Entre Ciencia e Ingeniería, vol. 16, no. 31, enero-junio de 2022, páginas 17-24. DOI: https://doi.org/10.31908/19098367.2655
ISSN 1909-8367 (Impreso), ISSN 2539-4169 (En línea)

Inteligencia artificial para el control de tráfico en redes de datos: Una Revisión[1]

Artificial intelligence for traffic control in data networks: A Review

D. A. León, J. G. Martínez, I. A. Ardila y D. J. Mosquera

Recibido: mayo 31 de 2021 – Aceptado: mayo 15 de 2022

***Resumen*—El control del tráfico en las redes de datos ha cobrado gran importancia en los últimos tiempos debido al uso masivo que se le está dando a las redes informáticas en distintos ámbitos de la sociedad. Con el fin de realizar un control de tráfico efectivo, se suele hacer uso de diferentes técnicas que permiten, entre otras cosas, clasificar, predecir y monitorear el tráfico de la red. Estas técnicas han ido evolucionando y actualmente se apoyan en métodos de inteligencia artificial, lo cual ha permitido mejorar los resultados obtenidos con las técnicas convencionales. El presente artículo recopila los diferentes aportes realizados por el campo de la inteligencia artificial al mejoramiento de estas técnicas y a la gestión de redes en general. Se describen los aportes realizados en aspectos tales como la seguridad, la predicción y clasificación del tráfico de datos, así como la optimización del ruteo en una red informática.**

***Palabras clave*— Gestión de tráfico, técnicas de control de tráfico, inteligencia artificial, aprendizaje automático, aprendizaje profundo.**

***Abstract*— Traffic control in data networks has recently become very important due to the massive use of computer networks in different areas of society. Different techniques are usually used to carry out effective traffic control, allowing, among other things, to classify, predict and monitor network traffic. These techniques have evolved and are currently supported by artificial intelligence tools, which have made it possible to improve the results obtained with conventional techniques. This paper collects the different contributions made by the field of artificial intelligence to the improvement of these techniques and network management in general. The article describes the contributions made in aspects such as security, prediction, and classification of data traffic, as well as the optimization of routing in a computer network.**

***Keywords*—Traffic management, traffic control techniques, artificial intelligence, machine learning, deep learning.**

I. Introducción

En los últimos tiempos las redes de datos han tomado gran importancia en el diario vivir de los seres humanos. Dichas redes son usadas en una gran cantidad de campos tales como la industria de las telecomunicaciones, la industria bancaria, la industria turística, la industria del entretenimiento, entre muchas otras. Sin embargo, con el aumento del uso de las redes de datos se han acrecentado también una serie de problemas con los cuales los administradores de redes deben tratar para garantizar el buen funcionamiento de la mismas, entre dichos problemas podemos encontrar aspectos como evitar la congestión de la red, garantizar que todos los recursos de la red se compartan de manera eficiente entre los usuarios y las aplicaciones, además de detectar tráfico inusual que pueda afectar el funcionamiento de la misma.

Con el fin de cumplir tales exigencias, se suele hacer uso de una serie de técnicas que permiten, entre otras cosas, controlar el flujo de la red, realizar la clasificación del tráfico, realizar predicciones de comportamientos futuros y detectar posibles amenazas. Estas técnicas están en constante evolución y en los últimos años han venido siendo objeto de estudios que pretenden mejorar su eficacia y brindar resultados más acordes a la realidad.

Uno de los grandes inconvenientes que presentan las técnicas convencionales es que, dada la inmensidad de datos que se deben tener en cuenta, se hace difícil priorizar la información de manera tal que se puedan atacar los aspectos más críticos que afectan el funcionamiento de la red. Lo anterior hace que

[1]Producto derivado del grupo de investigación "ORION", apoyado por la Universidad Distrital Francisco José De Caldas a través del proyecto de Ingeniería en Telemática.

D. A. León, Universidad Distrital Francisco José De Caldas, Bogotá, Colombia, email: daleons@correo.udistrital.edu.co.
J. G. Martínez, Universidad Distrital Francisco José De Caldas, Bogotá, Colombia, email: jgmartinezc@correo.udistrital.edu.co.
I. A. Ardila, Universidad Distrital Francisco José De Caldas, Bogotá, Colombia, email: iaardila@udistrital.edu.co.
D. J. Mosquera, Universidad Distrital Francisco José De Caldas, Bogotá, Colombia, email: djmosquerap@udistrital.edu.co.

Como citar este artículo: D. A. León, J. G. Martínez, I. A. Ardila y D. J. Mosquera. Inteligencia artificial para el control de tráfico en redes de datos: Una Revisión, Entre Ciencia e Ingeniería, vol. 16, no. 31, pp. 17-24, enero-junio 2022. DOI: https://doi.org/10.31908/19098367.2655.

se dificulte la toma de decisiones y por ende se presenten demoras en la ejecución de planes de corrección, prevención y optimización, lo que termina por afectar los indicadores de desempeño de la red.

Actualmente, la inteligencia artificial brinda las herramientas necesarias para poder analizar grandes volúmenes de datos de una manera más precisa y eficiente, sin necesidad de la intervención humana, de manera tal que podría ayudar a solucionar problemas relacionados a la administración de redes, entre ellos los problemas de control de tráfico mencionados anteriormente.

De acuerdo con lo anterior, el trabajo propuesto pretende realizar una revisión bibliográfica sobre los aportes realizados por la inteligencia artificial en la gestión o control de tráfico de las redes, con el fin de informar al lector sobre dichos aportes y las implicaciones de su implementación. Para esto, el artículo está organizado de la siguiente forma: en la sección II se presenta el concepto de control de tráfico, en la sección III se exponen los retos asociados al control de tráfico, en la sección IV se muestran algunas de las aplicaciones de la IA en el control de tráfico. Por último, en la sección V se plantean las conclusiones.

