



Universidad Popular Autónoma del Estado de Puebla

Vicerrectoría de Investigación y Posgrados

Decanato de Ingeniería y Negocios

Departamento de Ingeniería y Tecnologías de la Información

Posgrados en Ingeniería Mecatrónica

**DISEÑO DE UN SISTEMA DE DETECCIÓN DE INTRUSOS BASADO EN
MACHINE LEARNING PARA DETECTAR INFILTRACIONES DEL TIPO
GRAY HOLE**

EN REDES VANET.

Tesis que para obtener el Grado de

Doctor en Ingeniería Mecatrónica

Presenta:

M.C. JUAN ANTONIO ARÍZAGA SILVA

Asesores:

Dr. Luis Rosales Roldán

Dr. Alfredo Toriz Palacios

Dr. Ricardo Iván Álvarez
Tamayo

Puebla, México

14 de marzo de 2022



UPAEP – Secretaría General

Dirección General de Apoyos Académicos

Dirección del Centro de Recursos para el Aprendizaje y la Investigación.

Biblioteca Central - **Karol Wojtyla**

Tesis Digitales Restricciones de uso:

DERECHOS RESERVADOS ©

PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de textos, imágenes, gráficas, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente de donde la obtuvo mencionando el autor o autores involucrados en el documento.

Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



Universidad Popular Autónoma del Estado de Puebla

Vicerrectoría de Investigación y Posgrados

Decanato de Ingeniería y Negocios

Departamento de Ingeniería y Tecnologías de la
Información

Posgrados en Ingeniería Mecatrónica

Doctorado en Ingeniería Mecatrónica

Se aprueba la Tesis:

DISEÑO DE UN SISTEMA DE DETECCIÓN DE INTRUSOS BASADO EN MACHINE
LEARNING PARA DETECTAR INFILTRACIONES DEL TIPO GRAY HOLE EN REDES
VANET.

Comité Asesor:

Dr. Luis Rosales Roldán

Director de Tesis

Dr. Alfredo Toriz Palacios

Asesor/Tutor de Tesis

Dr. Ricardo Iván Álvarez Tamayo

Asesor/Tutor de Tesis

Puebla, México

14 de marzo de 2022

Agradecimientos

Agradezco al CONACYT por el apoyo recibido a través de la beca de Posgrado con la Industria número 179109/637863, que permitió la elaboración de esta Tesis.

DEDICATORIA

Dedicada a los que se fueron durante estos cuatro años:

Juan Antonio

y

José Carlos.

Aldo, por supuesto.

Tabla de contenido

Agradecimientos	III
DEDICATORIA.....	IV
Resumen.....	¡Error! Marcador no definido.
1 Vehicular Ad-hoc Networks (VANETs)	9
Introducción	9
1.1.1 Evolución histórica de las redes vehiculares	10
1.2 Redes VANET	11
1.2.1 Tecnología inalámbrica	11
1.2.2 Tipos de Comunicaciones vehiculares	12
1.2.4 IEEE 802.11p	17
1.3 Protocolos de enrutamiento de las redes VANET	18
1.4 El estándar IEEE 1609.x.....	20
1.5 Arquitectura de la próxima generación de VANET 5G.....	22
.....	23
1.6 Simulación de redes VANET	24
Referencias	26
2 Seguridad en redes VANET	29
2.1 Introducción	29
2.2 Trabajos previos	29
2.3 Amenazas de seguridad.	33
2.3.1 Riesgos sobre los usuarios	33
2.3.2 Amenazas a nivel de información.	33
2.3.3 Amenazas a la confidencialidad, integridad y disponibilidad de la información.	33

2.3.4 Amenazas a la infraestructura (RSU y Entidad central)	34
2.4 Conclusiones	35
Referencias	36
3 Cómputo inteligente en la seguridad de redes vehiculares Ad-hoc (VANET)	38
3.1 Introducción	38
3.2 Estado del arte en cómputo inteligente para ciberseguridad de las VANET	39
3.3 Clasificación de riesgos y vulnerabilidades en las VANET	40
3.3.1 Autenticidad	40
3.3.2 Disponibilidad	42
3.3.3 Confidencialidad	43
3.3.4 Integridad	43
Conclusiones	44
Referencias	45
4 Diseño del Sistema de Detección de Intrusos basado en Machine Learning	49
Resumen	49
Trabajos relacionados	49
Planteamiento del problema (el ataque de agujero gris)	52
Metodología	53
Configuración experimental	54
Generación de flujos y características	57
Resultados y análisis	58
Análisis	60
Conclusiones	65

Abstract

Wireless communication technologies are significant elements in the development of intelligent transportation systems (ITS) and cybersecurity is today's primary concern in vehicular communications. Vehicles are now able to connect to road infrastructure (V2I), other vehicles (V2V), or another element (V2X) spontaneously, vehicles are now nodes in a wireless network ready to interchange information.

The exchange of information between nodes in these vehicular networks must remain secure, reliable, and available. The lack of infrastructure and centralized management makes vehicular ad hoc networks (VANET) vulnerable to irregular behaviours that significantly threaten different aspects of network security. In this paper, an intelligent Intrusion Detection System (IDS) is proposed to prevent Gray-Hole attacks in VANET.

The advantage of the proposed IDS over existing systems is that it detects attacks before they cause considerable damage. Different Machine Learning algorithms were used to predict the behaviour of attacking nodes, being Random Forest the one that exhibited the best response to different performance metrics. Different simulations were developed using the NS3 network simulator to demonstrate that the IDS achieves high detection accuracy.

1 Vehicular Ad-hoc Networks (VANETs)

Introducción

En la última década, los automóviles han dejado de ser considerados un medio de transporte para convertirse en centros de datos móviles, con la intención de prevenir accidentes e incrementar la seguridad de los pasajeros; más aún se han diseñado una gran variedad de aplicaciones móviles lo que ha creado un nuevo mercado ante la necesidad de transferir la información de los autos o los ocupantes a través de Internet u otro tipo de red. La nueva generación de automóviles será capaz de realizar conexiones con su entorno, a través de lo que se ha dado a conocer como redes vehiculares, estas conexiones son diversas en aplicaciones y de diferente naturaleza: entre vehículos (V2V), con la infraestructura (V2I), con los peatones (V2P), entre otras; todas ellas agrupadas con el nombre de VANET (Vehicular Ad-Hoc Networks)[1].

Las redes VANET son un tipo de redes Ad-Hoc o de redes móviles distribuidas de naturaleza inalámbrica, formadas por dos tipos de nodos: estáticos y móviles. Los nodos estáticos, son elementos fijos emplazados a lo largo de las carreteras llamados RSU (Road-Side Unit), cuya función es la de enviar, recibir y retransmitir paquetes para aumentar el rango de cobertura de la red pudiendo también ofrecer acceso a Internet. Los nodos móviles son los vehículos equipados con un dispositivo electrónico llamado OBU (On Board Unit) para poder comunicarse con otros vehículos o con las RSU. Estos tipos de nodos tienen la capacidad de enviar, recibir y retransmitir mensajes entre ellos [2].

En este capítulo se expondrán las bases tecnológicas sobre las cuales están asentadas las redes VANET; los diversos estándares de comunicación desarrollados; los protocolos de enrutamiento, específicos para las redes Ad-hoc. Así como las redes VANET Heterogéneas (HetVNETs), principalmente aquellas que aparecerán en la convergencia de las redes Vehiculares y la red de comunicación 5G.

1.1.1 Evolución histórica de las redes vehiculares

El concepto de utilizar las comunicaciones de radio para comunicarse desde un vehículo con el fin de mejorar la seguridad ha existido mucho tiempo antes de las comunicaciones digitales. Un ejemplo es la patente "Sistemas de advertencia de radio para uso en vehículos" en 1922 y emitida en 1925. Basada en el concepto de comunicación por radio de punto a punto entre dispositivos iguales instalados en dos vehículos diferentes[3,4].

Sin tomar en cuenta a la industria militar, las comunicaciones vehiculares no tuvieron mayor desarrollo durante casi todo el siglo veinte, fue hasta finales de la década de los ochenta y principios de los noventa cuando el RDS Radio data system (sistema de datos de radio), un estándar de protocolo de comunicación para integrar pequeñas cantidades de información digital en transmisiones de radio de Frecuencia Modulada (FM) convencionales, en 1984, se convirtió en la primera infraestructura digital para comunicación de vehículos (I2V), y se introdujo en los Estados Unidos como Sistema de datos de transmisión de radio (RBDS) unos años más tarde. En 1990, RDS se convirtió en un estándar europeo. Tanto RDS como RBDS transportan datos a 1.187,5 bits por segundo en una subportadora de 57 kHz. Alrededor de 2005, luego de largas pruebas, RDS se mejoró para proporcionar RDS-Traffic Message Channel (TMC). Cada incidente de tráfico está codificado en binario y se envía como un mensaje TMC. Cada mensaje consta de un código de evento, un código de ubicación, la duración esperada del incidente, la extensión afectada y otros detalles. Sin embargo, la comunicación se seguía manteniendo unidireccional [3].

1.2 Redes VANET

1.2.1 Tecnología inalámbrica

Las redes Ad-Hoc pertenecen a un grupo especial de redes inalámbricas (wireless) en las cuales no existe un control centralizado, así como tampoco una infraestructura (infrastructureless networks), en principio no existe una conexión con el “mundo exterior”, los nodos de una red Ad-Hoc solamente se comunican entre ellos. La única infraestructura que existe en un momento determinado es la que se encuentra formada por los mismos nodos de la red. Dentro de las Redes AdHoc, cada nodo que forma la red puede actuar como ruteador (router) para encaminar la información hacia otro nodo de la red [5].

Puede hacerlo de forma directa o enviando mensajes hacia otros nodos que se encuentren físicamente más cercanos o en línea de vista con el nodo destino (multihop). Existe el envío de la información de tipo multicasting, para enviar información a todos los nodos de la red A continuación se detallan algunas de las principales características inherentes de las redes Ad-Hoc:

1. Multisalto: Una red multisalto es una red donde el camino del nodo emisor al nodo receptor puede atravesar varios nodos intermedios. Esto debido a que no se encuentran en línea de vista o que la potencia de transmisión del nodo transmisor no alcance la distancia física del receptor.
2. Conservación de la energía: Los nodos, por regla general, tienen una batería limitada y a no ser que dispongan de algún mecanismo para tal efecto, no tienen capacidad de recarga, por tanto, es muy importante conservar al mínimo sus comunicaciones y sus cálculos para prolongar la autonomía de las baterías.
3. Escalabilidad: En las redes distribuidas el número de nodos puede crecer hasta llegar a varios miles. Como no existe un punto de acceso concreto, la incorporación y descarte de nodos tiene que ser un proceso sencillo y transparente
4. Seguridad: Las redes inalámbricas son vulnerables a ataques, y las redes Ad-Hoc lo son especialmente debido a su naturaleza distribuida. Pueden padecer tanto ataques activos como pasivos, el atacante puede emular a un nodo legítimo y capturar la

información de cualquier otro nodo o retransmitir información falsa. El transporte seguro de mensajes (SMT) intenta solucionar este problema utilizando la redundancia en la red.

1.2.2 Tipos de Comunicaciones vehiculares

Las redes vehiculares son un tipo de redes Ad-Hoc Móviles formadas por dos tipos de nodos: estáticos y móviles. Los nodos estáticos, son elementos fijos emplazados a lo largo de las carreteras llamados RSU (Road-Side Unit), cuya función es la de enviar, recibir y retransmitir paquetes para aumentar el rango de cobertura de la red pudiendo también ofrecer acceso a Internet. Los nodos móviles son los vehículos equipados con un dispositivo electrónico llamado OBU (On Board Unit) para poder comunicarse con otros vehículos o con las RSU. Estos tipos de nodos tienen la capacidad de enviar, recibir y retransmitir mensajes entre ellos. Con estos elementos se han definido diferentes tipos de conexiones o escenarios de comunicación en las redes vehiculares: la comunicación intervehicular o vehículo a vehículo (V2V), en la que los automóviles intercambian mensajes directamente, la comunicación vehículo a infraestructura (V2I) comunicación vehículo a RSU (V2R) y comunicación vehículo a Peatón (V2P) [6]:

Vehicle to vehicle V2V

Este tipo de comunicación se refiere a la comunicación directa o basada en multisaltos entre vehículos en una red VANET.

Vehicle to Infrastructure V2I

El escenario de comunicación V2I hace referencia a la conexión existente entre los vehículos y la infraestructura a lo largo de la carretera.

Vehicle to Pedestrian V2P

V2P hace referencia a la comunicación entre los nodos de una red VANET y los peatones que circulan en un ambiente urbano.

Vehicle to (Smart)Grid V2G

Este es un sistema en el que los vehículos eléctricos se comunican con la red eléctrica para devolver electricidad a la red o acelerar la velocidad de carga del vehículo. Será un elemento en algunos modelos de autos que se conectan a la red y se utiliza como un modulador de red eléctrica para ajustar de manera dinámica la demanda de energía.

Vehicle to Home V2H

Este escenario ha sido propuesto por la ITU para la convergencia de Redes de Nueva Generación con redes VANET. Hace referencia a la comunicación entre un nodo de una red vehicular con un nodo de una red fija en el hogar a través de una infraestructura de red de próxima generación (Next Generation Network) NGN.

1.2.2.1 IEEE WAVE

En 1999, la Comisión Federal de Comunicaciones de los Estados Unidos (FCC) asignó un bloque de espectro en la banda 5.850-5.925 GHz para aplicaciones destinadas principalmente a las comunicaciones vehiculares. Los 75 Mhz. de este bloque se encuentran asignados, los 5 Mhz de la parte baja como una banda de guardia y los 70 Mhz restantes divididos en 7 canales de 10 Mhz. Existen en la actualidad bandas similares en Europa y Japón [4].

El estándar emergente para comunicaciones vehiculares son las Comunicaciones Dedicadas de Corto Alcance (DSRC). DSRC se basa en la tecnología IEEE 802.11 y ha sido titulado bajo el nombre de IEEE 802.11p. El objetivo de DSRC es apoyar operaciones de alta carga de transferencia de datos en las comunicaciones incluir información de tráfico, información de accidentes, condiciones de la carretera, mensajes de seguridad entre vehículos, cobro de peaje, manejo a través del pago, etc.

El estándar DSRC resultó atractivo debido al gran ancho de banda y la posibilidad de utilizar múltiples canales. Cabe aclarar que IEEE 802.11p no es el único estándar para DSRC en Europa y Japón los estándares son diferentes en tipo de comunicación, banda de frecuencia de radio utilizada para la comunicación, número de canales y su separación, velocidad de transmisión de datos, cobertura y su modulación [13]. Otro

estándar, el ASTM E2213-03 es la especificación para las telecomunicaciones y el intercambio de información entre los sistemas de carretera y vehículo [7].

