 12, on remarque que toutes les probabilités (Pr < LR) des variables sont inférieures à 0.0001 sauf pour la variable « ReceivedBroadcasts » qui est égale à 0.995.

Analyse de Type III (Variable Attacked)					
Source	DDL	Khi² (Wald)	Pr > Wald	Khi² (LR)	Pr > LR
ReceivedBro	1	##################	1,000	0,000	0,995
RXTXLostPac	1	##################	1,000	1965,814	< 0,0001
SlotsBackoff	1	##################	1,000	3604,365	< 0,0001
SNIRLostPacl	1	##################	1,000	3604,365	< 0,0001
TimesIntoBa	1	##################	1,000	3604,365	< 0,0001
totalBusyTin	1	##################	1,000	3604,365	< 0,0001

Figure 12 Analyse de type III

Pour les coefficients normalisés (fig.13), on voit l'importance de chaque variable par rapport aux autres. Plus la valeur absolue de la variable est grande, plus la variable est significative et son poids relatif est important.

Dans notre cas, la variable « RXTXLostPackets » est la plus importante par rapport aux autres variables. XLSTAT prend un point de coupure égale à 0.5. Pour notre système, on voit que le nombre de 0 bien classé est de 190, le nombre de 1 bien classé est de 50, le nombre de faux

zéros classés est 0 et le nombre de faux 1 classés est 0. Donc 100% des zéros sont bien classés, et le pourcentage de bonnes prédictions pour 1 est de 100%. On a donc 100% d'exactitude. D'après ces résultats, on peut considérer que le modèle construit est de précision acceptable.

Coefficients normalisés (Variable Attacked) :

Source	Valeur	Ecart-type	Khi² de Wald	Pr > Khi²	Wald Borne inf. (95%	ald Borne sup. (95%)
ReceivedBro	-6,070	0,000				
RXTXLostPac	24,389	0,000				
SlotsBackoff	0,472	0,000				
SNIRLostPacl	-1,357	0,000				
TimesIntoBa	-5,646	0,000				
totalBusyTir	1,514	0,000				
TotalLostPac	0,000	0,000				

Figure 13 Coefficients normalisés

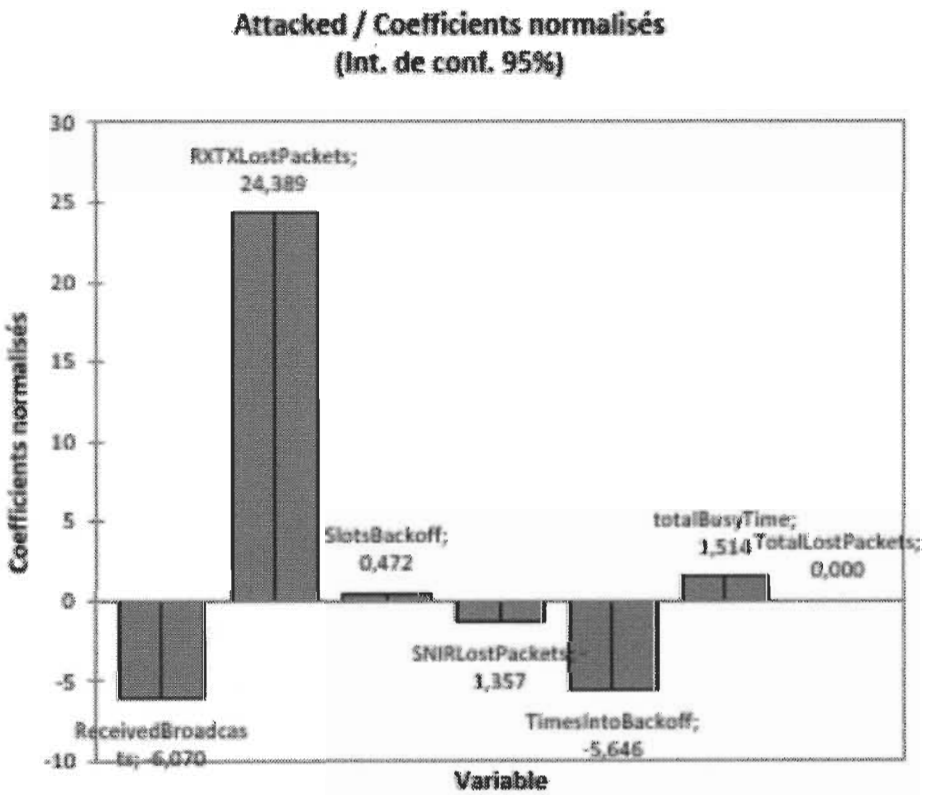

Figure 14 Diagramme des coefficients normalisés par rapport à la variable Attacked

de \ Vers	0	1	Total	% correct
0	190	0	190	100,00%
1	0	50	50	100,00%
Total	190	50	240	100,00%

Figure 15 Tableau de classification pour l'échantillon d'estimation (Variable Attacked)

Après révision des résultats, on remarque que la variable «RXTXLostPackets» qui est fortement corrélée avec la variable binaire «Attacked» (0.849) est la variable qui permet de donner ces prédictions avec une parfaite exactitude, car, après avoir classé les données en ordre croissant par rapport à la variable «RXTXLostPackets», on constate que les véhicules non attaqués ont très peu de paquets perdus par rapport aux véhicules attaqués.