II. Control de tráfico

Ocasionalmente, el control de tráfico de datos en las redes de comunicaciones es percibido como alguna de sus funciones o comprendido sólo parcialmente, es decir, se puede llegar a considerar el control de tráfico como el control del flujo de datos, el control de la congestión de la red o como la capacidad de permitir o denegar el acceso a los recursos de la red a sus usuarios. Pero como veremos a continuación el control de tráfico es un término más general y comprende entre otros los anteriores aspectos mencionados.

A. Control de la congestión

Para Chang y Su [1] el control de la congestión se refiere al mecanismo o conjunto de acciones tomadas por la red para afrontar las situaciones en las cuales el tráfico total ofrecido excede la capacidad misma de la red (congestión). Así mismo, la gestión del tráfico según el IETF [2], se refiere a la implementación de acciones que permita evitar o reducir la condición de congestión de tráfico en la red, a través de la clasificación.

Realizar una oportuna gestión de la congestión del tráfico es fundamental en una red y sus acciones deben tomarse de manera proactiva, ya que una vez congestionada la red, las acciones correctivas pueden ser muy complejas [3].

B. Control de flujo

De manera sencilla el control de flujo es la capacidad de la red de controlar o restringir el flujo de datos que se transportan por ella de la manera más conveniente para la red [4]. Dentro de las funciones de este tipo de control encontramos la medición de esos flujos de datos, la cual nos proporciona información que permite: "comprender el comportamiento de las redes existentes, planificar el desarrollo y la expansión de la red, cuantificar el rendimiento de la red, verificar la calidad del servicio de la red y atribuir el uso de la red a los usuarios" [5].

C. Control de tráfico

En las redes de transmisión de datos, la gestión o control del tráfico de datos implica varias acciones que buscan prevenir y/o reducir la congestión (gestión de la congestión) a partir de, por ejemplo, técnicas como el control de flujo; para cumplir con los objetivos de rendimiento de la red [1]. Entonces, el control del tráfico de la red es el proceso de administrar, controlar o reducir el tráfico de la red. Este proceso se realiza con el fin de utilizar de manera eficiente el ancho de banda de la red, mantener baja la latencia, ofrecer calidad de servicio y compartir de manera justa los recursos entre muchos usuarios y aplicaciones mediante la gestión del flujo de tráfico [6].

D. Técnicas de control de tráfico

Existen algunas técnicas comúnmente empleadas para controlar el flujo de datos que circulan por una red de datos, en este apartado tomamos el enfoque de Noormohammadpour y Raghavendra [6], para mencionar algunas de esas técnicas. Los autores sugieren un grupo de seis técnicas:

Control de transmisión: Esta técnica controla el flujo de datos enviados por la red, cuyo objetivo es reducir la latencia, reducción de la tasa de error/retrasos, maximizar la utilización y la equidad en la repartición de recursos. Por otro lado, los retos que se deben asumir con esta técnica son la ráfaga de tráfico y el reordenamiento de paquetes [6].

Conformación del tráfico o Traffic Shaping: consiste en "reprogramar (retrasar) los paquetes (o tramas) hasta que alcanzan el ancho de banda especificado o los límites de ráfagas. Dado que tales retrasos implican colas que casi siempre son finitas y, una vez que están llenas, el exceso de tráfico casi siempre se descarta, la configuración del tráfico casi siempre implica también la vigilancia del tráfico" [6].

Priorización: en esta técnica se establecen ciertos niveles o clasificaciones a los flujos de datos según su prioridad y de acuerdo con esta, son manejados, permitiendo reducir la latencia y la tasa de error/retrasos [6].

Balanceo de carga: Esta técnica permite aumentar el nivel de uso de la red en general, al distribuir la carga de los flujos de datos de manera equilibrada, evitando que ciertos enlaces de la red no se utilicen y por el contrario otros enlaces se vean saturados o congestionados [6].

Multipathing: permite fraccionar un flujo de datos en varios subflujos, para que estos sean enviados a través de diferentes rutas. En consecuencia, el receptor se debe encargar de tomar cada subflujo y posteriormente ordenarlos, aumentando la utilización [6].

Programación o Scheduling: Su objetivo principal es minimizar el tiempo que tarda un flujo de datos desde que sale del emisor hasta llegar a su destino final. Esto a partir de una formulación de la optimización de tiempo de respuesta de la red, que contempla varias métricas como la utilización, latencia y equidad de recursos, por lo que suele ser costoso a

nivel computacional [6].

III. Retos del control de tráfico

Uno de los grandes inconvenientes que se presentan al realizar el control de tráfico de una red es la cantidad y heterogeneidad de los datos que se deben tratar para controlar dicho tráfico. Esto se debe a que, como explican Li y Moore [7], el tráfico de una red es el producto de un sistema en el que interviene diferentes factores, entre los cuales podemos encontrar los hosts, las aplicaciones, los protocolos y por supuesto a los usuarios que interactúan entre si haciendo uso de los servicios brindados por la red. De esta manera, en cuanto más complejo sea el sistema sobre el cual funciona la red mayor será el tamaño y diversidad de la información que tendremos que tratar para poder tener control sobre el tráfico de la misma.

De igual manera, como indica Amaral et al. [8], se pueden presentar casos en el que dentro de una red que incluya una amplia gama de aplicaciones, las muestras de flujos no se encuentren correctamente etiquetadas, lo que representa un gran inconveniente debido a que se dificulta encontrar relaciones entre los datos que obtenemos del tráfico de la red y, por ende, ejercer un control de tráfico efectivo.

Por último, al realizar el proceso de control de tráfico se puede presentar algunos eventos que dificultan la medición del tráfico de la red [6]:

- La aparición de ráfagas en el flujo de datos que dificultan la medición del tráfico ya que los resultados pueden variar dependiendo del lugar de la red en el cual se realice dicha medición.
- Cuellos de botella que se generan cuando varios hosts envían el tráfico hacia un mismo destino de manera concurrida.
- El uso excesivo de CPU en procesos como el reordenamiento de paquetes.