Las especificaciones de la capa de control acceso al medio (MAC) y las especificaciones de capa física (PHY) del estándar DSRC se combina con estándar IEEE 1609.x para formar un estándar universalmente aceptado llamado Acceso inalámbrico en entornos vehiculares (WAVE). Esto puede verse en la Figura 1.1. El estándar IEEE 1609 define el modelo de comunicación, mecanismo de seguridad, estructura de administración, acceso físico y arquitectura general para comunicaciones inalámbricas para los componentes básicos (RSU, OBU).

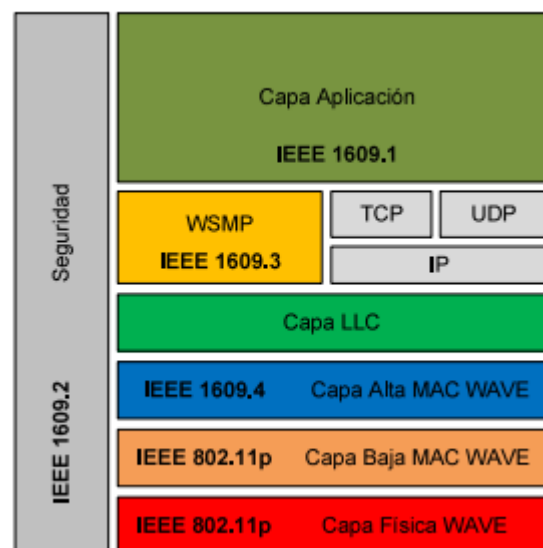


Figura 1.1 Arquitectura de WAVE [8]

1.2.2.2 La familia de estándares 802

Los protocolos WAN IEEE 802.11 [9] son una parte de los estándares que definen las redes de área local (LAN) y las redes de área metropolitana (MAN) definidos por la IEEE 802. Los estándares de esta familia poseen una sola Capa de control de enlace (Logical Link Control layer LLC) la cuál es estandarizada como 802.2 sobre la cual se encuentra el protocolo IP con sus protocolos de enrutamiento.

Por debajo de la capa LLC, como se muestran en la figura 1.2, se encuentran la capa de Control de Acceso a los Medios (MAC) y la correspondiente capa física (PHY), estas están empaquetadas juntas en el mismo subgrupo estándar. Existen muchos

subgrupos estándar como el de Ethernet y LAN inalámbrica que se especifican en el estándar 802.11.

ISO/OSI Network Layers

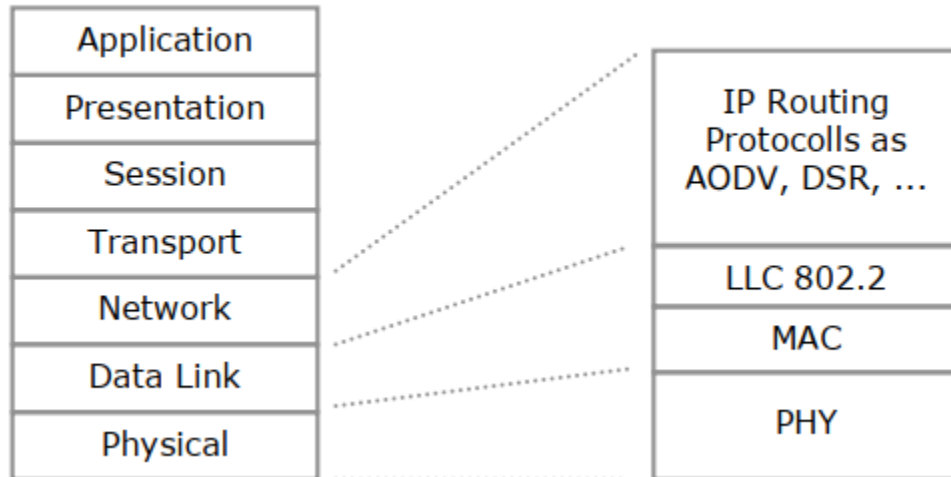


Figura 1.2 Modelo de capas ISO/OSI [9]

1.2.2.3 El estándar 802.11

Dentro del IEEE 802.11 hay definidos una serie de grupos de trabajo que se encargan de investigar y desarrollar diferentes temas relacionados con la 802.11. Estos grupos son los que se muestran en la Tabla 1-1.

Estándar	Descripción
802.11a	5GHz OFDM PHY – 54 Mbps
802.11b	2.4GHz CCK PHY – 11 Mbps
802.11c	802.11 bridging
802.11d	International roaming
802.11e	QoS/efficiency enhancements
802.11f	Inter AP protocol
802.11g	2.4GHz OFDM PHY – 54 Mbps
802.11h	5GHz regulatory extensions

802.11i	Security enhancements
802.11j	Japan 5GHz band extensions
802.11k	Radio resource measurement
802.11m	Maintenance
802.11n	High throughput PHY
802.11p	WAVE
802.11r	Fast Basic Service Set Transition
802.11ac	5GHz Gigabit

Tabla 1-1 Grupos de trabajo del Estándar 802.11

Estas extensiones o grupos de trabajo se han desarrollado para mejorar la calidad del servicio (QoS) y la seguridad para la capa MAC. Las extensiones de la capa física (PHY) redefinen en su mayoría la forma en que la capa física funciona. [9]

1.2.3.3 La capa PHY

El protocolo IEEE 802.11p utiliza en la capa física (PHY) Multiplexación por División de Frecuencias Ortogonales (OFDM) tal como lo hace el IEEE 802.11a/g [10]. Define dos subcapas a nivel de capa física: la subcapa PLCP (Physical Layer Convergence Procedure) y subcapa PMD (Physical Medium Dependent).

La PLCP es responsable de la comunicación de la capa física con la capa MAC, convierte las Unidades de datos de protocolo (PDU) de la capa MAC en una trama OFDM [11].

La subcapa PMD define características y métodos de transmisión y recepción de información entre dos o más estaciones a través del medio inalámbrico [11]. Especifica técnicas de modulación, codificación y la operación de transmisión.

1.2.3.4 La capa MAC

El protocolo de Control de Acceso al Medio (MAC) en IEEE 802.11p está basado en el mecanismo de Acceso al Canal Distribuido Mejorado (EDCA). A diferencia de otros estándares 802.11, el IEEE 802.11p permite comunicaciones que pueden o no contar

con un conjunto de servicios básicos (BSS); es decir, a través del modo de infraestructura o el modo ad hoc[12].

En el modo ad hoc, cada nodo (es decir, vehículo) actúa como punto de acceso en el SRS y emite periódicamente una señal de reconocimiento (beacon). Además, para transmitir un paquete, cada vehículo emplea el esquema de acceso múltiple por detección de portadora (CSMA/CA) como mecanismo de acceso al medio.

1.2.4 IEEE 802.11p

Para sintetizar, el Estándar IEEE 802.11p divide su espectro de 75 MHz. en un canal de guardia de 5MHz y 7 canales de 10 MHz. cada uno. De estos canales uno de ellos se utiliza como canal de control (CCH), que se reserva para los mensajes cortos de alta prioridad o los datos de gestión [12] mientras que, de los 6 restantes, 4 se utilizan como canales de servicio y dos de ellos como canales de seguridad, como puede observarse en la Figura 1.3.

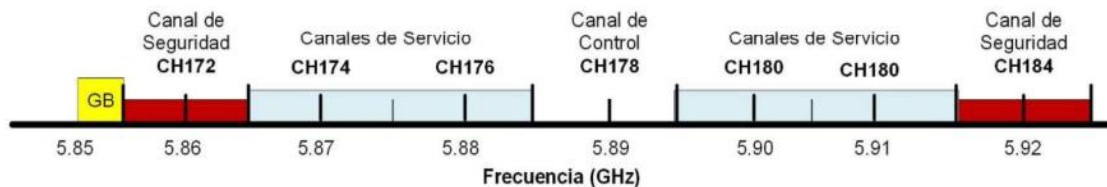


Figura 1.3 Canales en IEEE802.11p [12]

Un dispositivo IEEE 802.11p que implementa OFDM de 10 MHz puede transmitir y recibir datos a velocidades de 3, 4,5, 6, 9, 12, 18, 24 y 27 Mb/s. Las capacidades de comunicación de datos a velocidades de 3, 6 y 12 Mb/s son obligatorias en WAVE [9]. La tabla 1-3 muestra los parámetros que se definen en la capa PHY del IEEE 802.11p.

Especificaciones IEEE 802.11p

Tasa de datos (Mbps)	3, 4, 5, 6, 9, 12, 18, 24, 27
Modulación	BPSK, QPSK, 16QAM, 64QAM
Tasa de codificación	1/2, 1/3, 3/4
Subportadoras	52
Duración de símbolo	8us
Tiempo de Guarda	1.6us
Duración de Preámbulo	32us
Espaciamiento de Subportadora	0.15625 MHz

Tabla 1-3 Estándar 802.11p

1.3 Protocolos de enrutamiento de las redes VANET

Las VANET son una colección de nodos móviles y fijos que forman una red por un periodo de tiempo muy breve sin la necesidad de tener una infraestructura de red o una administración centralizada, esto hace que la topología de la red no sea fija sino que varíe con el tiempo. En este tipo de redes, los nodos se comportan al mismo tiempo como ruteadores y participan en el descubrimiento y mantenimiento de las rutas para diseminar mensajes a través de la red. Este comportamiento permite la creación de una desde cero sin la participación o intervención de ningún usuario.

1.3.1 Clasificación de Protocolos de enrutamiento

Cuando un nodo en la red desea comunicarse con cualquier otro nodo, si no existe línea de vista o transmisión directa entre los nodos, es necesario que se tenga establecido una ruta que permita conectar a través de los diversos nodos de la red, el nodo fuente y el nodo destino.

Ante esta tarea los protocolos de ruteo tradicionales no pueden adaptarse correctamente a las redes AdHoc por lo que se desarrollaron protocolos de ruteo específicos que permiten a los nodos de la red.

Los algoritmos de enrutamiento en las redes ad Hoc tienen la tarea de descubrir y mantener las rutas entre nodos. Estos pueden clasificarse siguiendo varios criterios en cuatro tipos distintos dependiendo del modo de descubrimiento de las rutas [13]:

- El alcance: unicast, broadcast o multicast, geocast, etc.
- El modo de descubrimiento de rutas: proactivo, reactivo, híbrido.
- Tipo de algoritmo que implementan: vector de distancias, estado de enlace, etc.

Basados en alcance:

Existen dos agrupaciones de protocolos basado en su alcance: protocolos unicast y protocolos multicast.

Los protocolos Unicast son aquellos que buscan crear las rutas enviando su información a un solo nodo a la vez, los protocolos multicast envían información a un grupo de nodos.

Los protocolos geocast son un caso particular de protocolos multicast que son muy usados en redes VANET. Son protocolos donde los grupos están organizados en función de la posición geográfica de los nodos de la red. Se usan protocolos geocast cuando se desea mandar un mensaje a un grupo de vehículos de una determinada zona geográfica, para advertir sobre riesgos, por ejemplo, en la carretera [13].

Basados en el modo de descubrimiento de las rutas.

Protocolos reactivos

El enrutamiento reactivo establece una ruta sólo cuando es necesario que un nodo se comunique con otro. Sólo mantiene las rutas que están en uso en ese momento, reduciendo así la carga de la red. El enrutamiento reactivo suele tener una fase de descubrimiento de rutas en la que los paquetes de consulta se inundan en la red en busca de una ruta. La fase finaliza cuando se encuentra una ruta al destino [14].

Protocolos proactivos

Los protocolos proactivos intentan conocer en cada momento, e independientemente de las necesidades de enrutamiento, una idea precisa del estado de la red. Algunos

paquetes de control son enviados de manera periódica para mantener el estado de las rutas entre cualquier par de nodos.

Esta característica procura una respuesta rápida ante solicitudes de ruta y suele ofrecer un buen comportamiento en los escenarios de movilidad alta. Sin embargo, para mantener una actualización permanente de las rutas, se introduce una sobrecarga importante de la red con los mensajes de control.[14]

Protocolos híbridos

Existen protocolos híbridos que combinan esos dos esquemas utilizando, por ejemplo, proactividad en las cercanías del nodo considerado, pero buscando las rutas bajo demanda para los nodos más alejados.

1.4 El estándar IEEE 1609.x

Como se ha visto, el estándar de facto DSRC utiliza en la subcapa de Control de Enlace Lógico (que está dentro de la capa de enlace) el estándar establecido IEEE 802.2. En las partes superiores del estándar y para la extensión de la subcapa MAC, DSRC utiliza una clase de normas definidas por el Grupo de Trabajo IEEE 1609. La subcapa MAC apoya la funcionalidad multicanal empleando la norma IEEE 1609.4 [15], la cual define como un dispositivo cambia entre los distintos canales del DSRC.

Por encima de la capa de enlace, la norma de servicios de red IEEE 1609.3 [16] define dos pilas de protocolos (que comparten la misma pila inferior en las capas de enlace físico y de datos): El Protocolo de mensajes cortos WAVE (WSMP) y la versión estándar del Protocolo de Internet 6 (IPv6). El WSMP ha sido diseñado para soportar la red y la capa de transporte funciona en un entorno vehicular inalámbrico. Los paquetes enviados a través del WSMP son referidos como “mensajes cortos de WAVE” (WSM). Uno de estos mensajes es BSM, por el cual los vehículos difunden periódicamente su información de estado clave, como la posición, la velocidad, la aceleración y el rumbo.

Los protocolos de la capa de aplicación podrían soportar aplicaciones de seguridad tanto como aplicaciones no seguras. Las aplicaciones no relacionadas con la seguridad pueden ser provistas por el sistema de comunicación de Internet y el DSRC [17]. La norma IEEE 1609.11 es un ejemplo de una norma de capa de aplicación no relacionada con la seguridad [18]. IEEE 1609.11 El intercambio de datos de pago electrónico por vía aérea proporciona un servicio común interoperable para la identidad de los dispositivos, la autenticación de los pagos y la transferencia de datos de pago. La norma de servicios de seguridad IEEE 1609.2 es otra norma de la clase de normas IEEE1609, que define las técnicas de cifrado y autenticación para los WSM [19].

1.5 Arquitectura de la próxima generación de VANET 5G

Las redes IEEE802.11p como se ha visto en este capítulo, han sido desarrolladas exclusivamente para las redes VANETs. Sin embargo, el crecimiento de otro tipo de redes y el uso simultáneo de diferentes tecnologías de acceso radioeléctrico presentan graves dificultades en cuanto al control y la gestión de la infraestructura. Más aun, teniendo en cuenta los inconvenientes de las redes IEEE 802.11p, como la escasa escalabilidad, la baja capacidad y la conectividad intermitente, las redes móviles LTE (Long Term Evolution) han sido propuestas como una posible solución para apoyar las aplicaciones en los vehículos [13].