3.3-Erreur relative, RMS, MAV

Afin de construire un système de prédiction efficace, on a choisi de travailler avec trois méthodes mathématiques. Ces méthodes sont utilisées dans un algorithme pour pouvoir prédire si un nœud est sous l'attaque DOS ou non.

Pour appliquer ces méthodes mathématiques sur nos données, on a besoin des données d'entrainement. On a pris un pourcentage de 15% du total des échantillons attaqués et 15% des échantillons non attaqués (figure 16), ce qui nous donne 23 véhicules non attaqués (VNA) et 8 véhicules attaqués (VA). Le logiciel MATLAB est utilisé pour la programmation.

Les méthodes ainsi que l'algorithme de prédictions utilisées sont définis et détaillés dans les paragraphes qui suivent.

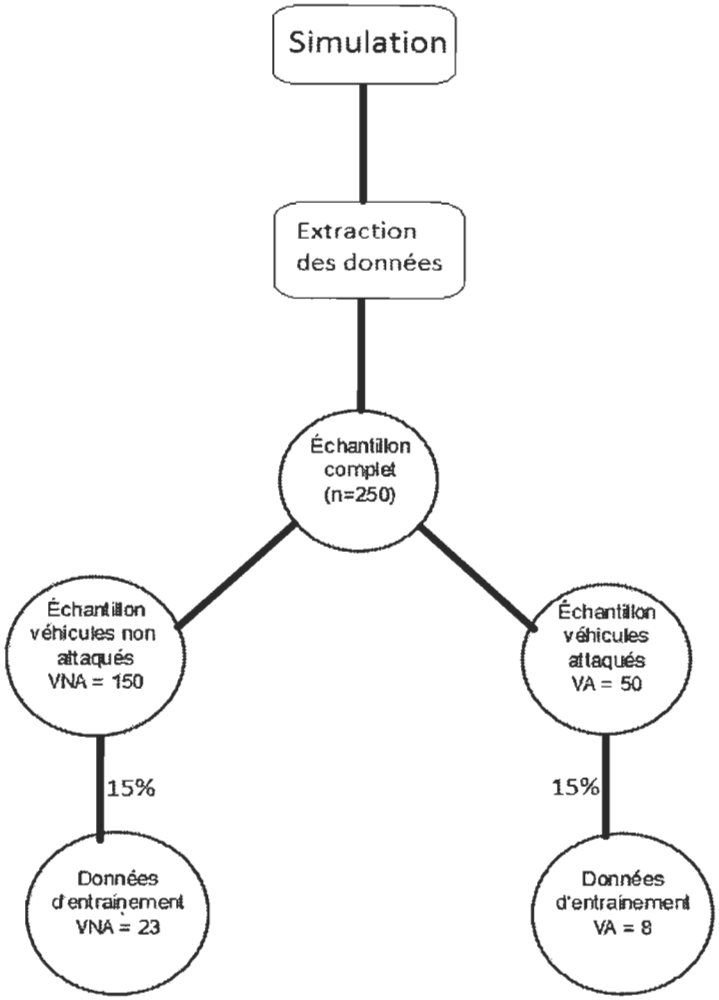

Figure 16 Présentation des données

3.4.1- Erreur relative

En statistique, l'erreur quadratique moyenne (MSE) ou l'écart quadratique moyen d'un estimateur (d'une procédure d'estimation d'une quantité non observée) mesure la moyenne des carrés des erreurs ou des écarts, c'est-à-dire, la différence entre l'estimateur et ce qui est estimé. La différence se produit en raison du caractère aléatoire ou parce que l'estimateur ne tient pas compte de l'information qui pourrait produire une estimation plus précise. La MSE est une mesure de la qualité d'un estimateur : elle est toujours positive, et les valeurs proches de zéro sont meilleures [30]. Plus l'erreur au carré est faible, plus on se rapproche de la ligne du meilleur ajustement. Ceci dépend des données.

$$\mathrm{MSE} = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(x_j - \bar{x})^2}$$

3.4.2- RMS

Dans les statistiques, la racine de la moyenne des carrés des écarts, RMS (Root Mean Square), connu aussi sous l'appellation de la moyenne quadratique, est définie comme étant la racine carrée de la moyenne arithmétique des carrés des écarts d'un ensemble de nombres. La RMS peut également être définit par une fonction continue variante en termes d'une intégrale des carrés des valeurs instantanées pendant un cycle [27], [28].