IV. Aplicaciones de la IA en el control de tráfico

A. Clasificación del tráfico de una red

Para la IETF [9], la clasificación del tráfico es un proceso automatizado que clasifica el tráfico de la red de acuerdo con diferentes parámetros como lo son el puerto, el protocolo, etc. La clasificación de tráfico juega un papel importante en la seguridad y la gestión de la red, debido a que permite detectar anomalías y desmejoras en la calidad del servicio con el fin tomar acciones tales como bloquear los flujos y recursos de la red [10].

Existen cuatro métodos principales de clasificación de tráfico:

Inicialmente, tenemos la clasificación de tráfico basada en puertos, en la cual se hace uso del puerto asignado por la Autoridad de Números Asignados de Internet (IANA) para realizar la identificación de las aplicaciones que hacen uso de la red [8]. Este método fue ampliamente usado en sus inicios, pero pasó a estar deprecado luego de que aparecieran aplicaciones que hacían uso de mecanismos de asignación de puertos dinámicos y de técnicas que permitían ocultar la información relacionada al puerto de la aplicación.

En segundo lugar, tenemos a la clasificación basada en inspección profunda de paquetes (DPI), en la cual se inspecciona la información de los paquetes en busca de patrones que identifiquen la aplicación que generó el tráfico [8]. Este método tiene algunas desventajas entre las que podemos encontrar el alto consumo de recursos que se usa en la inspección de los paquetes y la imposibilidad de operar sobre paquetes generados por aplicaciones que hacen uso de mecanismos de encriptación.

Finalmente, tenemos los métodos basados en estadística y los basados en el comportamiento, en los cuales se suelen usar técnicas de Machine Learning que permiten clasificar el tráfico extrayendo patrones de datos. Normalmente, este proceso de clasificación consta de cuatro etapas: (1) la recolección de datos, en la cual se captura el tráfico de la red haciendo uso de una herramienta construida para este fin; (2) la extracción de características, en la cual se seleccionan las características que se usarán para realizar la clasificación del tráfico; (3) el preprocesamiento de los datos, en el cual la data extraída es transformada a un formato que pueda ser procesado y (4) la ejecución del algoritmo de aprendizaje automático, en la cual se procesa la data formateada en la etapa anterior con el fin de generar los modelos de clasificación deseados [10].

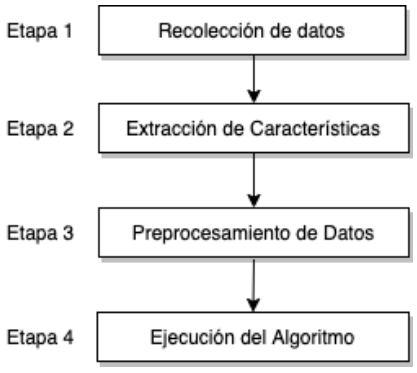

Fig. 1. Etapas del proceso de clasificación haciendo uso de técnicas de Machine Learning [10].

Para realizar la clasificación de tráfico se puede hacer uso de sistemas de aprendizaje supervisado, sistemas de aprendizaje no supervisado y sistemas híbridos de aprendizaje [8]. Si se quiere clasificar el tráfico a través de un sistema de aprendizaje no supervisado, se debe tener una data correctamente etiquetada, de lo contrario no va a ser posible realizar una caracterización de la misma y, por ende, no se va a poder realizar dicha clasificación. Este problema no se presenta en los sistemas de aprendizaje no supervisado, ya que en dichos sistemas se hace uso de técnicas de representation learning, con las cuales es posible descubrir características no visibles a partir de datos brutos para que posteriormente se pueda realizar el proceso de clasificación [11]. Es necesario tener en cuenta que en los sistemas de aprendizaje no supervisado no se toma en cuenta la data etiquetada, lo cual puede llevar a restarle precisión al modelo de clasificación

resultante. Finalmente, los sistemas de aprendizaje semi supervisados permiten suplir aquellos escenarios en los que contamos con una pequeña cantidad de datos etiquetados y una gran cantidad de datos sin etiquetar, de esta manera podemos usar ambos tipos de datos para generar un modelo de clasificación con una mejor de precisión de la que obtendríamos si solo hiciéramos uso de la data no etiquetada (sistema no supervisado).

Existen varios algoritmos de clasificación basados en machine learning, entre los algoritmos más conocidos podemos encontrar las máquinas de vector soporte (SVM) [12],[13],[14],[15],[16] los algoritmos basados en redes neuronales [17],[18],[19],[20] algoritmos basados en árboles de decisión [21],[22],[23] y los algoritmos de agrupamientos tales como K-means y esperanza-maximización (EM) [24], [25].

B. Predicción de tráfico de la red

Para Vinayakumar, Soman y Poornachandran [26], la predicción de tráfico tiene como objetivo determinar el tráfico que tendrá una red en el futuro basándose en los datos de tráfico que se han obtenido hasta un momento determinado. La predicción del tráfico de la red informática puede ser crucial para los proveedores de red y la gestión de la red informática en general. Es de gran interés en varios dominios, como aplicaciones adaptativas, control de congestión, control de admisión y asignación de ancho de banda [27]. Además, al mejorar esta tarea, se pueden crear herramientas efectivas de control de tráfico y detección de anomalías, lo que da como resultado ganancias económicas a partir de una mejor gestión de recursos [28].

Existen varias técnicas de inteligencia artificial que son usadas para predecir el tráfico de una red, entre las cuales podemos encontrar técnicas basadas en redes neuronales tales como el perceptrón multicapa (MLP) [28][29], el codificador automático apilado de aprendizaje profundo (SAE) [30], [31] y las redes neuronales recurrentes (RNN) [32], [33], [34]. Como se evidencia en [27], estas tres técnicas han demostrado ser capaces de ajustar y predecir el tráfico de una red con alta precisión, por lo cual se convierten en una herramienta de gran ayuda en el proceso del control de tráfico.

C. Seguridad de la red

Cuando hablamos de seguridad es imprescindible hablar de una gestión de los riesgos, el caso de la seguridad en las redes de datos no es la excepción y ésta debe contemplar procesos como el monitoreo de la red, seguridad de la protección, prevención de intrusiones, gestión de incidentes, configuración segura y la educación y concientización de usuarios en el uso adecuado de la red; para garantizar confidencialidad, integridad y disponibilidad [35].