Cuando las tecnologías de comunicación LTE se han integrado en las redes vehiculares, los resultados de las simulaciones y los experimentos revelaron que la interferencia reduce el rendimiento de las redes de vehículos basados en LTE [20]. Para superar este problema, se propuso la tecnología de transmisión por ondas milimétricas para conectar a los usuarios dentro de los vehículos. Por otra parte, se propusieron las Redes Definidas por Software (Software Defined Networks, SDN) como una tecnología de red efectiva, capaz de soportar la naturaleza dinámica de las redes vehiculares.

Los diversos estudios sobre redes VANET y SDN han llevado a la aparición del paradigma de Redes Vehiculares definidas por Software (Software-Defined Vehicular Networks, SDVN). La arquitectura de las SDVN principalmente comprende el concepto de separar en planos distintos la arquitectura de las redes VANET. En consecuencia, la visión más simplificada de la arquitectura de la SDVN comprende principalmente tres planos (el plano de aplicación, el plano de control y el plano de datos) con sus correspondientes interfaces conectadas, como se muestra en la Figura 1.4.

Por otro lado, con la aparición de los vehículos autónomos, son necesarios algunos requisitos más rigurosos (por ejemplo, el retardo de transmisión debe ser inferior a 1

ms) y en general para los sistemas de transporte inteligente (STI) y para las comunicaciones V2V y V2I [12]. Para cumplir estos rigurosos requisitos, se espera que las tecnologías de comunicación móvil 5G, la computación en nube y las SDN se integren en las futuras redes vehiculares. Por lo tanto, es necesario diseñar una nueva arquitectura de red para las redes vehiculares de 5G.

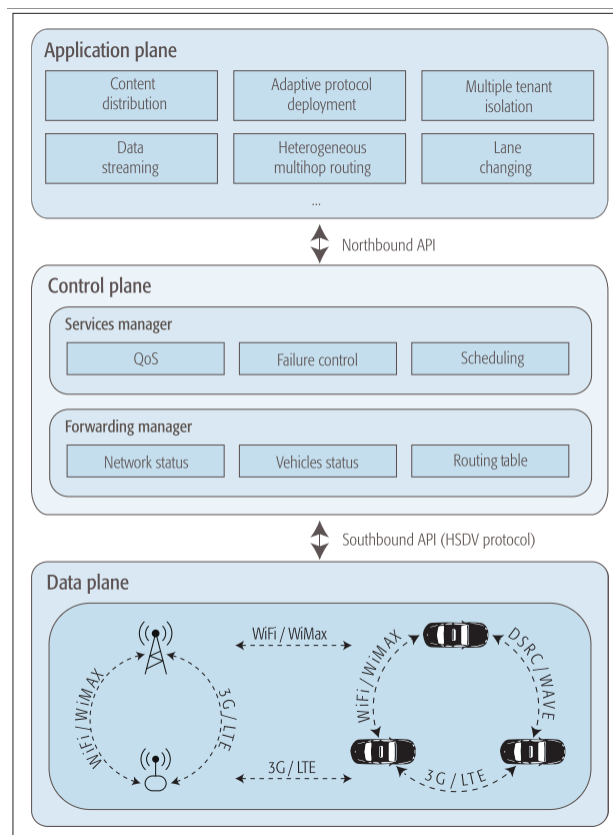


Figura 1.4 Estructura de las Redes Vehiculares definidas por Software [23]

1.6 Simulación de redes VANET

Las herramientas más populares para realizar simulaciones en el desarrollo de soluciones de ciberseguridad en redes VANET son el simulador de tráfico y movilidad SUMO, el simulador de red NS2 y su actualización NS3. Cabe destacar que la mayoría de los equipos de trabajo en general utilizan herramientas de simulación del tipo código abierto (Python, Veins, java, Sumo, NS2, NS3) esto debido a dos factores, el primero es que existe una amplia comunidad de desarrolladores que apoya y da soporte a las diversas herramientas, al igual que diversas empresas y universidades que brindan apoyo para el desarrollo de las mismas; por otro lado al ser código abierto es más sencillo crear un desarrollo ad hoc sin depender del pago de una licencia o problemas de derechos de autor..

El NS2 es una de las herramientas de simulación más aceptadas en la investigación académica e industrial de redes [21]. Se desarrolló como parte de un proyecto de la Agencia de Proyectos de Investigación Avanzada de Defensa (DARPA). Se trata de un simulador de código abierto, que utiliza C++ y lenguaje de comandos de herramientas (TCL) de programación. En NS2, hay muchos módulos y estándares disponibles para diferentes escenarios de red. Así, NS2 se ha convertido en un simulador estándar de facto para las comunicaciones de red por cable e inalámbricas. Sin embargo, el uso de NS2 para simular diferentes escenarios de redes locales inalámbricas (Wireless Local Area Network, WLAN), redes móviles ad hoc (Mobile Ad-Hoc Network, MANET), redes de sensores inalámbricos (Wireless Sensor Network, WSN), etc. es obvio debido a la disponibilidad de varios módulos de red y estándares inalámbricos. La versión oficial de NS2 cuenta con dos modelos de movilidad para la simulación de VANET, a saber, el modelo de autopista y el modelo Manhattan (Grid). Aquí, los vehículos se mueven de forma independiente y también es posible el posicionamiento o la colocación de dos o más vehículos. Con respecto a las comunicaciones VANET, NS2 se enfrenta a algunos retos como la implementación de nodos de alta movilidad nodos, entornos de simulación con más de cien nodos, etc. [22].

Hoy en día el uso del NS3, una herramienta que ha evolucionado de NS2, ha aumentado drásticamente y parece que será una de las mejores herramientas de simulación en el futuro. En comparación con el NS2, el NS3 tiene una arquitectura menos compleja con un núcleo de simulación optimizado. El NS3 también elimina el uso de la programación TCL. En el NS3 se pueden utilizar diferentes lenguajes informáticos como C, C++ y Python. Esto hace que el NS3 sea capaz de manejar escenarios de simulación a gran escala con más de diez mil nodos.

El NS3 es fácil de usar y de depurar los errores gracias a su documentación. También satisface la necesidad de un escenario de simulación completo, desde los ajustes de configuración básicos hasta archivo de trazas o la compilación y el análisis detallados. Su arquitectura también contempla el desarrollo y despliegue de marcos de simulación suficientes para utilizarlo como emulador de red, es decir, se puede crear una réplica exacta del entorno realista e interconectarlo con aplicaciones del mundo real y sus protocolos [22].

El NS3 admite redes basadas en el Protocolo de Internet (IP) o no basadas en IP, con una visión de las simulaciones de IP móvil. Dispone de diferentes modelos de comunicación como Wi-Fi, Wi-Max y Long Term Evolution (LTE) con el apoyo de varios protocolos de enrutamiento estático y dinámico. NS3 también es compatible con IEEE 802.11p, que es el estándar WAVE para las comunicaciones vehiculares.

Referencias

- [1] A. Sari, O. Onursal, and M. Akkaya, "Review of the Security Issues in Vehicular Ad Hoc Networks (VANET)," *Netw. Syst. Sci.*, vol. 8, pp. 552–566, 2015, doi: 10.4236/ijcns.2015.813050.
- [2] S. Tanwar, J. Vora, S. Tyagi, N. Kumar, and M. S. Obaidat, "A systematic review on security issues in vehicular ad hoc network," *Secur. Priv.*, vol. 1, no. 5, p. e39, 2018, doi: 10.1002/spy2.39.
- [3] Arizaga-Silva J. Alonso-Perez M. Álvarez-González R., "Redes VANET Vehicular Ad-Hoc Networks, la conectividad de los autos. Primera parte." 2018 Revista Visión Politécnica. Universidad Politécnica de Puebla.
- [4] Campolo, C. e. a. *Vehicular Ad Hoc Networks* (Springer Int., 2015).
- [5] Pathan, A.-S. K. *Security of Self-Organizing Networks: MANET, WSN, WMN, VANET* (Auerbach Publications, Boston, MA, USA, 2010), 1st edn.
- [6] Al-Sultan, S., Al-Doori, M. M., Al-Bayatti, A. H. & Zedan, H. A comprehensive survey on vehicular ad hoc network. *J. Netw. Comput. Appl.* 37, 380–392 (2014). URL <http://dx.doi.org/10.1016/j.jnca.2013.02.036>. DOI 10.1016/j.jnca.2013.02.036.
- [7] V. Jindal, K. Mahavidyalaya, and P. Bedi, "Vehicular Ad-Hoc Networks: Introduction, Standards, Routing Protocols and Challenges," *Int. J. Comput. Sci. Issues*, vol. 13, no. 2, pp. 44–55, 2016, doi: 10.20943/01201602.4455.
- [8] k. Alobuela, c. Padilla. "análisis y documentación de la implementación de la capa física ieee 802.11a y 802.11p para redes ad hoc en el simulador omnet++," escuela politécnica nacional facultad, 2019.
- [9] "Vehicular Ad Hoc Networks," *Hans J. Wirel. Commun.*, vol. 03, no. 01, pp. 22–44, 2013, doi: 10.12677/hjwc.2013.31005.
- [10] D. T. Fakultät and D. Eckhoff, "Simulation of Privacy-Enhancing Technologies in Vehicular Ad-Hoc Networks," pp. 13–17, 2016.

- [11] A. . M. Abeldime y L. Wu, The Physical Layer of the IEEE 802.11p WAVE Communication Standard: The Specifications and Challenges, San Francisco, 2014.
- [12] IEEE, IEEE Standard for Information technology--Telecommunications and information exchange between systems Local and metropolitan area networks-- Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE, 2016.
- [13] Z. El-Rewini, K. Sadatsharan, D. F. Selvaraj, S. J. Plathottam, and P. Ranganathan, "Cybersecurity challenges in vehicular communications," Vehicular Communications, vol. 23. Elsevier Inc., Jun. 01, 2020, doi: 10.1016/j.vehcom.2019.100214.
- [14] W. Viriyasitavat, "A framework for reliable and efficient communications in vehicular networks," ProQuest Diss. Theses, p. 210, 2012, [Online]. Available: <http://search.proquest.com/docview/1151827010?accountid=41453>.
- [15] IEEE Std, "IEEE standard for wireless access in vehicular environments (WAVE)- multi-channel operation, IEEE std 1609.4-2016," 2016.
- [16] IEEE Std, "IEEE Standard for Wireless Access in Vehicular Environments (WAVE)- Networking Services, IEEE Std. 1609.3-2010," IEEE Std.
- [17] H. Hartenstein and K. Laberteaux, VANET: Vehicular Applications and InterNetworking Technologies. Wiley Online Library, 2010
- [18] IEEE Std, "IEEE Standard for Wireless Access in Vehicular Environments (WAVE)—Over-the-Air Electronic Payment Data Exchange Protocol for Intelligent Transportation Systems (ITS), IEEE Std. 1609.11-2010," IEEE Std.
- [19] IEEE Std, "IEEE Standard for Wireless Access in Vehicular Environments Security Services for Applications and Management Messages, IEEE Std. 1609.2- 2013," IEEE Std.
- [20] X. Ge, Z. Li, and S. Li, "5G software defined vehicular networks," IEEE Commun. Mag., vol. 55, no. 7, pp. 87–93, 2017, doi: 10.1109/MCOM.2017.1601144.

- [21] K. M. Ali Alheeti, A. Gruebler and K. D. McDonald-Maier, "An intrusion detection system against malicious attacks on the communication network of driverless cars," *2015 12th Annual IEEE Consumer Communications and Networking Conference (CCNC)*, 2015, pp. 916-921, doi: 10.1109/CCNC.2015.7158098.
- [22] S. Mallisery, M. M. M. Pai, M. Mehbadi, R. M. Pai, and Y. S. Wu, "Online and offline communication architecture for vehicular ad-hoc networks using NS3 and SUMO simulators," *J. High Speed Networks*, 2019, doi: 10.3233/JHS-190615.
- [23] S. Correia, A. Boukerche, and R. I. Meneguette, "An architecture for hierarchical software-defined vehicular networks," *IEEE Commun. Mag.*, vol. 55, no. 7, pp. 80–86, 2017, doi: 10.1109/MCOM.2017.1601105.

2 Seguridad en redes VANET

2.1 Introducción

En los últimos años, los automóviles han dejado de ser considerados un medio de transporte para convertirse en centros de datos móviles con la intención de incrementar la seguridad de los pasajeros, prevenir accidentes y brindar mejores servicios [1]; para esto se han diseñado una gran variedad de aplicaciones móviles lo que ha creado un nuevo mercado, ante la necesidad de transferir la información de los autos o los ocupantes a través de Internet u otro tipo de red.

La nueva generación de automóviles será capaz de realizar conexiones con su entorno, a través de lo que se ha dado a conocer como redes vehiculares, estas conexiones son diversas en aplicaciones y de diferente naturaleza: entre vehículos (vehicle-2-vehicle,V2V), con la infraestructura (vehicle-2-Infrastructure,V2I), con los peatones (vehicle-2-pedestrian,V2P), entre otras; todas ellas agrupadas con el nombre de VANET's. (Vehicular Ad-Hoc Networks) [2].

Hay muchos tipos de ataques que pueden ocurrir en una red VANET, cada elemento dentro de la red necesita un enfoque de ciberseguridad diferente según su naturaleza. En este capítulo, se pretende, desde el punto de vista de ciberseguridad establecer los elementos de alto valor, los activos (assets), que hacen de las redes VANET el blanco de ataques cibernéticos; por último, se exponen los diferentes tipos de amenazas a los que pueden estar expuestos cada uno de los diferentes activos en la red

2.2 Trabajos previos

En una VANET, cada vehículo se define como un nodo de la red y está equipado con una unidad de comunicación a bordo denominada OBU (On-Board Unit) y una unidad de aplicación llamada AU (Application Unit). Mediante la OBU los vehículos son capaces de comunicarse de forma inalámbrica entre sí, así como con las unidades de la carretera (Road Side Units) RSUs. Se espera que las redes VANET soporten una amplia gama de aplicaciones prometedoras tales como servicios basados en la ubicación. Sin embargo, la naturaleza de la difusión del medio inalámbrico permitiría a

un agente adverso espiar las comunicaciones que contengan los identificadores de nodo, y estimar las ubicaciones de los nodos de comunicación con suficiente precisión para rastrear los nodos [5-7].

La implementación de la seguridad en las redes VANET tiene desafíos únicos. El caso es más complicado debido a los diferentes requisitos de las distintas aplicaciones. La seguridad para la difusión segura de la información requiere un enfoque diferente al requerido para aplicaciones de gestión del tráfico debido a que existen muchos tipos diferentes de ataques que pueden ocurrir dentro de una red VANET. La clasificación de los tipos de ataques dentro de la literatura especializada depende en gran medida del grupo de investigación por lo que en este artículo se mostrará una aproximación a los problemas de ciberseguridad basada en los activos (assets) presentes dentro de una red VANET. Ejemplos de la diversidad de clasificaciones se presentan a continuación: Los investigadores Du y Zhu proponen en su trabajo [8] un modelo de árbol de ataque donde asignan a cada nodo del árbol una probabilidad de sufrir un ataque por parte de un agente externo. La probabilidad en cada nodo depende de tres atributos: costo de ataque, dificultad técnica y dificultad de descubrimiento.