Dans le cas d'un ensemble de n valeurs$\{x_1, x_2, \dots, x_n\}$, l'expression analytique de la moyenne quadratique de x est :

$$x_{rms} = \sqrt{\frac{1}{n}(x_1^2 + x_2^2 + \cdots + x_n^2)}$$

Ou bien :

$$x_{rms} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}$$

3.4.3- MAV

Pour la méthode « Mean Absolute Value », la formule ci-dessous est utilisée telle qu'elle est définit dans [31].

$$MAV = \frac{1}{N}\sum_{n=1}^{N} |x|$$

3.4.4- Algorithmes

Un algorithme pour chacune des méthodes citées ci-dessus est écrit afin de trouver la meilleure méthode de prédiction. Le taux d'erreur de chaque méthode est calculé également.

Les fichiers nécessaires pour les calculs sont : le fichier contenant toutes les données, le fichier des échantillons de données d'entrainement des véhicules non attaqués et le fichier des échantillons de données d'entrainement des véhicules attaqués. Ces fichiers sont créés et importés. A noter que les données choisies comme données d'entrainement sont supprimées du fichier contenant toutes les données.

3.4.5.1- Algorithme erreur relative

Pour cet algorithme, on considère chaque ligne de données comme un vecteur et on calcule l'erreur relative entre chaque ligne de données d'entrainement des véhicules non attaquées et tous les véhicules dans la base de données. On les enregistre dans un fichier (matrice [208 x 23]), puis on cherche le minimum de ces valeurs pour chaque véhicule et on l'enregistre. Si la valeur de l'erreur relative entre deux véhicules est petite, les deux véhicules représentent donc les mêmes informations et vice-versa.

Ensuite on fait une boucle pour faire le même traitement entre les données d'entrainement des véhicules attaqués et tous les véhicules dans la base de données.

Pour chaque véhicule de la base de données, on enregistre l'erreur relative minimale des véhicules attaqués et des véhicules non attaqués. À la fin on fait une boucle pour comparer les erreurs relatives de tous les véhicules et on décide à quelle catégorie appartient le nœud sélectionné.

3.4.5.2- Algorithme RMS, MAV

Pour les méthodes RMS et MAV, on a utilisé le même algorithme que celui de l'erreur relative. La différence est qu'au lieu de chercher le minimum des erreurs relatives, on calcule le RMS pour chaque véhicule de la base de données avec les échantillons d'entrainement attaqués et non attaqués. On prend ensuite le minimum entre les deux pour déduire le groupe auquel appartient chaque nœud. La même procédure est réalisée pour la méthode MAV (figure 17).

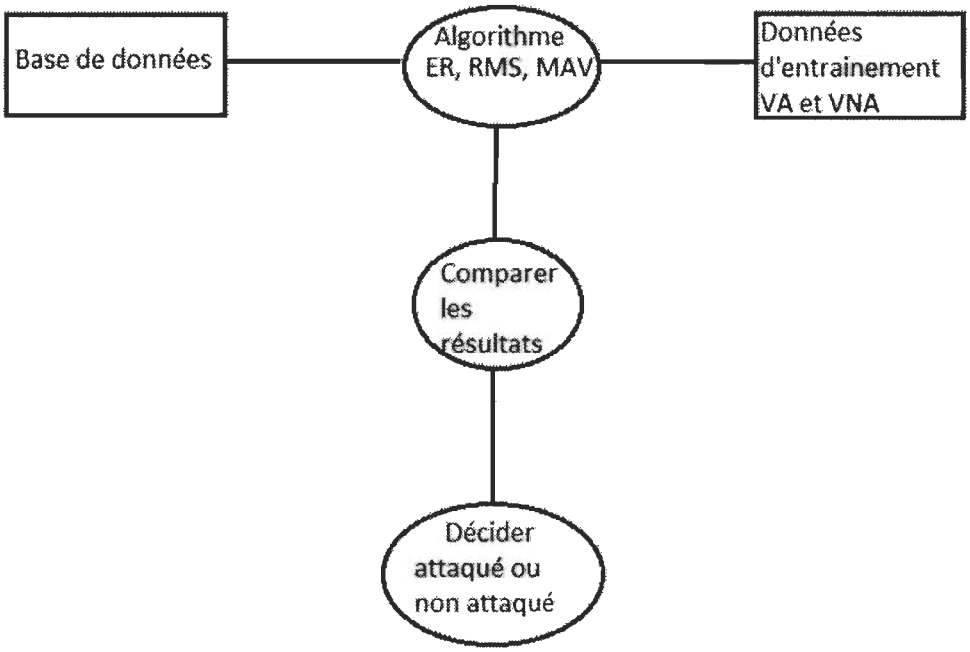

Figure 17 Les étapes de calcul

3.4.5- Résultats et discussions

a. Résultats erreur relative

Pour la méthode de l'erreur relative, le taux d'erreur obtenu est de 0% pour les véhicules non attaqués et de 2.38% d'erreur pour les véhicules attaqués (figures 18 et 19). Cette méthode permet bien de prédire les véhicules non attaqués.

```
%=== Résultats de la méthode : Erreur relative  ======

   - véhicule  1 non attaqué   -

   - véhicule  2 non attaqué   -

   - véhicule  3 non attaqué   -

   - véhicule  4 non attaqué   -

   - véhicule  5 non attaqué   -
```