Infortunadamente, a la par de la evolución a las técnicas y herramientas que protegen las redes, han evolucionado también, las técnicas y herramientas que los ciberdelincuentes emplean para atacar los sistemas. Por lo anterior, la seguridad no deja de ser una constate preocupación en la gestión del tráfico de red, lo que ha motivado la permanente investigación en el campo, dejando en evidencia el gran interés en las técnicas proporcionadas por la inteligencia artificial (IA) para confrontar los retos que la seguridad informática afronta en la actualidad. La IA brinda flexibilidad y capacidad de aprendizaje al software que ayuda a combatir los delitos cibernéticos, a partir de numerosos métodos inspirados en la naturaleza "como inteligencia computacional, redes neuronales, agentes inteligentes, sistemas inmunes artificiales, aprendizaje automático, minería de datos, reconocimiento de patrones, lógica difusa, heurística, etc., que han jugado un papel cada vez más importante en la detección y prevención de delitos cibernéticos" [36].

Adicional a la evolución de las técnicas y herramientas de los ciberdelincuentes, la evolución de las redes en sí mismas, con la aparición de nuevas tecnologías como la quinta generación de redes móviles (5G), implica nuevos retos en la protección de redes, ya que los sistemas de protección actuales pueden convertirse en insuficientes e incluso obsoletos. Por esto, las redes 5G requerirán la inclusión de nuevas tecnologías basadas en el aprendizaje automático y aprendizaje profundo [37], [38],[39]. No obstante, para esto aun hace falta mucha investigación, pues implementar sistemas de aprendizaje automático en las redes de 5G puede incluso traer problemas de seguridad como permitir una mayor facilidad para el seguimiento de usuarios y de ataques de violación de la privacidad [40].

Dado que la clasificación del tráfico es uno de los pasos fundamentales para la detección de anomalías en la red o el sistema de detección de intrusos (IDS) y desempeña un papel importante en el dominio de la seguridad de la red [41], [42]; el aprendizaje automático se ajusta a esta necesidad permitiendo la predicción de comportamientos, basándose en propiedades o características previamente conocidas, aprendidas a partir de unos datos de entrenamiento. Por otro lado, en los escenarios en los que dichas propiedades son desconocidas en los datos se puede dar uso de los métodos del aprendizaje profundo, ya que estos se centran en el descubrimiento de nuevas propiedades [42], [43]. Pero como se menciona, ML y DL se basan en datos independientemente de que deban o no ser previamente etiquetados, por esto una adecuada elección y uso de los datos son requisitos previos para lograr realizar una investigación de seguridad relevante [44]. Algunos de los conjuntos de datos de entrenamiento con fines de seguridad se pueden consultar en [43], [44], [45].

1) Monitoreo de la red

Como indican Vinayakumar et al. [44], el monitoreo de la comunicación de red incluye sensores que recopilan metadatos de la red, herramientas que brindan una sesión de red detallada y un monitoreo profundo de paquetes o análisis basado en firmas. Durante el monitoreo, los eventos de registro discretos se usan comúnmente para proporcionar indicadores de ataque, donde los eventos de sistema de archivos y de nivel de proceso se usan a menudo para identificar ciertos tipos de técnicas diversas.

Dentro de los inconvenientes encontrados en el monitoreo de la red, encontramos que estos actúan o responden con

frecuencia de manera reactiva incluso en muchos sistemas que usan métodos de ML. En respuesta a esto, se puede extender la capacidad de los IDS monitoreando los flujos de tráfico de red con técnicas de DL, ya que el aprendizaje profundo puede contribuir a la previsión proactiva, como lo propone [35].

2) Detección de intrusión

El IDS es el sistema que se encarga de monitorear o supervisar los elementos que comprenden la red, para entre otros verificar su estado y descubrir, determinar e identificar anomalías en el uso de la misma [35], [43] y [44]. Sin embargo, como se indica en [46] estos sistemas aún presentan algunos retos como mejorar la precisión de la detección, reducir las falsas alarmas y la detección de ataques desconocidos; por esto gran número de investigaciones en este campo se han centrado en desarrollar sistemas de detección que aprovechen los métodos de ML.

Con la aplicación de la inteligencia artificial, se puede lograr una clasificación, procesamiento y análisis efectivos de la información de datos de la red. A través de este proceso, algunos datos sospechosos que existen en la red podrían ser filtrados de manera efectiva y los usuarios también pueden obtener informes de inspección detallados [47].

La aplicación de la IA ofrece variadas ventajas dependiendo del método empleado en los sistemas de detección y prevención de intrusiones (IDPS) como el procesamiento de información de forma paralela en las redes neuronales artificiales, adaptación al entorno y las preferencias de usuario en los agentes inteligentes, auto adaptabilidad y auto organización en los sistemas inmunes artificiales, optimización con algoritmos genéticos y la interoperabilidad en la lógica difusa [36].

3) Métodos y algoritmos de IA para seguridad

Dentro de la literatura académica existe variada cantidad de investigaciones y desarrollo de sistemas basados en métodos principalmente de machine learning y deep learning para la detección de intrusos como parte del sistema de seguridad de la red.

Entre los algoritmos de ML considerados como poco profundos más comúnmente utilizados para la seguridad en redes encontramos: máquinas de vectores de soporte, k vecinos más próximos, árboles de decisión, Naïve Bayes, las redes neuronales artificiales y regresión logística, aunque podemos encontrar otros como clustering, aprendizaje en conjunto híbrido, entre otros. Así mismo, los algoritmos y métodos del aprendizaje profundo, más empleados son: red de creencias profundas (Deep Belief Network; DBN), redes neuronales recurrentes (Recurrent Neural Networks; RNN), redes neuronales convolucionales (Covolutional Neural Networks; CNN), red generativa antagónica (Generative Adversarial Network; GAN), entre otros [36], [43], [44], [45], [46]. Encontramos que en [36] se resume las ventajas de la utilización de algunos de los algoritmos mencionados en los sistemas de detección y prevención de intrusiones, [43] y [44] describen el conjunto de datos utilizados para la ciberseguridad y algunos de los algoritmos de ML y DL más comunes. En [46] se realiza un estudio en el que se compara los métodos superficiales o poco profundos sobre los métodos profundos o de deep learning en cinco características: tiempo de ejecución, número de parámetros, representación de características, capacidad de aprendizaje e interoperabilidad; mostrando mayores ventajas como es el caso del rendimiento en los algoritmos de DL.