Kaur et al. [9] presentan cinco diferentes categorías de ataques, aunque no muestran evidencia de por qué realiza esta división, antes bien cita trabajos anteriores [6,7,10]. Por otro lado, Tyagi y Dembla proponen solo dos clasificaciones de ataque, internos y externos, sin imponer un criterio explícito de dicha división [3]. Al igual que con otros sistemas, los desafíos de seguridad de los vehículos autónomos y las redes vehiculares se pueden clasificar en términos generales en ataques a la confidencialidad, integridad, privacidad y disponibilidad como lo mencionan Gerla y Reiher [11]; los otros sistemas a los que se refieren estos autores corresponden a sistemas ciberfísicos que soportan infraestructura crítica y sistemas de información gubernamental y empresarial donde es necesario asegurar los activos (valores tangibles e intangibles) de las organizaciones.

Desde el punto de vista de ciberseguridad, los siguientes elementos son considerados activos dentro de las redes VANET [12]:

1) Usuario de la red VANET.

Al ser una red concebida para incrementar la seguridad física de los automovilistas, estos constituyen, junto con su información privada, el activo más importante de la red.

2) Intercambio de información.

Al igual que otras redes, los usuarios de las redes Vehiculares también exigen seguridad en términos de integridad de datos, confidencialidad y disponibilidad (CIA por sus siglas en inglés: confidentiality, integrity and availability).

3) Vehículos.

Dentro de las redes VANET los automóviles representan un activo importante por las tareas que realizan al interior de la red. Un vehículo en la actualidad más que un medio de transporte se ha convertido en un centro de datos móvil. Al interior del vehículo se encuentran los sensores que coleccionan la información que después será transmitida a través de la red utilizando la OBU hacia los nodos adyacentes al vehículo.

El automóvil puede también, como se ha mencionado anteriormente, transmitir, recibir y rutear paquetes de información hacia los demás elementos de la red.

4) Unidades en carretera (RSU)

Este elemento sirve como puente entre el entorno de infraestructura y el entorno adhoc. En estos nodos se instalarán los sensores necesarios, las unidades de procesamiento y el sistema de comunicación para recibir información de otros nodos. Debido a su naturaleza estática es un activo con una alta probabilidad de ser atacado. Si la RSU se ve comprometida, los datos almacenados en el interior se ven comprometidos y no se puede garantizar la comunicación segura con la infraestructura.

5) Protocolos de comunicación de red.

Las comunicaciones en una red VANET son por su naturaleza principalmente inalámbricas, el estándar de facto para comunicaciones vehiculares son las

Comunicaciones Dedicadas de Corto Alcance (DSRC). El DSRC se basa en la tecnología IEEE 802.11 y ha sido titulado bajo el nombre de IEEE 802.11p. Estas comunicaciones pueden incluir información de tráfico, información de accidentes, condiciones de la carretera, mensajes de seguridad entre vehículos, cobro de peaje, manejo a través del pago, etc. Otros dos estándares, el ASTM E2213-03 y el IEEE 1609.x conforman el estándar conocido como Acceso inalámbrico en entornos vehiculares (WAVE) [2].

Por otro lado, también son importantes y son un activo de valor las comunicaciones intravehiculares para conexión de los sensores y las ECU al interior del auto y con la OBU que se realizan a través de los buses CAN, LIN y Ethernet Automotriz.

6) Entidad central.

La entidad central es otro nodo estático en la arquitectura VANET que incluye los servidores de aplicaciones que proporcionan diversas aplicaciones, como aplicaciones para evitar colisiones, actualizaciones del clima y el tráfico, etc. La entidad central se encuentra en el dominio de la infraestructura y desempeña un papel vital durante la comunicación V2I donde los mensajes son recibidos primero por el servidor de aplicaciones. Autentica el mensaje recibido y lo reenvía a otros vehículos a través de una ubicación geográfica amplia [12].

7) Terceros en la red.

Los terceros en la red representan a los distintos tipos de autoridades (policía y tránsito), los cuales se encuentran en la parte de la red VANET que representa la Infraestructura. También se incluye a los fabricantes de los automóviles, los cuales a través de canales dedicados de comunicación pueden tener acceso a la red intravehicular, al OBU o la Unidad de aplicación.

Es necesario asegurarse que los terceros sean confiables a través de algún elemento de seguridad, cuando esto sucede, son conocidos como TTP (trusted third parties)

2.3 Amenazas de seguridad.

Esta sección presenta diversas amenazas potenciales para los activos de las redes VANET

2.3.1 Riesgos sobre los usuarios

A nivel de Usuario de la red las principales amenazas están relacionadas con la confidencialidad de la información personal del usuario, así como su localización geográfica y otra información sensible.

Por ejemplo, se prevé que los servicios ofrecidos en VANET incluyan conexión a Internet y aplicaciones peer to peer para compartir archivos entre usuarios de la red.

Un agente adversario podría generar aplicaciones maliciosas que sustraigan más información de los usuarios o que incluso pueda existir robo de identidad.

Otra amenaza hacia los usuarios de las redes VANET, como cualquier otra red es aquella basada en ingeniería social (Physhing o spoofing).

2.3.2 Amenazas a nivel de información.

Siempre existen amenazas a la información donde el principal interés del atacante es comprometer su confidencialidad, integridad y autenticidad (CIA). Las amenazas a la información pueden explotarse siguiendo diferentes aspectos de seguridad

Al interior del vehículo, se ha comprobado que existen diversos ataques a las comunicaciones internas del vehículo, específicamente en el Bus de comunicaciones CAN [12], lo que podría dar lugar al envío de información errónea por parte de un usuario a los demás nodos de red, sobre todo en comunicaciones V2V y V2I, sobre diferentes aspectos relacionados a la velocidad y posición de un nodo móvil de la red [14].

2.3.3 Amenazas a la confidencialidad, integridad y disponibilidad de la información.

Un nodo de red puede actuar de manera maliciosa de diferentes formas, cada de una de ellas puede clasificarse como un intento de manipulación de la información.

Por ejemplo, cada nodo dentro de la red puede servir de puente como ruteador de la información entre dos nodos, cuando por algún motivo no conocido un nodo decide no retransmitir los paquetes de datos ante una solicitud es entonces cuando la disponibilidad de la información se ve alterada y la seguridad de los usuarios, automovilistas y peatones se ve en riesgo.

Por otro lado, la confidencialidad de los paquetes de datos puede verse vulnerada al existir algún canal inalámbrico inseguro con un encriptamiento que resulte ineficiente [15].

De igual forma la integridad de la información se ve comprometida cuando un atacante externo puede alterar, modificar e incluso borrar paquetes de datos transmitidos entre varios elementos de la red.

Existen diversos vectores de potenciales amenazas a la información [10,12]:

- Escuchas de datos confidenciales.
- Ataque de interferencia (Jamming).
- Ataques de suplantación
- Ataques de hombre en el medio Man in the middle (MITM).
- Ataques de suplantación de identidad (Spoofing).

2.3.4 Amenazas a la infraestructura (RSU y Entidad central).

La infraestructura, siendo la entidad estática en VANET, es una de las ubicaciones favoritas para que un atacante lance diferentes ataques de red del tipo Denegación de Servicio (DoS) y del tipo Persona en el medio (PITM), para generar alteración de mensajes en el canal y ataques de suplantación [17].

La forma de los ataques puede ser del tipo [12,14,19]:

- Modificación de mensajes con mensajes incorrectos y comprometidos
- Manipulación y caída de mensajes.
- Fuga de datos en el canal cableado del back-end

- Inundación de red con mensajes comprometidos
- Alteraciones de mensajes en ruta a otros vehículos a través de RSU y entidad central
- Ataques relacionados con la calidad (QoS) del mensaje

Como puede observarse en algunos dominios los tipos de ataques se superponen. Esto se debe a la naturaleza heterogénea de las redes VANET las cuales cubren diferentes tipos de protocolos, topologías y medios de transmisión

2.4 Conclusiones

En las últimas dos décadas, se han emprendido muchos proyectos VANET al rededor del mundo y se han desarrollado varios estándares VANET para mejorar las comunicaciones entre vehículos (V2V) y vehículo a otros (V2X). La existencia de redes VANET abre el camino para una amplia gama de aplicaciones y ha abierto la puerta a diversos riesgos de ciberseguridad.

En este capítulo se ha presentado una revisión de los elementos de alto valor dentro de las redes VANET; también se revisaron algunas de las principales amenazas de ciberseguridad en las que los investigadores han centrado su atención en los últimos años. Así como los diversos tipos de ataques que pueden surgir derivados de estas amenazas.

Referencias

- [1] Jindal, V. and Bedi, P. "Vehicular ad-hoc networks: Introduction, standards, routing protocols and challenges" (2016).
- [2] Arizaga-Silva J. Alonso-Perez M. Álvarez-González R., "Redes VANET Vehicular Ad-Hoc Networks, la conectividad de los autos. Primera parte." 2018 Revista Visión Politécnica. Universidad Politécnica de Puebla.
- [3] Tyagi and D. Dembla, "Advanced Secured Routing Algorithm of Vehicular Ad-Hoc Network," Wireless Personal Communications, vol. 102, no. 1, pp. 41–60, May 2018.
- [4] [4] "Mobile ad hoc network and vehicular ad-hoc network security," security of self-organizing networks, pp. 124–124, Sep. 2010. doi:10.1201/ebk1439819197-8
- [5] Al-Sultan, S., Al-Doori, M. M., Al-Bayatti, A. H. & Zedan, H. A comprehensive survey on vehicular ad hoc network. J. Netw. Comput. Appl. 37, 380–392 (2014).
- [6] Hoa La and A. Cavalli, "Security Attacks and Solutions in Vehicular Ad Hoc Networks: A Survey", International Journal on AdHoc Networking Systems, vol. 4, no. 2, pp. 1-20, 2014.
- [7] J. Isaac, S. Zeadally and J. Cámara, "Security attacks and solutions for vehicular ad-hoc networks", IET Communications, vol. 4, no. 7, p. 894, 2010.
- [8] Du and H. Zhu, "Security Assessment in Vehicular Networks," SpringerBriefs in Computer Science, 2013.
- [9] R. Kaur, T. P. Singh, and V. Khajuria, "Security Issues in Vehicular Ad-Hoc Network (VANET)," 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), May 2018.
- [10] Saini, M.S., & Singh, H.P. (2016). VANET , its Characteristics , Attacks and Routing Techniques : A Survey. International Journal of Science and Research (IJSR), vol. 5, no. 5, pp. 1595-1599, 2016.

- [11] M. Gerla and P. Reiher, "Securing the Future Autonomous Vehicle: A Cyber-Physical Systems Approach," *Securing Cyber-Physical Systems*, pp. 197–220, Sep. 2015.
- [12] Ahmad, F., Adnane, A. and N. L. Franqueira, V. (2016) "A Systematic Approach for Cyber Security in Vehicular Networks". *Journal of Computer and Communications*, 4, 38-62.
- [13] Nadeem Majeed, M. e. a. Vehicular ad-hoc networks history and future development arenas. *ITEE J.* (2013).
- [14] A. A. Celes and N. E. Elizabeth, "Verification Based Authentication Scheme for Bogus Attacks in VANETs for Secure Communication," 2018 International Conference on Communication and Signal Processing (ICCSP), Chennai, 2018, pp. 0388-0392. doi: 10.1109/ICCSP.2018.8524540
- [15] K. M. A. Alheeti, A. Gruebler and K. D. McDonald-Maier, "An intrusion detection system against malicious attacks on the communication network of driverless cars," 2015 12th Annual IEEE Consumer Communications and Networking Conference (CCNC), Las Vegas, NV, 2015, pp. 916-921. doi: 10.1109/CCNC.2015.7158098
- [16] R. Shringar Raw, M. Kumar and N. Singh, "Security Challenges, Issues and Their Solutions For Vanet", *International Journal of Network Security & Its Applications*, vol. 5, no. 5, pp. 95-105, 2013.
- [17] H. Doumenc, "Estudio comparativo de protocolos de encaminamiento en redes vanet," tech. rep., Universidad Politecnica de Madrid, 2008.
- [18] C. Campolo, A. Molinaro, & R. Scopigno (Eds.), *Vehicular ad hoc Networks: standards, solutions, and research* (Springer Int., 2015).

3 Cómputo inteligente en la seguridad de redes vehiculares Ad-hoc (VANET)

3.1 Introducción

Los sistemas inteligentes de transporte (ITS), los cuales se pretende que permitan una gestión eficiente y controlada del tráfico de vehículos y la movilidad urbana en general, dentro de las llamadas ciudades "inteligentes" se encuentran basados fundamentalmente en el desarrollo de nuevas tecnologías de comunicación. En ese sentido, las redes vehiculares ad-hoc se han desarrollado para aumentar la seguridad de los usuarios y vehículos en diversos entornos automotrices, así como también brindar nuevos servicios a los conductores y peatones.

El intercambio de información entre los diversos elementos de estas conexiones, V2V, V2I, V2G, V2P, se divide en dos tipos, crítico y no crítico; como ejemplos de los primeros se encuentran: notificación de desvíos, respuesta a un evento de emergencia, manejo de carriles, minimización de congestión vehicular, información de las condiciones de la carretera, priorización de los vehículos de emergencia como ambulancias y camiones de bomberos [1]. Los datos no críticos hacen referencia a aplicaciones para intercambio de información no esencial entre nodos de la red VANET (Servicios, infoentretenimiento, etc.).

Estos datos proporcionan información contextual sobre la cinética del vehículo (como la velocidad, la aceleración y la dirección), las condiciones de la carretera, el flujo de tráfico, los entornos inalámbricos, etc., que puede explotarse para mejorar el rendimiento de la red mediante una toma de decisiones adaptativa basada en los datos. Sin embargo, las estrategias de comunicación tradicionales no están diseñadas para manejar y explotar esa información.

La Inteligencia Artificial, por otro lado, ha desarrollado métodos altamente eficientes para modelado y análisis de grandes volúmenes de datos mediante la búsqueda de patrones y estructuras subyacentes. En particular el aprendizaje de máquina (*machine learning*) ha tenido un amplio progreso en diversas áreas de aplicación, como

clasificación [2][3], reconocimiento de voz [4] y procesamiento del lenguaje natural y predicción [5], así como dentro del campo de la ciberseguridad, donde según reportes actuales [6] el 74% de los encargados de ciberseguridad en empresas a nivel mundial utilizan alguna herramienta de *machine learning* para reforzar la seguridad de sus centros de trabajo.