Figure 18 Résultats de l'erreur relative

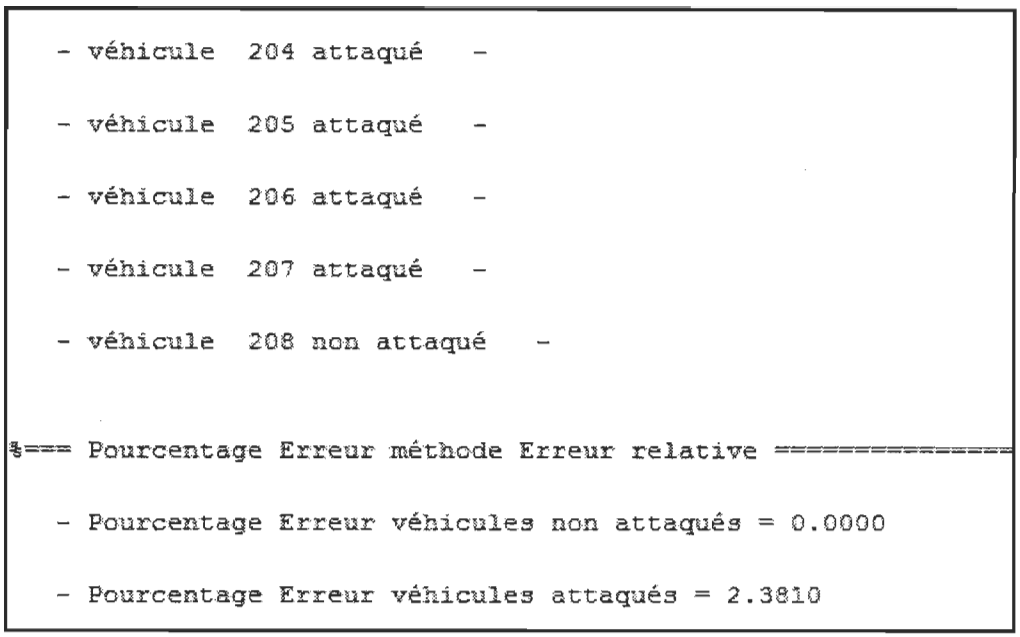

Figure 19 Taux d'erreur (Erreur relative)

b. Résultats RMS

La méthode Root Mean Square a permis de bien prédire les véhicules attaqués avec un taux d'erreur de 0%, par contre, pour les véhicules non attaqués le taux d'erreur est de 12.048% (figures 20 et 21).

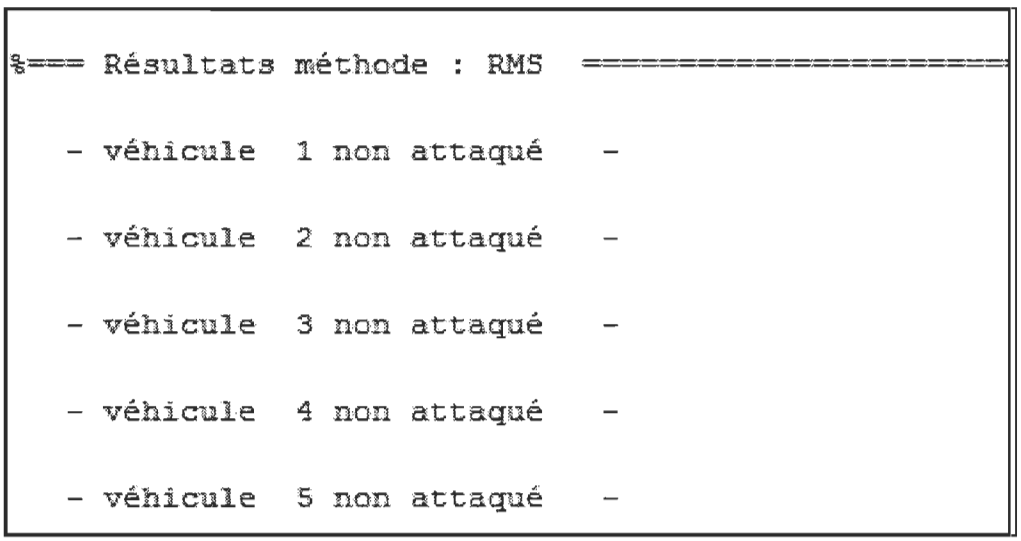

Figure 20 Résultats RMS

```
  - véhicule  204 attaqué   -
  - véhicule  205 attaqué   -
  - véhicule  206 attaqué   -
  - véhicule  207 attaqué   -
  - véhicule  208 attaqué   -

%=== Pourcentage Erreur méthode RMS ========================
  - Pourcentage Erreur véhicules non attaqués = 12.0482
  - Pourcentage Erreur véhicules attaqués = 0.0000
```