Algunos ejemplos de trabajos que han implementado los anteriores algoritmos mencionados u otros relacionados se describen brevemente a continuación:

[41] [48] hacen uso de técnicas de representación de imágenes para detección de malware; realizan la clasificación de tráfico de malware basado en las redes neuronales convolucionales con el enfoque de aprendizaje por representación. Este enfoque toma los datos del tráfico sin procesar y son representados en imágenes las cuales permiten clasificar el tipo de tráfico correspondiente, ya que se logró determinar que, para un tipo de tráfico particular, las imágenes generadas son muy similares entre sí; permitiendo distinguir el tráfico de malware del tráfico de otras aplicaciones de usuario (correo, video juegos, skype, etc) e incluso distinguir el tráfico entre diferentes tipos de malware.

Los Ataques DoS y DDoS llevan décadas preocupando a los gestores de seguridad en los sistemas, por su capacidad de dejar indisponibles los servicios, para estas preocupaciones en [49] y [50] nuevamente se propone la utilización de métodos de ML para la detección de estos ataques. En [49] los autores estudian la construcción de un modelo de ML basado en el análisis de regresión múltiple para la detección de ataques DDoS, mientras que en [50] se presenta un sistema basado en ML para la detección de DoS, en el que luego de explorar con varios algoritmos en la fase de clasificación determinan que random forest (RF) es quien muestra una mayor precisión. Se destaca que según los resultados obtenidos en [50] se obtiene una tasa de detección del 96%, con alta precisión y una tasa baja de falsas alarmas.

Algunas de las herramientas convencionales de detección de intrusiones en la red suelen presentar niveles considerables de falsas alarmas. Debido a esto, los autores en [51] deciden diseñar un enfoque para la detección de tráfico de red malicioso basado en un clasificador de red neuronal artificial que permite identificar patrones de shellcode, obteniendo mejores resultados de los métodos de detección que se basan en firmas, logrando una tasa de falsos positivos inferior al 2% durante pruebas con más de 400.000 muestras de contenido de archivos de tráfico de red benigno.

D. Optimización de ruteo

La optimización de enrutamiento permite un uso con mayor eficacia de los recursos de red y mejorar la calidad del servicio [52]. Por esto, suele ser de gran interés para los proveedores de internet, pero no solo se aplica a las redes de datos tradicionales sino también a otros tipos de red como la creciente Internet de las cosas IoT [53].

Existe algoritmos de optimización como la colonia de hormigas que puede ser muy efectivos en la optimización del enrutamiento y por esto varios protocolos de enrutamiento se basan en este tipo de algoritmos [54]. Pero estos algoritmos no son adecuados para optimizaciones en tiempo real debido a

que al ser un método heurístico implica un elevado costo computacional. Sin embargo, gracias a las virtudes de ML se puede optimizar el enrutamiento en tiempo real, principalmente con el aprendizaje con refuerzo [52].

Algunas soluciones para la optimización del tráfico sugieren optimización únicamente de los flujos que sean considerados críticos, lo que implica nuevamente un proceso de clasificación el cual puede ser de alta carga de computacional, por lo cual [55] proponen la CFR-RL (Critical Flow Rerouting-Reinforcement Learning). CFR_RL utiliza una red neuronal la cual se entrena mediante el aprendizaje automático por refuerzo con políticas de selección del tráfico critico en los diferentes tipos de tráfico y logra un rendimiento óptimo al redirigir entre el 10% y el 21,3% del tráfico total.

E. Algoritmos aplicados al control de tráfico

De acuerdo con las fuentes consultadas en el presente artículo, en la Tabla I se realiza un consolidado de los diferentes algoritmos y métodos usados para el control de tráfico.

TABLA I
ALGORITMOS APLICADOS AL CONTROL DE TRÁFICO.

Área de Aplicación	Algoritmos/Métodos	Documento
Clasificación del tráfico	Máquinas de vectores de soporte. (SVM)	Sun, Chen y Su. [12] Cao et al. [13] Dong. [14] Punitha y Mala. [15] Saber, Fergani y Abbas. [16]
	Redes neuronales	Ang et al. [17] Li et al. [18] Moreira et al. [19] Mengmeng, Xiangzhan y Likun. [20]
	Árboles de decisión	Li, et al. [21] Tong, Qu y Prasanna. [22] Nair y Sajeev. [23]
	K-means y esperanza-maximización	Singh. [24] Liu et al. [25]
Predicción del tráfico	Perceptrón multicapa (MLP): red neuronal	Prado, Salem y Santos. [28] Nikravesh et al. [29]
	Codificador automático apilado (SAE): red neuronal	Jin et al. [30] Li et al. [31]
	Redes neuronales recurrentes (RNN)	Jaffry y Hasan. [32] Ramakrishnan y Soni. [33] Andreoletti et al. [34]
Seguridad de la red	Red de creencias profundas (DBN)	Vinayakumar et al. [44] Lima et al. [46]
	Redes neuronales recurrentes (RNN)	Vinayakumar et al. [44] Lima et al. [46]
	Red generativa antagónica (GAN)	Lima et al. [46]
	Redes neuronales artificiales	Dilek, Çakır y Aydın. [36] Wang. [43] Sambangi y Gondi. [45] Lima et al. [46] Ly y Yao [51]
	Clustering	Wang. [43] Lima et al. [46]
	Árboles de decisión	Wang. [43] Vinayakumar et al. [44] Sambangi y Gondi. [45] Lima et al. [46]
	Naive Bayes	Wang. [43] Sambangi y Gondi. [45] Lima et al. [46]
	Máquina de vectores de soporte (SVM)	Wang. [43] Vinayakumar et al. [44] Sambangi y Gondi. [45] Lima et al. [46]
	K-vecinos más cercanos	Vinayakumar et al. [44] Sambangi y Gondi. [45] Lima et al. [46]
	Autoencoder	Lima et al. [46]
	Máquina de Boltzmann restringida (RBM)	Lima et al. [46]
	Red neutral profundas (DNN)	Lima et al. [46]
	Redes bayesianas	Wang. [43]
	Redes neuronales convolucionales (CNN)	Farah et al. [41] Vinayakumar et al. [44] Zhang et al. [48]
	Algoritmos de regresión múltiple	Saleem et al. [49]
	Random forest	Sanchez et al. [50]
	Modelos ocultos de Markov	Wang. [43]
Optimización de ruteo	Redes neuronales	Zhang et al. [55]