Machine learning proporciona un amplio conjunto de herramientas que pueden ser aprovechadas para explotar los datos producidos, consumidos y almacenados en las redes vehiculares. En este capítulo se realizará una revisión del estado del arte que guarda la aplicación de las herramientas del cómputo inteligente, principalmente el aprendizaje de máquina en diferentes contextos de las redes VANET.

3.2 Estado del arte en cómputo inteligente para ciberseguridad de las VANET

Diversos grupos de investigadores [7-10] han realizado revisiones sistemáticas de la seguridad de las comunicaciones intervehiculares de forma general, mientras que otros han escrito sobre temas más específicos, como en [11] donde la revisión se centra en sistemas de detección de Intrusos (IDS) realizados con algoritmos de Machine Learning; ninguno ha sido específicamente para revisar las investigaciones de ciberseguridad automotriz en las redes VANET utilizando herramientas de cómputo inteligente

Dicho esto. Tras una revisión de la literatura se observa que la clasificación de los tipos de riesgos y ataques reportados depende en gran medida del grupo de investigación.

3.3 Clasificación de riesgos y vulnerabilidades en las VANET

Ejemplos de la diversidad de clasificaciones se presentan a continuación: Los investigadores Du y Zhu proponen en su trabajo [12] un modelo de árbol de ataque donde asignan a cada nodo del árbol una probabilidad de sufrir un ataque por parte de un agente externo. La probabilidad en cada nodo depende de tres atributos: costo de ataque, dificultad técnica y dificultad de descubrimiento.

Kaur et al [13] presentan cinco diferentes categorías de ataques, aunque no muestran evidencia de por qué realiza esta división, antes bien cita trabajos anteriores [10,14,15]. Por otro lado, Tyagi y Dembla proponen solo dos clasificaciones de ataque, internos y externos, sin imponer un criterio explícito de dicha división [16].

En el trabajo de Ahmad et al [17] se intenta identificar los riesgos y las vulnerabilidades en las redes VANET catalogando los activos (assets) de la red en tres grupos diferentes: Infraestructura, información y Sistema vehicular. Esto con el fin, a decir de los investigadores, de facilitar la evaluación de la seguridad y análisis de las amenazas.

Cabe destacar que otros grupos de investigadores que trabajan en la seguridad de sistemas ciberfísicos, no exclusivamente en la industria automotriz, como lo mencionan Gerla y Reiher [18] utilizan la clasificación CIA la cual propone que, en términos generales, los ataques a la seguridad de cualquier Sistema ciberfísico son ataques a la confidencialidad (privacidad), integridad, autenticidad y disponibilidad de la información. En la Tabla 3.1 se muestran los ejemplos de diversos ataques utilizando la clasificación CIA.

3.3.1 Autenticidad

La autenticidad asegura la legitimidad del origen del mensaje. La información necesaria para poder transitar con seguridad a través de una red de carreteras necesita ser verificada. En ese sentido, el trabajo de [20] propone varios algoritmos de Machine Learning aplicados a la información proveniente de las imágenes de diferentes cámaras del centro de Londres con el fin de determinar de forma acertada el flujo vehicular.

Tabla 3-1 Tipos de amenazas en las Redes VANET [19]

Categoría	Ataque	Objetivo
Integridad	Manipulación de mensajes	
	Enmascarado	Cambiar el contenido de los mensajes para enviar información errónea o datos falsos
	Agujero negro	
	Agujero gris	
	Fabricación	
	Malware	
Autenticidad	Sibil Attack	
	Manipulación de mensajes	
	Enmascarado	Permitir que los usuarios no verificados se conecten a la red, permitiendo que se comuniquen sin la identificación correcta
	Agujero negro	
	Agujero gris	
	Worm hole	
	Fabricación	
	Malware	
Confidencialidad	Replay attack	
	Escucha	Leer y revelar el contenido de los mensajes de los usuarios que viajan a través de las redes para ejecutar diferentes ataques
	Revelación de identidad	
	Análisis de Tráfico	
	Malware	
Disponibilidad	Denegación de servicio (DoS)	Esto permite a los atacantes remover el vehículo de la red y hacer que no se encuentre disponible para el resto de los vehículos en la VANET.
	Agujero negro/gris	
	Spamming	

Jamming y malware

En [21] utilizan un algoritmo deep reinforcement learning que toma la información de las velocidades de los automóviles que lo rodean para tomar la decisión de cambio de carril. El cambio de carril no es una tarea trivial ya que no solo depende de la dinámica del automóvil sino también de la geometría del camino y las velocidades de los automóviles circundantes.

El trabajo de [22] sugiere un algoritmo de decisión descentralizado para verificar la información obtenida por los mensajes enviados por diversos nodos de la VANET para crear de manera colaborativa un mapa de alta calidad para automóviles con manejo autónomo.

Los autores que presentan en [23] realizan el entrenamiento de una red neuronal de tres capas para desarrollar un Sistema de detección de intrusos en las redes VANET extrayendo el comportamiento normal y anormal en la red utilizando el archivo de rastreo obtenido de la simulación ya que contiene diferentes datos (características) que se pueden utilizar para análisis. Estas características describen el comportamiento normal o malicioso de los nodos. Los autores claman obtener un 98% de precisión al utilizar 21 características del archivo de rastreo.

3.3.2 Disponibilidad

Un nodo de red puede actuar de manera maliciosa de diferentes formas, cada de una de ellas puede clasificarse como un intento de manipulación de la información. Por ejemplo, cada nodo dentro de la red puede servir de puente como ruteador de la información entre dos nodos, cuando por algún motivo no conocido un nodo decide no retransmitir los paquetes de datos ante una solicitud es entonces cuando la disponibilidad de la información se ve alterada y la seguridad de los usuarios, automovilistas y peatones se ve en riesgo.

Los resultados reportados en [24] plantean una reducción en la transmisión de datos a través de los nodos de la red utilizando para ello un algoritmo adaptivo (SBSD) que permite mejoras, tanto en la entrega como en el tiempo de respuesta durante las simulaciones de la red.

En el trabajo presentado en [25] se describe un IDS basado en técnicas de ML diseñadas para detectar tanto los ataques de "spoofing" como los de "jamming" en un entorno de automóviles conectados (VANET). El motor de detección del IDS presentado se basa en los algoritmos ML RF, k-NN y OCSVM.

Otro trabajo que presenta una medida contra un ataque de "jamming" utilizando diversos algoritmos ML (RandomForests/SVM/AdaBoost) está desarrollado en [26]. Los resultados se dividen en dos escenarios. Experimentos escenarios interiores y exteriores. Utilizando una tarjeta de desarrollo para generar una herramienta antijamming con una exactitud del 99% en el mejor de los casos para interiores y una de 78% en exteriores.

3.2.3 Confidencialidad

En [27] los investigadores propusieron un modelo de detección de mal comportamiento basado en un conjunto que sustituye a los umbrales estáticos por umbrales dinámicos que se determinan en línea. Como tal, el modelo propuesto es capaz de hacer frente a la naturaleza dinámica de la red. Se utilizaron los filtros de Kalman y Hampel para ajustar automáticamente estos umbrales. El modelo fue entrenado usando los datos provenientes de los clasificadores estadísticos, parámetros de contexto, consistencia, plausibilidad y características de comportamiento.

En [10] se analizó la intervención de intrusos malintencionados en un ataque del tipo persona en medio (person in the middle) en las redes VANET. El principal objetivo de estos intrusos es engañar a la comunicación en general difundiendo contenidos maliciosos tanto a los vehículos conectados como a los autónomos en la red. Para abordar estas cuestiones, los autores propusieron un marco basado en Blockchain que puede garantizar la secrecía y la transparencia en la red, ya que la información se almacena y se rastrea en la misma cadena

3.2.4 Integridad

En los trabajos reportados en [28,29,30] se investigó la manera de hacer más eficientes los protocolos de comunicación para la diseminación segura de la información. A decir de los investigadores en [31], un algoritmo de agrupación eficiente debe tener en

cuenta toda la información necesaria relacionada con la movilidad de los nodos. Los investigadores proponen un Protocolo de Agrupación Ponderada Adaptativa (AWCP), especialmente diseñado para redes vehiculares, que tiene en cuenta la identificación de la autopista, la dirección de los vehículos, la posición, la velocidad y el número de vehículos vecinos para mejorar la estabilidad de la topología de la red, convirtiendo el problema de elevar el QoS de la red en un problema de optimización multiobjetivo, plantean el uso de algoritmo genético clasificado no dominado NSGA-II para encontrar los parámetros óptimos.

En [29] un algoritmo adaptivo, diseñado por los investigadores, aprende a agrupar automóviles cercanos para disminuir el retraso en la entrega de paquetes dentro de una arquitectura híbrida que integra LTE con la comunicación multi-hop basada en IEEE 802.11p.

El trabajo de [30] propone un algoritmo de enrutamiento autoadaptativo seguro (RSAR) basado en el algoritmo Q-Learning. Combinando los parámetros de fiabilidad y ajustando la función heurística, el RSAR consigue un buen rendimiento con VANET.

Conclusiones

Los trabajos presentados en el presente capítulo abordan diversos temas de cómputo inteligente que atienden a diversos problemas de ciberseguridad. Eso habla de que este es un tema de interés en la actualidad. Ofreciendo un área de oportunidad para desarrollar mayor investigación.

Dos puntos quedan por abordar. El primero es que, muchos equipos de trabajo utilizan sus propias métricas para realizar la validación de sus proyectos. Y que no existe una clasificación de los riesgos de ciberseguridad basado en la normativa vigente. La clasificación de los tipos de riesgos y ataques reportados depende en gran medida del grupo de investigación.

Referencias

- [1] Macher, G. et al., 2017. Automotive SPICE, Safety and Cybersecurity Integration. Computer Safety, Reliability, and Security, pp.273–285. Available at: http://dx.doi.org/10.1007/978-3-319-66284-8_23.
- [2] L. I. R. Ruiz, W. Beccaro, B. G. P. Evaristo and F. J. Ramirez Fernandez, "Tactile Sensing Glove-Based System for Objects Classification Using Support Vector Machine," in IEEE Latin America Transactions, vol. 16, no. 6, pp. 1658-1663, June 2018. doi: 10.1109/TLA.2018.8444383
- [3] M. Melgarejo, C. Parra and N. Obregon, "Applying Computational Intelligence To The Classification Of Pollution Events," in IEEE Latin America Transactions, vol. 13, no. 7, pp. 2071-2077, July 2015. doi: 10.1109/TLA.2015.7273760
- [4] G. Chau and G. Kemper, "One Channel Subvocal Speech Phrases Recognition Using Cumulative Residual Entropy and Support Vector Machines," in IEEE Latin America Transactions, vol. 13, no. 7, pp. 2135-2143, July 2015.
doi: 10.1109/TLA.2015.7273769
- [5] L. P. Del Bosque and S. E. Garza, "Prediction of Aggressive Comments in Social Media: an Exploratory Study," in IEEE Latin America Transactions, vol. 14, no. 7, pp. 3474-3480, July 2016. doi: 10.1109/TLA.2016.7587657
- [6] A. Satoh, Y. Nakamura, and T. Ikenaga, "Ssh dictionary attack detection based on flow analysis," in Applications and the Internet (SAINT), 2012 IEEE/IPSJ 12th International Symposium on, July 2012, pp. 51–59.
- [7] M. Dibaei et al., "Attacks and defences on intelligent connected vehicles: a survey," Digital Communications and Networks. Chongqing University of Posts and Telecommunications, 2020, doi: 10.1016/j.dcan.2020.04.007.
- [8] S. Tanwar, J. Vora, S. Tyagi, N. Kumar, and M. S. Obaidat, "A systematic review on security issues in vehicular ad hoc network," Secur. Priv., vol. 1, no. 5, p. e39, 2018, doi: 10.1002/spy2.39

- [9] M. Abu Talib, S. Abbas, Q. Nasir, and M. F. Mowakeh, "Systematic literature review on Internet-of-Vehicles communication security," *Int. J. Distrib. Sens. Networks*, vol. 14, no. 12, Dec. 2018, doi: 10.1177/1550147718815054.
- [10] Saini, M.S., & Singh, H.P. (2016). VANET , its Characteristics , Attacks and Routing Techniques : A Survey, *International Journal of Science and Research (IJSR)*, vol. 5,no. 5, pp. 1595-1599, 2016.
- [11] F. Goncalves et al., "A Systematic Review on Intelligent Intrusion Detection Systems for VANETs," *Int. Congr. Ultra Mod. Telecommun. Control Syst. Work.*, vol. 2019-October, 2019, doi: 10.1109/ICUMT48472.2019.8970942.
- [12] Du and H. Zhu, "Security Assessment in Vehicular Networks," *SpringerBriefs in Computer Science*, 2013.
- [13] R. Kaur, T. P. Singh, and V. Khajuria, "Security Issues in Vehicular Ad-Hoc Network(VANET)," 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), May 2018.
- [14] Hoa La and A. Cavalli, "Security Attacks and Solutions in Vehicular Ad Hoc Networks: A Survey", *International Journal on AdHoc Networking Systems*, vol. 4, no. 2, pp. 1-20, 2014.
- [15] J. Isaac, S. Zeadally and J. Cámara, "Security attacks and solutions for vehicular ad-hoc networks", *IET Communications*, vol. 4, no. 7, p. 894, 2010.
- [16] Tyagi and D. Dembla, "Advanced Secured Routing Algorithm of Vehicular Ad-Hoc Network," *Wireless Personal Communications*, vol. 102, no. 1, pp. 41–60, May 2018.
- [17] Ahmad, F., Adnane, A. and N. L. Franqueira, V. (2016) "A Systematic Approach for Cyber Security in Vehicular Networks". *Journal of Computer and Communications*, 4, 38-62.
- [18] M. Gerla and P. Reiher, "Securing the Future Autonomous Vehicle: A Cyber-Physical Systems Approach," *Securing Cyber-Physical Systems*, pp. 197–220, Sep. 2015.