Figure 21 Taux d'erreur (RMS)

c. Résultats MAV

Avec la méthode de MAV, on a réussi à prédire les véhicules attaqués avec un taux d'erreur de 0%, mais pour les véhicules non attaqués le taux d'erreur est grand et est égal à 34.337%, comme le montre les figures 22 et 23.

```
%=== Résultats méthode : MAV  ============================
   - véhicule  1 non attaqué   -
   - véhicule  2 non attaqué   -
   - véhicule  3 non attaqué   -
   - véhicule  4 non attaqué   -
```

Figure 22 Résultats MAV

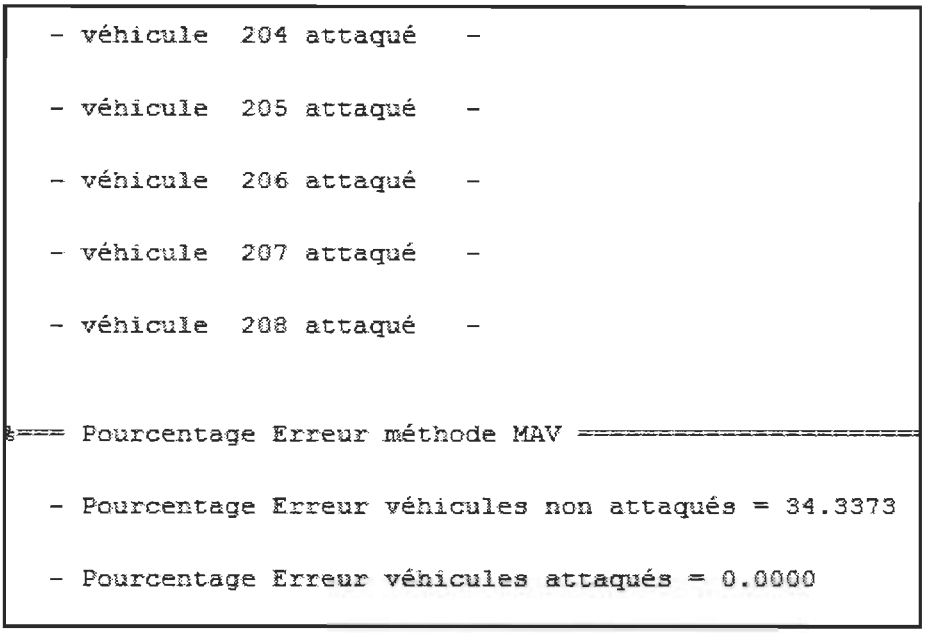

Figure 23 Taux d'erreur (MAV)

L'erreur relative, Root Mean Square, et Mean Absolute Value sont donc des méthodes simples à comprendre et à implémenter, pour cela il suffit de trouver le bon algorithme pour les utiliser et avoir de bons résultats.

La comparaison des résultats montre que la méthode de l'erreur relative est la meilleure. Tandis que les méthodes RMS et MAV présentent un taux d'erreur un peu élevé pour les véhicules non attaqués. Ceci peut être expliqué par l'incapacité de l'algorithme de prédire l'attaque dans ces conditions, surtout qu'un simple changement des données d'entrainement peut affecter le résultat.

4- Réseau de neurones

Dans cette partie du réseau de neurones, le pattern de reconnaissance et de classification implémenté dans MATLAB est utilisé.
Les 15% des données sont consacrées pour l'entrainement, 15% pour la validation, et le reste pour faire le test de prédiction. Notre modèle de réseau contient 10 couches (fig. 24).

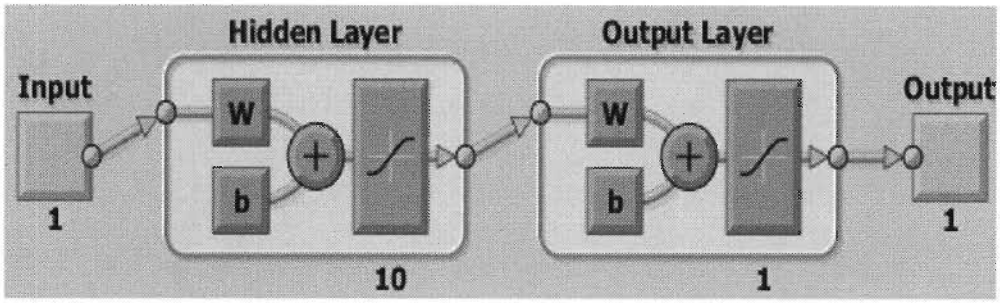

Figure 24 Réseau de neurone utilisé

4.1-Résultats

Le diagramme de performance (fig.25) montre que le réseau converge dans une solution d'air faible. Cette figure montre les différents types d'erreurs qui se sont produites dans le réseau formé final. Dans notre cas, ceci n'indique aucun problème d'entrainement; le diagramme montre qu'il y a eu une très bonne phase d'entrainement. Les courbes de validation et de test sont similaires.