V. CONCLUSIONES

La inteligencia artificial ha contribuido de manera significativa en el mejoramiento de las técnicas usadas en el proceso de control de tráfico de redes de datos, en aspectos tales como la clasificación y predicción del tráfico, la seguridad en las redes y la optimización del ruteo. Esto debido a las bondades proporcionadas principalmente por los algoritmos de aprendizaje automático y aprendizaje profundo (velocidad, precisión, autoaprendizaje, entre otros), los cuales permiten a los sistemas de gestión de redes trabajar de una manera más autónoma, proactiva y eficiente, haciendo que las redes funcionen de manera óptima, mejorando con ello la calidad del servicio ofrecido a los usuarios.

Por otro lado, se puede apreciar una clara tendencia hacía la implementación de algoritmos de aprendizaje profundo sobre algoritmos de aprendizaje automático tradicional, debido a que con el aprendizaje profundo se requiere menos preparación de los datos obtenidos de la red para poder generar los modelos de clasificación necesarios en los procesos de optimización de ruteo e identificación de intrusiones. Entre los algoritmos más usados se destacan los basados en redes neuronales, debido a su gran capacidad en el procesamiento de información de forma paralela y su alta precisión.

De igual manera, se evidencia que, independientemente de los algoritmos aplicados, la fase de clasificación debe ser bien estructurada y planificada de acuerdo con el problema específico, ya que de esta fase depende la eficiencia en la ejecución del algoritmo y/o la precisión de sus resultados.

Para finalizar, se debe indicar que a pesar de que se han optimizado los mecanismos de procesamiento de la información y del constante avance en el aumento de los porcentajes de precisión y en la reducción de las tazas de error, sigue existiendo un pequeño porcentaje de error que en algunos escenarios hace necesaria la supervisión humana.

Referencias

[1] Y. D. Chang and D. H. Su, Study of Traffic Control and Congestion Control in Broadband ISDN, Gaithersburg: National Institute of Standards and Technology, 1992. Available: https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir5000.pdf

[2] S. Poretsky, J. Perser, S. Erramilli and S. Khurana, "Terminology for Benchmarking Network-layer Traffic Control Mechanisms", Internet Engineering Task Force, RFC 4689, Oct. 2006. Available: https://tools.ietf.org/html/rfc4689

[3] A. Farrel, J. P. Vasseur and J. Ash, "A Path Computation Element (PCE)-Based Architecture", Internet Engineering Task Force, RFC 4655, Aug. 2006. Available: https://tools.ietf.org/html/rfc4655

[4] D. Russell, The Principles of Computer Networking, UK: Cambridge University Press, 1989. Available: https://books.google.com.co/books?id=ReDwVJGlxLsC

[5] N. Brownlee, C. Mills and G. Ruth, "Traffic Flow Measurement: Architecture", Internet Engineering Task Force, RFC 2722, Oct. 1999. Available: https://tools.ietf.org/html/rfc2722

[6] M. Noormohammadpour and C. S. Raghavendra, "Datacenter Traffic Control: Understanding Techniques and Tradeoffs," in *IEEE Commun. Surveys & Tutorials*, vol. 20, no. 2, pp. 1492-1525, 2018, doi: 10.1109/COMST.2017.2782753

[7] W. Li and A. W. Moore, "A Machine Learning Approach for Efficient Traffic Classification," *2007 15th International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*, 2007, pp. 310-317, doi: 10.1109/MASCOTS.2007.2.

[8] P. Amaral, J. Dinis, P. Pinto, L. Bernardo, J. Tavares and H. S. Mamede, "Machine Learning in Software Defined Networks: Data collection and traffic classification," *2016 IEEE 24th International Conference on Network Protocols (ICNP)*, 2016, pp. 1-5, do: 10.1109/ICNP.2016.7785327.

[9] S. Blake, D. Black, M. Carlson and E. Davies, "An Architecture for Differentiated Services", Internet Engineering Task Force, RFC 2475, Dec. 1998. Available: https://datatracker.ietf.org/doc/html/rfc2475.

[10] M. Shafiq, X. Yu, A. A. Laghari, L. Yao, N. K. Karn and F. Abdessamia, "Network Traffic Classification techniques and comparative analysis using Machine Learning algorithms," *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*, 2016, pp. 2451-2455, doi: 10.1109/CompComm.2016.7925139.

[11] Y. Bengio, A. Courville and P. Vincent, "Representation Learning: A Review and New Perspectives," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798-1828, 2013, doi: 10.1109/TPAMI.2013.50.

[12] G. Sun, T. Chen and Y. Su, "Internet Traffic Classification Based on Incremental Support Vector Machines", in *Mobile Netw Appl*, vol. 23, no. 4, pp. 789–796, 2018, doi: 10.1007/s11036-018-0999-x.

[13] J. Cao, D. Wang, Z. Qu, H. Sun, B. Li and C. Chen, "An Improved Network Traffic Classification Model Based on a Support Vector Machine", in *Symmetry*, vol. 12, no. 2, pp. 301, 2020, doi: 10.3390/sym12020301.