- [19] M. Abu Talib, S. Abbas, Q. Nasir, and M. F. Mowakeh, "Systematic literature review on Internet-of-Vehicles communication security," *Int. J. Distrib. Sens. Networks*, vol. 14, no. 12, Dec. 2018, doi: 10.1177/1550147718815054.
- [20] R. Ketabi, M. Al Qathrady, B. Alipour, and A. Helmy, "Vehicular traffic density forecasting through the eyes of traffic cameras; a spatio-temporal machine learning study," *DIVANet 2019 - Proc. 9th ACM Symp. Des. Anal. Intell. Veh. Networks Appl.*, pp. 81–88, 2019, doi: 10.1145/3345838.3356002.
- [21] S. Yang et al, "Message Relaying and Collaboration Motivating for Mobile Crowdsensing Service: An Edge-Assisted Approach," *Wireless Communications and Mobile Computing*, vol. 2018, pp. 1-13, 2018.
- [22] H. An and J. Il Jung, "Decision-making system for lane change using deep reinforcement learning in connected and automated driving," *Electron.*, vol. 8, no. 5, 2019, doi: 10.3390/electronics8050543.
- [23] K. M. A. Alheeti, A. Gruebler and K. D. McDonald-Maier, "An intrusion detection system against malicious attacks on the communication network of driverless cars," 2015 12th Annual IEEE Consumer Communications and Networking Conference (CCNC), Las Vegas, NV, 2015, pp. 916-921. doi: 10.1109/CCNC.2015.7158098
- [24] A. Ouksel and D. Lundquist, "Reducing redundant data transmissions in wireless ad hoc networks: comparing aggregation and filtering," *Wireless Networks*, vol. 21, (7), pp. 2155-2168, 2015.
- [25] D. Kosmanos et al., "Intrusion Detection System for Platooning Connected Autonomous Vehicles." Accessed: Nov. 24, 2020. [Online].
Available: <https://ieeexplore.ieee.org/abstract/document/8908528/>.
- [26] O. Punal, I. Aktas, C. J. Schnellke, G. Abidin, K. Wehrle, and J. Gross, "Machine learning-based jamming detection for IEEE 802.11: Design and experimental evaluation," *Proceeding IEEE Int. Symp. a World Wireless, Mob. Multimed. Networks 2014*, doi: 10.1109/WoWMoM.2014.6918964.

- [27] Ghaleb, Maarof, Zainal, Alrimy, Alsaeedi, and Boulila, "Ensemble-Based Hybrid Context-Aware Misbehavior Detection Model for Vehicular Ad Hoc Network," *Remote Sens.*, vol. 11, no. 23, p. 2852, Dec. 2019, doi: 10.3390/rs11232852.
- [28] M. Hadded et al, "A multi-objective genetic algorithm-based adaptive weighted clustering protocol in VANET," in 2015, . DOI: 10.1109/CEC.2015.7256998.
- [29] S. Ucar, S. C. Ergen and O. Ozkasap, "Multihop-Cluster-Based IEEE 802.11p and LTE Hybrid Architecture for VANET Safety Message Dissemination," *IEEE Transactions on Vehicular Technology*, vol. 65, (4), pp. 2621-2636, 2016.
- [30] D. Zhang, T. Zhang and X. Liu, "Novel self-adaptive routing service algorithm for application in VANET," *Applied Intelligence* (Dordrecht, Netherlands), vol. 49, (5), pp. 1866-1879, 2018;2019;
- [31] M. Hadded et al, "A multi-objective genetic algorithm-based adaptive weighted clustering protocol in VANET," in 2015, . DOI: 10.1109/CEC.2015.7256998.

4 Diseño del Sistema de Detección de Intrusos basado en Machine Learning

Resumen

Las tecnologías de comunicación inalámbrica son elementos importantes en el desarrollo de los sistemas de transporte inteligentes (ITS) y la ciberseguridad es hoy la principal preocupación en las comunicaciones inalámbricas. Los vehículos son ahora capaces de conectarse a la infraestructura vial (V2I), a otros vehículos (V2V) o a otro elemento (V2X) de forma espontánea, los vehículos son ahora nodos de una red inalámbrica. El intercambio de información entre los nodos de estas redes vehiculares debe seguir siendo seguro, fiable y disponible. La falta de infraestructura y de gestión centralizada hace que las redes vehiculares ad hoc (VANET) sean vulnerables a comportamientos irregulares que amenazan significativamente diferentes aspectos de la seguridad de la red. En este trabajo se propone un sistema inteligente de detección de intrusos (IDS) para prevenir los ataques de tipo Gray-Hole en las VANET. La ventaja del IDS propuesto sobre los sistemas existentes es que detecta los ataques antes de que causen daños considerables. Se utilizaron diferentes algoritmos de Machine Learning para predecir el comportamiento de los nodos atacantes, siendo Random Forest el que mejor respuesta mostró a las diferentes métricas de rendimiento. Se desarrollaron diferentes simulaciones utilizando el simulador de red NS-3 para demostrar que el IDS alcanza una alta precisión de detección.

Trabajos relacionados

El aprendizaje automático es una forma eficaz de diseñar IDS con gran precisión en comparación con otras tecnologías. Se han realizado muchas investigaciones para abordar el comportamiento erróneo en las VANETS.

Alheeti et al. [1] diseñaron un IDS inteligente que predice los ataques de denegación de servicio (DoS) en las simulaciones de VANET. La técnica de predicción de intrusiones se basa en el Análisis Discriminante Lineal (Linear

Discriminant Analysis, LDA) y el Análisis Discriminante Cuadrático (Quadratic Discriminant Analysis, QDA). Los autores utilizan 21 características extraídas de los resultados de las simulaciones, pero no dan más información al respecto. Los mejores resultados que afirman es una tasa de precisión del 95% cuando se utiliza el LDA.

Coussement et al. diseñaron un IDS para identificar comportamientos anormales que puedan tener un impacto negativo en los sistemas de comunicaciones externas de los vehículos móviles [2]. El sistema propuesto examina los paquetes de comunicación entrantes y salientes para detectar ataques. Utiliza un mecanismo de toma de decisiones para asegurar la información sensible de las VANET. El sistema de seguridad propuesto se basa en una combinación de dos esquemas IDS y la agrupación de nodos móviles.

En [3] se presenta un IDS para ataques Black Hole. Los autores consideran cuatro características en el diseño del esquema. La técnica de detección se basa en un modelo probabilístico; según los autores un vehículo está conectado con cierta probabilidad a la VANET si está rodeado al menos por n vehículos cooperativos y no hay vecinos Black Hole conectados a él. En lugar de utilizar métricas basadas en detecciones positivas y negativas, los autores utilizan métricas como la sobrecarga de enrutamiento (ROH), la tasa de pérdida de paquetes (PLR) y el retraso de extremo a extremo (E2E) para garantizar que el IDS funciona correctamente puesto que estas características permiten determinar si hay un agente intermedio interfiriendo con la entrega de paquetes de datos.

Un método de detección de intrusos basado en el aprendizaje profundo en VANET [5]. Implementan un sistema de detección de intrusos para un modelo de comunicación vehicular utilizando algoritmos de aprendizaje profundo. Para la extracción de características se utiliza el algoritmo CNN (Convolutional Neural Network) y para la clasificación el algoritmo LSTM (Long Short Term Memory). Los autores compararon su método con métodos anteriores en un conjunto de datos públicos y en un conjunto de datos VANET simulado de la vida real.

Grover et al [7] compara diferentes algoritmos de aprendizaje automático para detectar ataques Sybil. Como otros, utilizan características conocidas como el número de paquetes transmitidos, recibidos, abandonados y capturados, pero también otras nuevas como la verificación de la desviación de la velocidad para verificar la posición de los nodos falsos y la presencia de RSU en forma de semáforos para observar el intercambio de información entre los nodos. Los clasificadores Random Forest (RF) y J-48 presentan los mejores resultados basados en las características seleccionadas para la detección binaria y multiclase.

Siguiendo la misma línea [8] los autores presentan un IDS basado en dos algoritmos de aprendizaje automático, SVM y K neihbour; las características fueron extraídas de un conjunto de datos etiquetados con información de los Mensajes Básicos de Seguridad (BSM) de VANET, definidos en el estándar SAE J2735. Las características se utilizaron para detectar y clasificar cinco tipos diferentes de comportamientos erróneos conocidos. El trabajo demostró que SVM podía ser utilizado con éxito como módulo de clasificación con una tasa de clasificación correcta (CCR) del 88% y KNN con el 78%.

En [10], los autores utilizaron características extraídas del archivo de trazas pcap para detectar el ataque. El sistema de detección de intrusiones/anomalías utilizó el archivo de trazas como datos auditables para detectar el ataque. El IDS propuesto utiliza una versión fuzzificada de los datos del fichero pcap para alimentar una red neuronal para detectar el ataque de agujero negro.

Los autores de [17] proponen un modelo robusto de reputación distribuida basado en un filtro bayesiano. El modelo permite a los nodos establecer perfiles (por ejemplo, malicioso, honesto) sobre sus vecinos y detectar comportamientos maliciosos (por ejemplo, agujero negro, agujero gris).

En un enfoque diferente, se ha estudiado la teoría de juegos para generar un tipo diferente de IDS para prevenir comportamientos maliciosos en los nodos de las VANETS como se puede ver en [9, 18]

Planteamiento del problema (el ataque de agujero gris)

Los ataques de agujero negro y agujero gris se dan en las redes ad hoc, Las redes ad hoc utilizan diferentes tipos de protocolos de enrutamiento, de los cuales una clase se conoce como protocolos basados en tablas. Las tablas de enrutamiento se crean y mantienen mediante la difusión controlada de mensajes multicast y unicast. En el caso del protocolo AODV se utilizan dos mensajes de control principales para el descubrimiento de rutas, que son la solicitud de ruta (RREQ) y la respuesta de ruta (RREP).

Cuando el origen necesita enviar datos para comunicarse con el destino, y no hay una ruta disponible, entonces inunda de mensajes RREQ a sus vecinos. Los vecinos responden con RREP si tienen la mejor ruta hacia el destino. Si no la tienen, envían más mensajes RREQ a sus vecinos y el proceso de descubrimiento de rutas continúa hasta que el RREQ llega al destino o al nodo intermedio que ha obtenido la ruta. La fuente envía paquetes de datos al recibir el RREP.

Los mensajes RREP contienen dos métricas utilizadas por AODV para elegir la mejor ruta entre los nodos, que son: los números de secuencia y el recuento de saltos. Los números de secuencia dan información sobre la frescura de los paquetes de control y el recuento de saltos mantiene el número de nodos entre el origen y el destino.

Un atacante Gray/black hole es capaz de difundir un mensaje RREP falso en respuesta a una petición RREQ válida con información falsa sobre el número de secuencia y el recuento de saltos. El nodo Gray/black Hole compromete el número de secuencia del destino estableciendo esta métrica a un valor máximo posible para eludir al origen y hacerle creer que un nodo diferente al legítimo debería ser el siguiente salto. El recuento de saltos se establece en un valor tan bajo como 2 junto con el número de secuencia de destino para que la fuente se marque como siguiente salto [16].

De este modo, el atacante redirige todos los paquetes hacia sí mismo para dejarlos caer y bloquear a otros nodos para que se comuniquen con el nodo atacado. Este comportamiento puede realizarse con un nodo objetivo o con todos los nodos. El agujero gris es casi lo mismo que el agujero negro, pero el comportamiento puede cambiar a normal en función de los nodos objetivo o simplemente de vez en cuando. Así que el ataque de agujero gris es más difícil de abordar que el de agujero negro.

Metodología

La metodología propuesta en este trabajo utiliza algoritmos de clasificación de aprendizaje automático para obtener un alto rendimiento de predicción, precisión de clasificación y baja latencia de detección en el cálculo de la confianza en VANETs bajo ataques de "agujero gris". La Figura 1 describe el Sistema de Detección de Intrusos (IDS) propuesto.

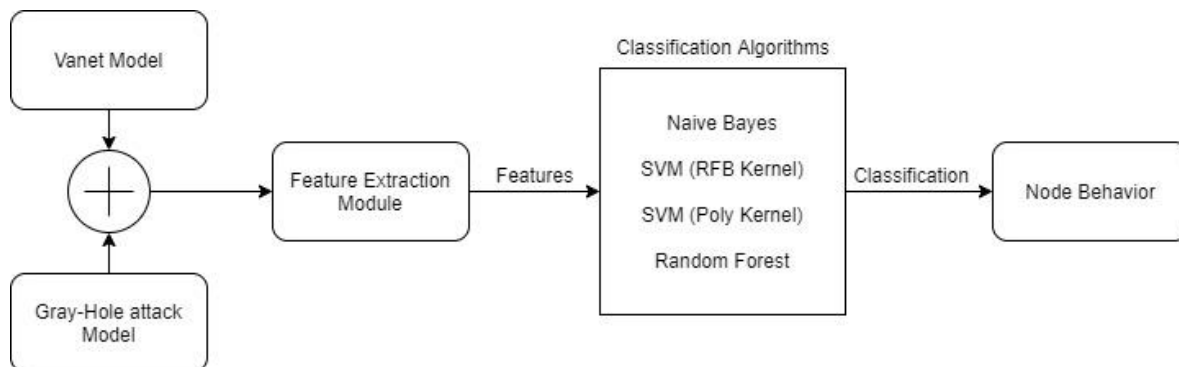


Figura 4.1 Diseño propuesto para clasificar los nodos legítimos y los que se comportan mal

Esta tarea de predicción/clasificación se realiza a partir de los atributos o características relevantes que existen con mayor frecuencia para clasificar una muestra como nodo legítimo o nodo con mal comportamiento [7]. Estas

características se obtuvieron tras realizar una secuencia de simulaciones en un escenario VANET basado en el simulador NS-3.

Configuración experimental

La información sobre los datos de los paquetes de red se registra y almacena en archivos con formato pcap. Este formato de archivo, ampliamente conocido que se utiliza para guardar de forma ordenada los paquetes de datos (*data packets*) de diferentes sistemas de comunicación.

Como se observa en la Figura 4.1, la primera parte del proceso es realizar el modelo de la red VANET, esto se hace a través de diversas simulaciones llevadas a cabo en el el simulador NS3 integrado con SUMO. NS3 es un simulador de red ampliamente utilizado y SUMO es un simulador de tráfico [1,7] NS-3 proporciona múltiples características, incluyendo: simulaciones de comunicación y de protocolo de red, simulaciones de comunicación de redes vehiculares, control de configuraciones cliente-servidor.

EN NS-3 cada dispositivo de red generado es un nodo, Esta abstracción es representada en C++ a través de una class llamada *Node*. Esta clase contiene todos los métodos para manejar y controlar lo que sucede con los nodos durante las simulaciones.

En NS-3 durante una simulación además de generarse los nodos que interactúan entre sí, las aplicaciones generadas por los usuarios se ejecutan en los nodos de ns-3 para realizar simulaciones. Estas aplicaciones están modeladas a través de una clase llamada *Application*

En el mundo real cuando una persona desea conectar una computadora a una red LAN o Wireless LAN esto se hace utilizando un dispositivo llamado tarjeta de red o *NIC* (Network Interface Card), la cuál permite conectar el nodo (la computadora) a una red a través de un canal. En NS-3 existe el mismo concepto, una clase llamada *NetDevice* es instalada en cada nodo para permitir que este se conecte a través de un canal, representado por un Objeto de la clase *Channel*.

Esta arquitectura en el simulador permite la libertad de que un solo nodo pueda estar conectado a más de un canal utilizando múltiples dispositivos de red (*NetDevices*).