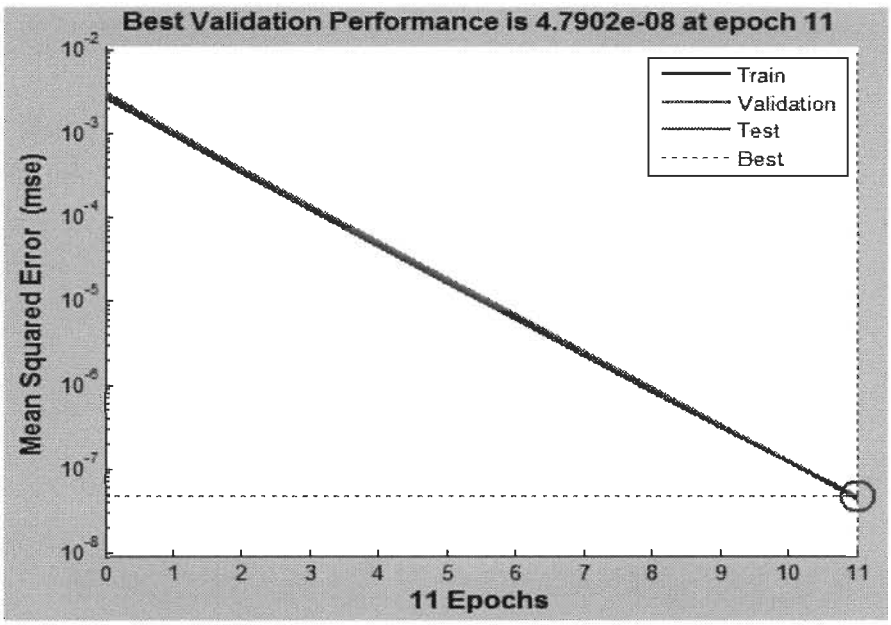

Figure 25 Performance du réseau

La matrice de confusion (fig.26) montre le nombre de véhicules (attaqués et non attaqués) qui ont été correctement classés en vert. Elle montre également le nombre de véhicules mal classés en rouge. Le pourcentage total des classifications correctes et incorrectes est indiqué en bleu. Nous avons 100% de classifications correctes et 0% de classification incorrecte. Ce résultat est dû à l'effet significatif de l'attaque DOS sur le réseau, qui a modifié le comportement de nombreux véhicules, ce qui a facilité la classification.

En conclusion et comme le montrent les résultats obtenus, les prédictions sont parfaitement satisfaisantes : le réseau de neurones et la régression logistique sont plus adaptés à ce type d'études que les autres méthodes.

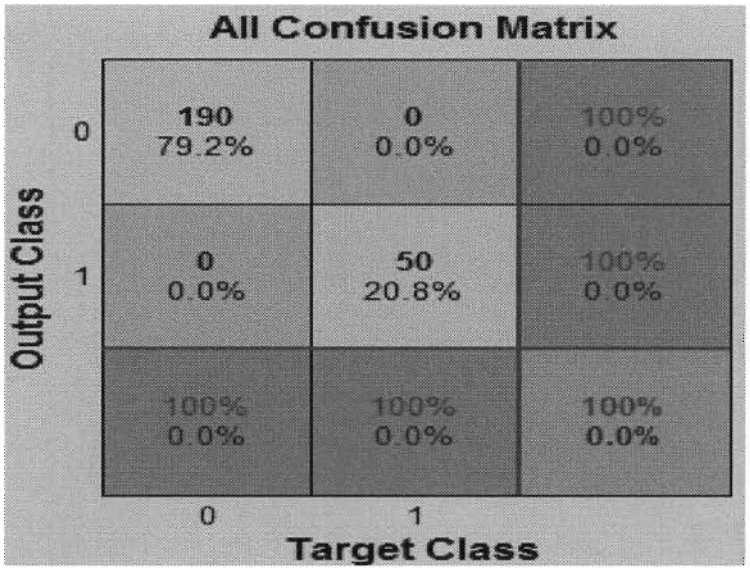

Figure 26 Matrice de confusion

5- Conclusion

La simulation de l'attaque DOS sur le réseau VANET montre un grand effet de l'attaque sur le réseau.

Les résultats de prédictions obtenus par la méthode de la régression logistique et la méthode du réseau neurones donnent une grande précision pour la détection de l'attaque DOS dans le réseau VANET par rapport aux autres méthodes statistiques.

CONCLUSION ET PERSPECTIVES

Les attaques dans le réseau VANET sont un danger primordial qu'il faut absolument détecter. Elles menacent la vie privée des conducteurs et des passagers et peuvent causer des accidents et de la congestion sur la route. Il est donc très important de mettre en œuvre des protocoles et des mécanismes de sécurité pour contrôler les entités du réseau afin de préserver la sécurité des conducteurs et des passagers, ainsi que des véhicules.