[14] S. Dong, "Multi class SVM algorithm with active learning for network traffic classification," in *Expert Systems with Applications*, vol. 176, no. 1, 2021, doi: 10.1016/j.eswa.2021.114885.

[15] V. Punitha and C. Mala, "Traffic classification in server farm using supervised learning techniques", in *Neural Comput & Applic*, vol. 33, no. 4, pp. 1279–1296, 2020, doi: 10.1007/s00521-020-05030-2.

[16] A. Saber, B. Fergani and M. Abbas, "Encrypted Traffic Classification: Combining Over-and Under-Sampling through a PCA-SVM," *2018 3rd International Conference on Pattern Analysis and Intelligent Systems (PAIS)*, 2018, pp. 1-5, doi: 10.1109/PAIS.2018.8598480.

[17] M. Ang, E. Valla, N. Neggatu and A. Moore, "Network Traffic Classification via Neural Networks," Universidad de Cambridge, Cambridge, Reino Unido, Rep. Téc. TR-912, 2017.

[18] R. Li, X. Xiao, S. Ni, H. Zheng and S. Xia, "Byte Segment Neural Network for Network Traffic Classification," *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*, 2018, pp. 1-10, doi: 10.1109/IWQoS.2018.8624128.

[19] R. Moreira, L. F. Rodrigues, P. F. Rosa, R. L. Aguiar and F. d. O. Silva, "Packet Vision: a convolutional neural network approach for network traffic classification," *2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, 2020, pp. 256-263, doi: 10.1109/SIBGRAPI51738.2020.00042.

[20] G. Mengmeng, Y. Xiangzhan and L. Likun, "Robot Communication: Network Traffic Classification Based on Deep Neural Network," in *Frontiers in Neurorobotics*, vol. 15, no. 1, 2021, doi: 10.3389/fnbot.2021.648374.

[21] N. Li et al., "Network Traffic Classification and Control Technology Based on Decision Tree," I*nternational Conference on Applications and Techniques in Cyber Intelligence ATCI*, 2019, pp. 1701-1705, doing: 10.1007/978-3-030-25128-4_217.

[22] D. Tong, Y. R. Qu and V. K. Prasanna, "Accelerating Decision Tree Based Traffic Classification on FPGA and Multicore Platforms," in *IEEE Transactions on Parallel and Distributed Systems*, vol. 28, no. 11, pp. 3046-3059, 2017, doi: 10.1109/TPDS.2017.2714661.

[23] L. M. Nair and G. P. Sajeev, "Internet Traffic Classification by Aggregating Correlated Decision Tree Classifier," *2015 Seventh International Conference on Computational Intelligence, Modelling and Simulation (CIMSim)*, 2015, pp. 135-140, doi: 10.1109/CIMSim.2015.15.

[24] H. Singh, "Performance Analysis of Unsupervised Machine Learning Techniques for Network Traffic Classification," *2015 Fifth International Conference on Advanced Computing & Communication Technologies*, 2015, pp. 401-404, doi: 10.1109/ACCT.2015.54.

[25] S. Liu, J. Hu, S. Hao *and T. Song, "Improved EM method for internet traffic classification," 2016 8th International Conference on Knowledge and Smart Technology (KST)*, 2016, pp. 13-17, doi: 10.1109/KST.2016.7440488.

[26] R. Vinayakumar, K. P. Soman and P. Poornachandran, "Applying deep learning approaches for network traffic prediction," *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2017, pp. 2353-2358, doi: 10.1109/ICACCI.2017.8126198.

[27] T. Prado, J. Salem and A. Santos, "Computer network traffic prediction: a comparison between traditional and deep learning neural networks," in *International Journal of Big Data Intelligence*, vol. 3, no. 1, pp. 28-37, 2016, doi: 10.1504/IJBDI.2016.073903.

[28] T. Prado, J. Salem and A. Santos, "Multilayer Perceptron and Stacked Autoencoder for Internet Traffic Prediction," *International Conference on Network and Parallel Computing*, 2014, pp. 61-71, doi: 10.1007/978-3-662-44917-2_6.

[29] A. Y. Nikravesh, S. A. Ajila, C. Lung and W. Ding, "Mobile Network Traffic Prediction Using MLP, MLPWD, and SVM," *2016 IEEE International Congress on Big Data (BigData Congress)*, 2016, pp. 402-409, doi: 10.1109/BigDataCongress.2016.63.

[30] Y. Jin, W. Xu, P. Wang and J. Yan, "SAE Network: A Deep Learning Method for Traffic Flow Prediction," *2018 5th International Conference on Information, Cybernetics, and Computational Social Systems (ICCSS)*, 2018, pp. 241-246, doi: 10.1109/ICCSS.2018.8572451.

[31] P. Li, Z. Chen, L. T. Yang, J. Gao, Q. Zhang and M. J. Deen, "An Improved Stacked Auto-Encoder for Network Traffic Flow Classification," in *IEEE Network*, vol. 32, no. 6, pp. 22-27, 2018, doi: 10.1109/MNET.2018.1800078.

[32] S. Jaffry and S. F. Hasan, "Cellular Traffic Prediction using Recurrent Neural Networks," *2020 IEEE 5th International Symposium on Telecommunication Technologies (ISTT)*, 2020, pp. 94-98, doi: 10.1109/ISTT50966.2020.9279373.

[33] N. Ramakrishnan and T. Soni, "Network Traffic Prediction Using Recurrent Neural Networks," *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2018, pp. 187-193, doi: 10.1109/ICMLA.2018.00035.

[34] D. Andreoletti, S. Troia, F. Musumeci, S. Giordano, G. Maier and M. Tornatore, "Network Traffic Prediction based on Diffusion Convolutional Recurrent Neural Networks," *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2019, pp. 246-251, doi: 10.1109/INFCOMW.2019.8845132.