Para que la tarea de crear y manejar los nodos de una simulación sea más sencilla en NS-3 se tiene una clase llamada *NodeContainer*. Otras tareas como asignar direcciones IP, añadir una dirección MAC al *NetDevice*, instalar ese dispositivo de red en un nodo, configurar la pila de protocolos del nodo son realizadas por otras clases llamadas *Helpers*

Lo expresado anteriormente permite entender que para las simulaciones de redes VANET, en este trabajo, los nodos son los vehículos que se conectan en una red MANET puesto que son nodos móviles (*Mobility*). Cada nodo (*Node*) posee un dispositivo de red *NetDevice* que permite conectarse a través de un canal (*Channel*) inalámbrico (*Wireless*) utilizando un estándar, el IEEE 802.11p.

Una vez creada la simulación, algunos de estos nodos serán encargados de enviar información a través de la red generada a través de una aplicación (*Application*).

La Figura 4.1 se muestra el diseño del modelo de ataque de agujero gris; para diseñarlo se buscó información de otros trabajos similares basados en ataques Blackhole y dado que este tipo de ataque está basado en modificar la capa de Red del modelo OSI fue necesario investigar cómo se había hecho anteriormente.

Como ambos tipos de ataque tienen que ver con modificar a favor del atacante las tablas de ruteo mantenidas en cada nodo entonces se hizo necesario modificar la parte de la capa de red encargada de crear y mantener estas tablas. Los protocolos de ruteo, como se mostró en capítulos anteriores, son la base de esta tarea en las redes MANET.

El simulador NS-3, al ser de código abierto y ser mantenido por una amplia red de colaboradores, incluye algunos de los protocolos de ruteo más comunes; de estos el algoritmo AODV es el que más se usa dentro de la investigación de seguridad en redes VANET.

La modificación de este protocolo de ruteo supuso realizar cambios la instalación del simulador y no solo la creación de una nueva simulación. Los cambios sustanciales se llevaron a cabo en el archivo *aodv-routing-protocol.c* haciendo mejoras a los cambios propuestos y realizados por Shalini Satre en [32].

El cambio realizado modifica el protocolo de tal manera de que si el nodo asignado es un nodo adversario (nodo maligno) entonces responderá al mensaje búsqueda de cualquier nodo con un mensaje RREQ con un número de secuencia más alto y una secuencia de salto con solo un uno, con lo cual el nodo adversario hará creer a los demás nodos cercanos que su ruta es la mejor opción para crear su tabla de ruteo.

Una vez que el nodo es colocado en la tabla de ruteo de los nodos cercanos, todo el tráfico que pase por ahí se perderá porque no lo enviará a ningún otro lado. En inglés se utiliza el término *dropping packet*

Como la diferencia entre el ataque de agujero negro y el agujero gris estriba básicamente en la intermitencia del mal comportamiento del nodo en el ataque de agujero gris; al crear la simulación se estableció un parámetro aleatorio que permitía que el nodo malicioso dejara de serlo en cualquier momento dentro del tiempo de simulación; esto hace que el detectar este comportamiento malicioso sea más complicado para un sistema de detección de Intrusos.

Al realizar las simulaciones, el tiempo de simulación se varió entre 100 y 1000 segundos, el número de vehículos también fue una variable, variando entre 10 y 50. Los nodos maliciosos se cambiaron aleatoriamente en cada experimento. Cada evento se ejecutó 5 veces con diferentes valores de inicialización. Se calculó la media de estos resultados para obtener las múltiples muestras utilizadas en la clasificación.

Los resultados de la simulación, recogidos a través de los archivos PCAP, se utilizaron para obtener características relevantes sobre el comportamiento de los nodos, como se observa en la Figura 4.1.

Para extraer las características del tráfico de red se desarrolló un script en Python. Los archivos PCAP pueden ser leídos por muchas aplicaciones o lenguajes de programación, en este caso se seleccionó Python para extraer la información de 30 características de los flujos de comunicación entre los nodos, para esto se utilizó la biblioteca pyshark.

Generación de flujos y características

Un flujo se define en el sentido convencional como una serie de paquetes que comparten la misma penta-tupla (dirección IP y puerto de origen, dirección IP y puerto de destino y número de protocolo) [19]. Un script de Python extrae la información del flujo y calcula al mismo tiempo las características asociadas a cada flujo. Las características extraídas se dividen en tres conjuntos: características basadas en el tiempo, características relacionadas con los paquetes y características del protocolo de enrutamiento.

En la literatura, se pueden encontrar estos trabajos [20,21] que utilizan características extraídas de los flujos centradas en el tiempo. En este conjunto el tiempo es una medida del comportamiento del flujo, por ejemplo, el tiempo entre paquetes u obtener una medida del tiempo de arribo de datos, por ejemplo, bytes por segundo o paquetes por segundo.

Las características relacionadas con los paquetes son el número de paquetes transmitidos en el flujo, el número de paquetes transmitidos hacia adelante (del nodo A al B), el número de paquetes hacia atrás (del nodo B al nodo A), el periodo entre tramas (hacia adelante y hacia atrás) dos medidas para cada una, la media y la desviación estándar. Como los paquetes tienen diferente cantidad de información, los bytes de carga útil se extraen y se seleccionan como características de la misma manera que los paquetes (hacia adelante, hacia atrás, media y desviación estándar).

Las características del proceso de enrutamiento dependen del protocolo seleccionado utilizado en las simulaciones, la revisión de la literatura muestra que dos características clave se utilizan ampliamente, el recuento de saltos y el número de secuencia de destino, ya que se transmiten varios paquetes en un flujo,

la información seleccionada sobre estos dos son la media y la desviación estándar. En algunos trabajos [1,7,22] los autores utilizan una información fuzzificada de estas características.

Resultados y análisis

Se utilizaron varios algoritmos de aprendizaje automático, como se observa en la Figura 4.1, para predecir el comportamiento de los nodos atacantes: Diversas formas de Support Vector Machine, Random Forest, Logistic Regression y clasificadores Naive Bayes. También se utilizaron múltiples medidas para evaluar estos algoritmos, como precisión, recall, y f1-score que son de amplio uso en la literatura.

La precisión (p) y la recuperación (r) pueden calcularse como [24]:

$$p = \frac{TP}{TP+FP} \quad (1)$$

$$r = \frac{TP}{TP+FN} \quad (2)$$

Donde: TP (TN) significa verdadero positivo (verdadero negativo) y FP (FN) significa falso positivo (falso negativo).

La puntuación F1 puede interpretarse como una media armónica ponderada de la precisión y la recuperación [23].

$$F_1 = 2 \frac{p \cdot r}{p+r} \quad (3)$$

La Tabla 4-2 y la Tabla 4-3 sintetizan los resultados de las mediciones realizadas por los algoritmos de clasificación con diferentes métodos de escalado (Estándar y Min-Max).

Tabla 4-2 Resultados de los clasificadores con el método de escalado Estándar

	Polynomic SVM	Radial Basis SVM	Logistic Regression	Random Forest	Naive Bayes
Recall	0.717	0.731	0.739	0.956	0.724
Precision	0.704	0.68	0.685	0.958	0.758
F1 Score	0.709	0.686	0.684	0.955	0.736

Tabla 4-3 Resultados de los clasificadores con el método de escalado Min-Max.

	Polynomic SVM	Radial Basis SVM	Logistic Regression	Random Forest	Naive Bayes
Recall	0.76	0.72	0.76	0.98	0.72
Precision	0.735	0.7	0.74	0.98	0.76
F1 Score	0.74	0.71	0.72	0.98	0.74

En todos los algoritmos se utilizó la validación cruzada estratificada de 10 veces para diferenciar los datos de entrenamiento de los de prueba. En cada pliegue, el conjunto de datos se eligió aleatoriamente, como se hace habitualmente según la literatura [21].

Ambas tablas muestran que el mejor algoritmo para clasificar el mal comportamiento de los nodos en el ataque de Gray-hole una VANET es Random Forest. El clasificador Random Forest utilizado tiene 10 estimadores con una profundidad máxima de 10. Cuando se ajustan estos dos parámetros a 15

estimadores y una profundidad de 15, Random Forest consigue 0,9927 en términos de Recall, Precision y F1 Score.

Análisis

La tabla 3 muestra un comportamiento de predicción bajo para las SVM. Las diferentes funciones del núcleo tienen diferente capacidad de aprendizaje y rendimiento de generalización. Entre todas las funciones kernel, observamos que el kernel de la función de base radial (RBF) tiene una mayor capacidad de interpolación y es bueno para reflejar las propiedades locales de las muestras, y su desventaja es que no puede extraer las características globales de todas las muestras de entrenamiento [26]. De hecho, el sobreajuste de la SVM con el kernel RBF sería inevitable hasta cierto punto si no se toman medidas especiales para superarlo [27]. El parámetro ajustado de la escala gamma, tal y como se afirma en [28], no muestra ningún cambio en el rendimiento de RBF, por lo que, dadas las características extraídas, el límite se sitúa en torno al 74,6% de precisión.

Para encontrar una base teórica para el bajo desempeño de los SVM, principalmente de la RBF-SVM (Máquina de Soporte Vectorial con kernel de base radial) en este trabajo y haciendo un análisis similar al realizado en [27] se encontró lo siguiente:

Dado un conjunto de datos de entrenamiento con m muestras (X_k, Y_k) , $k = 1:m$ cada muestra con n distintas características, donde $X_k \in \mathbb{R}^n$ es una muestra con una etiqueta de dos clases $Y_k \in \{-1, 1\}$, dos métricas ampliamente conocidas: $d_{ij} = \|X_i - X_j\|$ la distancia euclidiana entre dos muestras diferentes y $\Delta_{ijk} = |X_{ki} - X_{kj}|$ la distancia Manhattan entre dos características dadas de muestras diferentes, denominada como la diferencia absoluta entre X_i y X_j en la k -ésima característica [27].

Sea $\delta_{ij} = \max_k [\Delta_{ijk}]$ la máxima diferencia absoluta (MAD) que mide la diferencia mayor de expresión entre todas las características para dos muestras.

Es fácil notar que, para todas las muestras del conjunto de datos, la relación entre la distancia euclidiana y la MAD está siempre entre 1 y $\sqrt[2]{n}$, eso significa que:

$$1 \leq \frac{d_{ij}}{\delta_{ij}} \leq \sqrt[2]{n} \quad (4)$$

Utilizando este resultado se llega a la siguiente afirmación, dado cualquiera δ_{ij} existe un d_{ij} tal que el cociente es igual a 1. Por lo que d_{ij} es el mínimo d_{min} . Sustituyendo este resultado en la expresión (4) se obtiene

$$d_{min} \leq d_{ij} \leq \sqrt[2]{n}\delta_{ij} \quad (5)$$

La distancia Euclidiana entre las muestras de entrenamiento X_i y X_j en el espacio de características de una RBF-SVM es la entrada K_{ij} en la matriz del núcleo de la máquina de aprendizaje en el proceso de aprendizaje, así:

$$\kappa_K(X_i, X_j) = e^{\frac{-\|X_i - X_j\|^2}{2\sigma^2}} \quad (6)$$

La expresión (6) puede reescribirse con base en la distancia euclidiana:

$$K_{ij} = \kappa_K(X_i, X_j) = e^{-d_{ij}^2/2\sigma^2} \quad (7)$$

Las expresiones (7) y (5) pueden combinarse para obtener el siguiente resultado

$$e^{-n\delta_{ij}^2/2\sigma^2} \leq e^{-d_{ij}^2/2\sigma^2} \leq e^{-d_{min}^2/2\sigma^2} \quad (8)$$

De acuerdo con (8) la entrada K_{ij} la entrada en la matriz del núcleo de la máquina de aprendizaje en el proceso de aprendizaje está limitada superiormente por el término d_{min} . Eligiendo $\sigma = 1$ como se afirma en la literatura conduce a la expresión $K_{ij} = \exp^{d_{min}^2/2}$ como el límite superior en el proceso de entrenamiento. Es interesante observar que $d_{min} = 10$ produce un valor de $K_{ij} = 2 \times 10^{-22} \sim 0$.

Típicamente, algunos términos K_{ij} se convierten en cero durante el entrenamiento [28]. Pero ¿Qué es lo que sucedería si casi todos los K_{ij} son cero o casi cero? EL algoritmo RBF-SVM siempre hará overfitting (sobreajuste) y siempre generalizará el resultado basado en cuál clase tiene mayoría dentro de

los datos de entrenamiento, La RBF-SVM, en este supuesto, pierde la generalización.

Para probar esto, dado $Y_k \in \{-1,1\}$, se llamará m_{-1} al número de muestras en el conjunto de entrenamiento donde $Y = -1$, y m_{+1} el número de muestras en el conjunto de entrenamiento donde $Y = 1$, y sea $m = m_{+1} + m_{-1}$ la longitud total de las muestras

Como se afirma en la literatura [27,28] la función de decisión de la RBF-SVM se calcula como:

$$f(X_i) = \text{sign}\left(\sum_{j=1}^m \alpha_j e^{d_{\min}^2/2} + b\right) \quad (9)$$

si casi todos los K_{ij} son cero o cerca de cero entonces está claro que la función de decisión dependerá totalmente del vector umbral b .

$$f(X_i) = \text{sign}(b) \quad (10)$$

Esto está claro porque el vector b está definido por el vector de pesos debido a la condición de Karush-Kuhn-Tucker (KKT) [29]:

$$b = \frac{1}{m} \sum_{i=1}^m (Y_i - \sum_{i=1}^m \alpha_i K_{ij}) \quad (11)$$

La ecuación 11, como se afirma en la literatura, es la media de todas las muestras del conjunto de entrenamiento.

Pero, cuando $K_{ij} \sim 0$:

$$b_0 = \frac{1}{m} \sum_{i=1}^m Y_i \quad (12)$$

Utilizando el hecho de que $Y_k \in \{-1,1\}$ esto lleva al siguiente resultado

$$b = \frac{1}{m} [\sum_{i=1}^{m_{+1}} 1 - \sum_{i=1}^{m_{-1}} 1] = \frac{m_{+1} - m_{-1}}{m} \quad (13)$$

Entonces se tiene que

$$f(X_i) = \text{sign}\left(\frac{m_{+1} - m_{-1}}{m}\right) \quad (14)$$

Según la definición de la función de *sign* [27] la función de decisión se simplifica aún más, de tal forma que

$$f(X_i) = \text{sign}(m_{+1} - m_{-1}) \quad (15)$$

Es evidente que la decisión sobre el tipo de clase se encontraría basado únicamente en conocer cuál clase tiene más muestras en el conjunto de entrenamiento. Si no hay un tipo mayoritario, la máquina de aprendizaje no puede determinar el tipo de clase de la muestra de entrada. Esto significa que el uso de la métrica de la distancia euclidiana entre las muestras del conjunto de entrenamiento, incluso antes del entrenamiento, dará una visión rápida de si habrá overfitting en una SVM-RBF.