L'attaque DOS sur le réseau VANET est un sujet de recherche en plein essor dans de nombreux travaux, mais elle reste jusqu'à présent sans solution définitive.

Dans ce manuscrit, nous nous sommes intéressés à l'étude de l'attaque DOS sur l'environnement du réseau VANET pour comprendre ses effets ainsi que le comportement du réseau dans de telles circonstances.

Les méthodes de prédiction de l'attaque DOS que nous avons testé sont les trois méthodes mathématiques suivantes: la méthode de l'erreur relative, la méthode RMS et la méthode MAV, ainsi que les méthodes de la régression logistique et du réseau de neurones. Les résultats ont montré que les deux dernières méthodes permettent de mieux analyser les données et de mieux prédire l'attaque DOS que les autres.

Chaque paramètre de la simulation de l'attaque a été étudié en utilisant la régression logistique afin de comprendre en détails les effets de l'attaque DOS. Les résultats de prédictions obtenus par le réseau de neurones et la régression logistique donnent une grande précision pour détecter l'attaque dans le réseau VANET. Ces résultats montrent également que l'attaque DOS a un grand effet sur le réseau VANET. Les véhicules perdent un grand nombre de paquets pendant la transmission. L'attaque DOS est capable d'empêcher des véhicules de recevoir des messages importants.

Cette étude a permis de distinguer les véhicules attaqués et non attaqués et de prédire l'attaque DOS sur le réseau VANET, ce qui va nous permettre, dans notre futur travail, de réduire le champ d'étude afin de détecter la source de l'attaque et même de prédire les véhicules des attaquants en utilisant les même méthodes prédictives.

REFERENCES

[1] ADETUNDJA ADIGUN « gestion de l'anonymat et de la traçabilité dans les réseaux véhiculaires sans fil », mémoire de maîtrise, Département de mathématiques et informatique appliquées. Université du Québec à Trois-Rivières, Octobre 2013.

[2] ROMAIN COUSSEMENT « Mécanisme d'aide à la décision pour les ids dans les réseaux VANETs », mémoire de maîtrise, Département de mathématiques et informatique appliquées, Université du Québec à Trois-Rivières, Janvier 2014.

[3] AHIZOUNE AHMED, " Un protocole de diffusion des messages dans les réseaux véhiculaires", Mémoire de master. Département d'informatique et de recherche opérationnelle, Faculté des arts et sciences, Université de Montréal. Avril 2011.

[4] NOUREDDINE CHAIB, "La sécurité des communications dans les réseaux VANET", Mémoire de master, faculté des sciences de l'ingénieur département d'informatique. Université Elhadi Lakhdar - Batna, 05 Septembre 2011.

[5] JONATHAN PETIT, "Surcoût de l'authentification et du consensus dans la sécurité des réseaux sans fil véhiculaires", Thèse de Doctorat, Université de Toulouse, 13 Juillet 2011.

[6] http://hdhili.weebly.com/uploads/9/8/9/6/9896432/chap4_vanet.pdf (Janvier 2015)

[7] GHASSAN SAMARA, WAFAA A.H. AL-SALIHY, R. SURES, «Security Analysis of Vehicular AdHoc Networks (VANET) » Second International Conference on Network Applications, Protocols and Services 22-23 Sept. 2010. Electronic ISBN: 978-0-7695-4177-8.

[8] NURAIN IZZATI SHUHAIMI, TUTUN JUHANA « Sécurity in véhicular ad hoc with identity- based cryptographie approach: a survey », 7th international conference on telecommunication system, services and applications (TSSA) 30-31 Oct. 2012. Electronic ISBN: 978-1-4673-4550-7

[9] QINGZI LIU, QIWU WU, LI YONG « Hierarchical Security Architecture of vanet », Department of Information Engineering, Engineering University of Armed Police Force, China. International Conference on Cyberspace Technology (CCT 2013) 23-23 Nov. 2013. Online ISBN: 978-1-84919-801-1. Publisher: IET

[10] VINEETHA PARUCHURI« Inter-Vehicular Communications: Security and Reliability Issues », Vineetha Paruchuri, Department of Computer Science, RV College of engineering, Bangalore, India. Conference on ICT Convergence (ICTC), 2011 International. 28-30 Sept. 2011, Electronic ISBN: 978-1-4577-1268-5.

[11] GHASSAN SAMARA, WAFAA A.H. ALI ALSALIHY « A New Security Mechanism for Vehicular Communication Networks », 978-1-4673-1677-4. Cyber Security, Cyber Warfare and Digital Forensic (CyberSec), 2012 International Conference on. 26-28 June 2012. Electronic ISBN: 978-1-4673-1426-8.

[12] EMÍLLIA BUBENÍKOVÁ, JÁN DURECH, MÁRIA FRANEKOVÁ « Security Solutions of Intelligent Transportation System's Applications with using VANET Networks » Department of Information and Control Systems, University of Žilina. Control Conference (ICCC), 2014 15th International Carpathian. 28-30 May 2014, IEEE. Electronic ISBN: 978-1-4799-3528-4.