Los datos extraídos utilizados en este trabajo muestran una mayoría de la clase comportamiento normal del 74%, mientras que la otra (mal comportamiento) tan solo del 26%. La tabla 3 muestra una aproximación muy clara a estos resultados, esto sucede porque la distancia euclidiana entre las diferentes muestras en los datos de entrenamiento utilizados en este trabajo presenta un valor mayor que el $d_{min} = 10$ que conduce al overfitting en el algoritmo RBF-SVM.

Con respecto a los resultados del kernel polinómico (POL-SVM), utilizando la expansión en serie de Taylor para expresar la ecuación (6) queda de manifiesto la

relación del kernel RBF con el kernel polinómico. Cada término de la serie es un monomio de X_i y X_j .

$$\kappa_K(X_i, X_j) = e^{\frac{-\|X_i - X_j\|^2}{2\sigma^2}} = e^{\frac{-\|X_i^2\|}{2\sigma^2}} e^{\frac{-\|X_j^2\|}{2\sigma^2}} e^{\frac{\langle X_i X_j \rangle}{\sigma^2}} \quad (16)$$

$$e^{\frac{\langle X_i X_j \rangle}{\sigma^2}} = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{\langle X_i X_j \rangle}{\sigma^2} \right)^k \quad (16 \text{ a})$$

$$\kappa_K(X_i, X_j) = e^{\frac{-\|X_i - X_j\|^2}{2\sigma^2}} = e^{\frac{-\|X_i^2\| + \|X_j^2\|}{2\sigma^2}} \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{\langle X_i X_j \rangle}{\sigma^2} \right)^k \quad (16 \text{ b})$$

Representando el kernel RBF como un producto interno de dos funciones se puede hacer $\kappa_K(X_i, X_j) = \langle \phi(X_i), \phi(X_j) \rangle$ donde cada $\phi(X)$ es una función del espacio de características.

$$\phi_{k,j}(X) = e^{\frac{-\|X^2\|}{2\sigma^2}} \frac{1}{\sigma^2 \sqrt{k!}} \prod_{i=0}^k X_{ji} \quad (17)$$

Si bien las expresiones (16a) y (16b) muestran una suma infinita, la teoría de Taylor permite realizar un truncamiento para obtener un polinomio de grado $r < k$. El error introducido por este truncamiento se encuentra definido en [30].

Por otro lado, el kernel polinómico es también una suma de monomios entre X_i y X_j . De manera más precisa, las características que corresponden al kernel polinómico también pueden escribirse como

$$\phi_{k,j}(X) = \sqrt{\binom{r}{k}} c^{r-k} \prod_{i=0}^k X_{ji} \quad (18)$$

Observando las ecuaciones 17 y 18, existen diferencias de escalamiento entre ambos polinomios, el más importante es el escalamiento que depende del grado k del polinomio, y que es el inverso del factorial del grado. Lo cual significa que, para monomios de más alto grado, el escalamiento es menor y que por lo tanto se entiende que el algoritmo POL-SVM tiende a basarse más en los monomios de menor grado.

De esto se desprende que el comportamiento del kernel polinómico se aproxima asintóticamente [25,26,27,30] al comportamiento del kernel RBF, y el incrementar el orden del polinomio no afecta este hecho, puesto que el factor de escalamiento se va aproximando a cero, y no tiene mayor incidencia en este comportamiento.

Por lo que, en vista de este resultado, el comportamiento del algoritmo POL-SVM en la tabla 3 corresponde con estos experimentos y con la experiencia en los resultados de presentados por otros autores [30].

Conclusiones

Los resultados de este trabajo de investigación se encuentran divididos de manera cualitativa y de manera cuantitativa.

De manera cualitativa, se ha realizado una revisión profunda del estado del arte en el desarrollo de soluciones de ciberseguridad para el área automotriz, específicamente en el área de las redes automotrices inalámbricas. Estas soluciones se encuentran englobadas dentro de lo que se conoce como cómputo inteligente, Redes Neuronales, Sistemas difusos, Aprendizaje automático.

De esta investigación se ha visto una tendencia hacia el paradigma del aprendizaje de máquina (ML) o aprendizaje automático. Dentro de esta área se ha verificado que los diferentes grupos de trabajo han diversificado su esfuerzo por encontrar qué algoritmo es el que mejor puede resolver las dificultades que representa la ciberseguridad de las VANET.

Los métodos y modelos encontrados son variados, dependiendo en gran medida del área de investigación del grupo de trabajo en cuestión; en ese sentido no existe una tendencia clara, más aun, las partes interesadas: fabricantes de automóviles, entidades gubernamentales, entidades de estandarización no gubernamentales y empresas, han encontrado también dificultad en unificar criterios en cuanto a ciberseguridad automotriz.

Un primer paso fue en 2015 cuando apareció la norma SAE J306, pero los hallazgos encontrados en este estado del arte permiten afirmar que ningún equipo

de trabajo la menciona en sus investigaciones ni se usa como base para el desarrollo de soluciones de ciberseguridad.

Un nuevo paso hacia una definición clara de medidas que permitan incrementar la ciberseguridad de los automóviles, incluidas las VANET es la norma ISO/SAE21434:2021 publicada en Septiembre de 2021. “Este documento especifica los requisitos de ingeniería para la gestión de los riesgos de ciberseguridad en relación con el concepto, el desarrollo del producto, la producción, el funcionamiento, el mantenimiento y el desmantelamiento de los sistemas eléctricos y electrónicos (E/E) de los vehículos de carretera, incluidos sus componentes e interfaces” [31].

Otra norma, todavía en desarrollo es la ISO/DPAS 5112 la cual regulará las directrices para la auditoría de la ingeniería de ciberseguridad en vehículos. En ambos casos se debe esperar el impacto que estos dos estándares tengan en las propuestas de investigación futuras.

En términos cualitativos, de la revisión del estado del arte se desprende otra afirmación; los equipos de trabajo que presentan los resultados de sus investigaciones de forma cuantitativa, esto como la capacidad de un agente inteligente de detectar un comportamiento incorrecto o abusivo dentro de las VANET, no presentan un análisis cualitativo de porqué sus investigaciones llegan a las conclusiones que presentan.

Existe un mínimo examen de los resultados que se presentan que permita realizar una evaluación en las metodologías propuestas, por ejemplo, si estas pueden o no ser mejoradas; cuáles son los criterios de selección de los algoritmos utilizados o resultados que no permiten reconocer cuáles son las fortalezas del algoritmo que mejor ajusta los parámetros elegidos.

Referencias

[1] Yang, X., Yi, X., Khalil, I., Zeng, Y., Huang, X., Nepal, S., & Cui, H. (2019). A lightweight authentication scheme for vehicular ad hoc networks based on MSR. Veh. Commun., 15, 16-27.

- [2] M. Saini, A. Alelaiwi, and A. E. Saddik, "How Close are We to Realizing a Pragmatic VANET Solution? A Meta-Survey," *ACM Computing Surveys*, vol. 48, no. 2, pp. 1–40, Nov. 2015.
- [3] C. Miller and C. Valasek, Remote exploitation of an unaltered passenger vehicle, 2015, [online] Available: <http://illmatics.com/Remote%20Car%20Hacking.pdf>.
- [4] Ahmad, F., Adnane, A. and Franqueira, V. (2016) A Systematic Approach for Cyber Security in Vehicular Networks. *Journal of Computer and Communications*, 4, 38-62. doi: 10.4236/jcc.2016.416004.
- [5] K. M. A. Alheeti, A. Gruebler, and K. McDonald-Maier, "Using discriminant analysis to detect intrusions in external communication for self-driving vehicles," *Digit. Commun. Networks*, vol. 3, no. 3, pp. 180–187, 2017, doi: 10.1016/j.dcan.2017.03.001.
- [6] R. Coussement, B. Amar Bensaber, and I. Biskri, "Decision support protocol for intrusion detection in VANETs," 2013, doi: 10.1145/2512921.2512928.
- [7] Z. Hassan, A. Mehmood, C. Maple, M. A. Khan, and A. Aldegheishem, "Intelligent detection of black hole attacks for secure communication in autonomous and connected vehicles," *IEEE Access*, vol. 8, pp. 199618–199628, 2020, doi: 10.1109/ACCESS.2020.3034327.
- [8] Y. Zeng, M. Qiu, D. Zhu, Z. Xue, J. Xiong, and M. Liu, "DeepVCM: A Deep Learning Based Intrusion Detection Method in VANET," 2019, doi: 10.1109/BigDataSecurity-HPSC-IDS.2019.00060.
- [9] J. Grover, N. K. Prajapati, N. Kumar Prajapati, V. Laxmi, and M. Singh Gaur, "Machine Learning Approach for Multiple Misbehavior Detection in VANET," 2011, doi: 10.1007/978-3-642-22720-2_68.
- [10] K. M. A. Alheeti, A. Gruebler, and K. D. McDonald-Maier, "An Intrusion Detection System against Black Hole Attacks on the Communication Network of Self-Driving Cars," 2016, doi: 10.1109/EST.2015.10.

- [11] M. Erritali, B. El Ouahidi, B. Hssina, B. Bouikhalene, and A. Merbouha, "An ontology-based intrusion detection for vehicular ad hoc networks," *J. Theor. Appl. Inf. Technol.*, 2013.
- [12] Vatambeti, R, Supriya, KS, Sanshi, S. Identifying and detecting black hole and gray hole attack in MANET using gray wolf optimization. *Int J Commun Syst.* 2020; 33: e4610. <https://doi.org/10.1002/dac.4610>
- [13] E. Eziam, K. Tepe, A. Balador, K. S. Nwizege, and L. M. S. Jaimes, Malicious Node Detection in Vehicular Ad-Hoc Network Using Machine Learning and Deep Learning. 2018 IEEE Globecom Workshops (GC Wkshps), IEEE, 2018, pp. 1–6,
- [14] K. M. A. Alheeti, A. Gruebler, and K. D. McDonald-Maier, "On the Detection of Grey hole and Rushing Attacks in Self-Driving Vehicular Networks."
- [15] E. Eziam, K. Tepe, A. Balador, K. S. Nwizege, and L. M. S. Jaimes, "Malicious Node Detection in Vehicular Ad-Hoc Network Using Machine Learning and Deep Learning," in 2018 IEEE Globecom Workshops, GC Wkshps 2018 - Proceedings, Feb. 2019, pp. 1–6, doi: 10.1109/GLOCOMW.2018.8644127.
- [16] Z. Hassan, A. Mehmood, C. Maple, M. A. Khan, and A. Aldegheishem, "Intelligent Detection of Black Hole Attacks for Secure Communication in Autonomous and Connected Vehicles," doi: 10.1109/ACCESS.2020.3034327.
- [17] Y. Begriche, R., L. Khoukhi, and C. Xiuzhen, "Bayesian-based model for a reputation system in vehicular networks," 2015 Int. Conf. Cyber Secur. Smart Cities, Ind. Control Syst. Commun., 2015, doi: 10.1109/SSIC.2015.7245332.
- [18] Sun, Zemin, Yanheng Liu, Jian Wang, Guofa Li, C. Anil, Keqiang Li, Xinyu Guo, Geng Sun, D. Tian and D. Cao. "Applications of Game Theory in Vehicular Networks: A Survey." *ArXiv abs/2103.11874* (2021)
- [19] A. Mabrouk, A. Kobbane, and M. El Koutbi, "Signaling game-based approach to improve security in vehicular networks." in *VEHITS*, 2018, pp. 495–500

- [20] A. McGregor, M. Hall, P. Lorier, and J. Brunskill, "Flow clustering using machine learning techniques," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 3015, pp. 205–214, 2004, doi: 10.1007/978-3-540-24668-8_21.
- [21] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," *ICISSP 2018 - Proc. 4th Int. Conf. Inf. Syst. Secur. Priv.*, vol. 2018-Janua, no. Cic, pp. 108–116, 2018, doi: 10.5220/0006639801080116.
- [22] A. H. Lashkari, G. D. Gil, M. S. I. Mamun, and A. A. Ghorbani, "Characterization of tor traffic using time-based features," in *ICISSP 2017 - Proceedings of the 3rd International Conference on Information Systems Security and Privacy*, 2017, vol. 2017-Janua, pp. 253–262, doi: 10.5220/0006105602530262.
- [23] H. Moudni, M. Er-Roudi, H. Mouncif, and B. El Hadadi, "Black Hole attack Detection using Fuzzy based Intrusion Detection Systems in MANET," *Procedia Comput. Sci.*, vol. 151, pp. 1176–1181, 2019, doi: 10.1016/j.procs.2019.04.168.
- [24] C. Goutte and E. Gaussier, "A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation," *Lect. Notes Comput. Sci.*, vol. 3408, pp. 345–359, 2005, doi: 10.1007/978-3-540-31865-1_25.
- [25] M. S. Alsahli, M. M. Almasri, M. Al-Akhras, A. I. Al-Issa, and M. Alawairdhi, "Evaluation of Machine Learning Algorithms for Intrusion Detection System in WSN," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 5, pp. 617–626, 2021, doi: 10.14569/IJACSA.2021.0120574.
- [26] Y. Liu and D. Pi, "A novel kernel SVM algorithm with game theory for network intrusion detection," *KSII Trans. Internet Inf. Syst.*, vol. 11, no. 8, pp. 4043–4060, 2017, doi: 10.3837/tiis.2017.08.016.
- [27] H. Han and X. Jiang, "Overcome Support Vector Machine Diagnosis Overfitting," *Cancer Inform.*, vol. 13s1, p. CIN.S13875, 2014, doi: 10.4137/cin.s13875.

- [28] E. A. Zanaty, "Support Vector Machines (SVMs) versus Multilayer Perception (MLP) in data classification," *Egypt. Informatics J.*, vol. 13, no. 3, pp. 177–183, 2012, doi: 10.1016/j.eij.2012.08.002.
- [29] C. Savas and F. Dovic, "The impact of different kernel functions on the performance of scintillation detection based on support vector machines," *Sensors (Switzerland)*, vol. 19, no. 23, pp. 1–16, 2019, doi: 10.3390/s19235219.
- [30] A. Cotter, J. Keshet, and N. Srebro, "Explicit Approximations of the Gaussian Kernel," pp. 1–11, 2011, [Online]. Available: <http://arxiv.org/abs/1109.4603>.
- [31] "ISO - 43.040.15 - Car informatics. On board computer systems", *Iso.org*, 2021. [Online]. Available: <https://www.iso.org/ics/43.040.15/x/>. [Accessed: 09- Nov- 2021]
- [23] M. Tahiliani and S. Satre, "[ns-3] Blackhole Attack Simulation in ns-3", *Mohittahiliani.blogspot.com*, 2021. [Online]. Available: <http://mohittahiliani.blogspot.com/2014/12/ns-3-blackhole-attack-simulation-in-ns-3.html>. [Accessed: 02- Nov- 2021]