[13] ADETUNDJI ADIGUN, BOUCIF AMAR BENSABER, ISMAIL BISKRI « Protocol of Change Pseudonyms for VANETs », 9th IEEE International Workshop on Performance and Management of Wireless and Mobile Networks. 2013 IEEE 38th Conference on Local Computer Networks Workshops (LCN Workshops). 21-24 Oct. 2013. Electronic ISBN: 978-1-4799-0540-9

[14] USHA DEVI GANDHI, R.V.S.M KEERTHANA "Request Response Detection Algorithm for Detecting DoS Attack in VANET", International Conference on Reliability, Optimization and Information Technology, 2014, MRIU, India, Feb 6-8 2014. Electronic ISBN: 978-1-4799-2995-5.

[15] ROSELINMARY.S, M.MAHESHWARI AND M.THAMARAISELVAN « Early Detection of DOS Attacks in VANET Using Attacked Packet Detection Algorithm (APDA) ». Information Communication and Embedded Systems (ICICES), 2013 International Conference, 21-22 Feb. 2013, Electronic ISBN: 978-1-4673-5788-3.

[16] LI HE, WEN TAO ZHU « Mitigating DOS Attacks against Signature-Based Authentication in VANETs », IEEE International Conference on Computer Science and Automation Engineering (CSAE), 2012. Electronic ISBN: 978-1-4673-0089-6

[17] KARAN VERMA, HALABI HASBULLAH, ASHOK KUMAR «An Efficient Defense Method against UDP Spoofed Flooding Traffic of Denial of Service (DOS) Attacks in VANET », 978-1-4673-4529-3/$31.00. Advance Computing Conference (IACC), 22-23 Feb. 2013 IEEE 3rd International. Electronic ISBN: 978-1-4673-4529-3.

[18] IKECHUKWU K. AZOGU, MICHAEL T. FERREIRA, JONATHAN A. LARCOM, HONG LIU « A New Anti-Jamming Strategy for VANET », Globcom Workshop – Vehicular Network Evolution. 2013. Electronic ISBN: 978-1-4799-2851-4.

[19] KARAN VERMA, HALABI HASBULLAH « IP-CHOCK (filter)-Based Detection Scheme for Denial of Service (DoS) attacks in VANET ». International Conference on Computer and Information Sciences (ICCOINS), 2014. Electronic ISBN: 978-1-4799-4390-6.

[20] KARAN VERMA, HALABI HASBULLAH, HEMANT KUMAR SAINI «Reference Broadcast Synchronization-Based Prevention to DoS attacks in VANET », Seventh International Conference on Contemporary Computing (IC3), 7-9 Aug. 2014. Electronic ISBN: 978-1-4799-5173-4.

[21] JALEL BEN-OTHMAN, LYNDA MOKDAD « Modeling and Verification Tools for jamming attacks in VANETS », Globcom 2014 – Wireless Networking Symposium. 2014. Electronic ISBN: 978-1-4799-3512-3

[22] MOHAMED NIDHAL MEJRI, JALEL BEN-OTHMAN, MOHAMED HAMDI « Survey on VANET security challenges and possible cryptographic solutions », Vehicular communication 1 (2014) 53-66. 2014. Volume 1, Issue 2, April 2014, Pages 53–66.

[23] ADIL MUDASIR MALLA, INDIARAVI KANT SAHU « Security attacks with an effective solution for DOS attacks in VANET ». International Journal of Computer Applications (0975 – 8887) Volume 66– No.22, March 2013.

[24] RAM SHRINGAR RAW, MANISH KUMAR, NANHAY SINGH « Security challenges, issues and their solutions for VANET » International journal of network security and its applications (IJNSA) 5(5):95-105. September 2013.

[25] LAROUSSI KARIM, AMAR BENSABER BOUCIF, MESFIOUI MHAMED, BISKRI ISMAIL « A probabilistic model to corroborate three attacks in Vehicular Ad hoc Networks », 2015 IEEE Symposium on Computers and Communication (ISCC), 6-9 July 2015. Electronic ISBN: 978-1-4673-7194-0.

[26] https://en.wikipedia.org/wiki/Intelligent_transportation_system (4 Août 2016)

[27] https://fr.wikipedia.org/wiki/Moyenne_quadratique (4 Août 2016)

[28] https://en.wikipedia.org/wiki/Root_mean_square (4 Août 2016)

[30] http://projets-gmi.univ-avignon.fr/projets//proj1112/M2/p05/cc5.pdf (9 juillet 2016)

[31] ANGKOON PHINYOMARK, CHUSAK LIMSAKUL, AND PORNCHAI PHUKPATTARANONT "A Novel Feature Extraction for Robust EMG Pattern Recognition" Journal of Computing, volume 1, issue 1, December 2009, ISSN: 2151-9617.