[35] G. Nguyen, S. Dlugolinsky, V. Tran and Á. López, "Deep Learning for Proactive Network Monitoring and Security Protection," in *IEEE Access*, vol. 8, pp. 19696-19716, 2020, doi: 10.1109/ACCESS.2020.2968718.

[36] S. Dilek, H. Çakır and M. Aydın, "Applications of artificial intelligence techniques to combating cyber crimes: a review," in *IJAIA*, vol. 6, no. 1, pp. 21-39, 2015, doi: 10.5121/ijaia.2015.6102.

[37] K. Saleem, G. M. Alabduljabbar, N. Alrowais, J. Al-Muhtadi, M. Imran and J. J. P. C. Rodrigues, "Bio-Inspired Network Security for 5G-

Enabled IoT Applications," in *IEEE Access*, vol. 8, pp. 229152-229160, 2020, doi: 10.1109/ACCESS.2020.3046325.
[38] J. Sanchez-Gomez et al., "Integrating LPWAN Technologies in the 5G Ecosystem: A Survey on Security Challenges and Solutions," in *IEEE Access*, vol. 8, pp. 216437-216460, 2020, doi: 10.1109/ACCESS.2020.3041057.
[39] A. Ly and Y. -D. Yao, "A Review of Deep Learning in 5G Research: Channel Coding, Massive MIMO, Multiple Access, Resource Allocation, and Network Security," in *IEEE OJ-COMS*, vol. 2, pp. 396-408, 2021, doi: 10.1109/OJCOMS.2021.3058353.
[40] J. Suomalainen, A. Juhola, S. Shahabuddin, A. Mämmelä and I. Ahmad, "Machine Learning Threatens 5G Security," in *IEEE Access*, vol. 8, pp. 190822-190842, 2020, doi: 10.1109/ACCESS.2020.3031966
[41] W. Wang, M. Zhu, X. Zeng, X. Ye and Y. Sheng, "Malware traffic classification using convolutional neural network for representation learning," *ICOIN*, pp. 712-717, 2017, doi: 10.1109/ICOIN.2017.7899588,
[42] Y. Zeng, H. Gu, W. Wei and Y. Guo, "Deep-Full-Range: A Deep Learning Based Network Encrypted Traffic Classification and Intrusion Detection Framework," in *IEEE Access*, vol. 7, pp. 45182-45190, 2019, doi: 10.1109/ACCESS.2019.2908225.
[43] A. L. Buczak and E. Guven, "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection," in *IEEE Commun. Surv*. Tutor., vol. 18, no. 2, pp. 1153-1176, 2016, doi: 10.1109/COMST.2015.2494502.
[44] Y. Xin et al., "Machine Learning and Deep Learning Methods for Cybersecurity," in *IEEE Access*, vol. 6, pp. 35365-35381, 2018, doi: 10.1109/ACCESS.2018.2836950.
[45] N. Farah, A. Rahman, A. Khan, M. Rafni, M. Shah and D. Farid, "Application of Machine Learning Approaches in Intrusion Detection System: A Survey," in *IJAIA*, Vol. 4, No. 3, pp. 9-18, 2015, doi: 10.14569/IJARAI.2015.040302.
[46] H. Liu and B. Lang, "Machine Learning and Deep Learning Methods for Intrusion Detection Systems: A Survey," in *Applied Sciences*, vol. 9, no. 20, p. 4396, 2019, doi: 10.3390/app9204396.
[47] Y. Wang, "Research on Application of Artificial Intelligence in Computer Network Technology under the Background of Big Data," *2020 Journal of Physics: Conference Series 1607 012093*, 2020. 10.1088/1742-6596/1607/1/012093.
[48] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran and S. Venkatraman, "Robust Intelligent Malware Detection Using Deep Learning," in *IEEE Access*, vol. 7, pp. 46717-46738, 2019, doi: 10.1109/ACCESS.2019.2906934.
[49] S. Sambangi and L. Gondi, "A Machine Learning Approach for DDoS (Distributed Denial of Service) Attack Detection Using Multiple Linear Regression," in *Proceedings*, vol. 63, no. 1, p. 51, 2020, doi: 10.3390/proceedings2020063051.
[50] F. S. de Lima, F. A. Silveira, A. Brito, G. Vargas-Solar and L. F. Silveira, "Smart Detection: An Online Approach for DoS/DDoS Attack Detection Using Machine Learning," *Secur. Commun. Netw.*, vol. 2019, pp. 1-15, 2019, doi: 10.1155/2019/1574749.
[51] A. Shenfield, D. Day and A. Ayesh, "Intelligent intrusion detection systems using artificial neural networks," in *ICT Express*, vol. 4, pp. 95-99, 2018, doi: 10.1016/j.icte.2018.04.003.
[52] I. Ampratwum, "An Intelligent Traffic Classification based optimized routing in SDN-IoT: A Machine Learning Approach," (M.S. thesis), Faculty of Engineering, University of Ottawa, Ottawa. 2020. https://ruor.uottawa.ca/bitstream/10393/40155/1/Ampratwum_Isaac_2020_thesis.pdf
[53] S. Xu, X. Wang, G. Yang, J. Ren and S. Wang, "Routing optimization for cloud services in SDN-based Internet of Things with TCAM capacity constraint," in *JCN*, vol. 22, no. 2, pp. 145-158, 2020, doi: 10.1109/JCN.2020.000006.
[54] H. Zhang, X. Wang, P. Memarmoshrefi and D. Hogrefe, "A Survey of Ant Colony Optimization Based Routing Protocols for Mobile Ad Hoc Networks," in *IEEE Access*, vol. 5, pp. 24139-24161, 2017, doi: 10.1109/ACCESS.2017.2762472.
[55] J. Zhang, M. Ye, Z. Guo, C. -Y. Yen and H. J. Chao, "CFR-RL: Traffic Engineering with Reinforcement Learning in SDN," *IEEE JSAC*, vol. 38, no. 10, pp. 2249-2259, 2020, doi: 10.1109/JSAC.2020.3000371.

Daruin Arley León Salas. Tecnólogo en sistematización de datos de la Universidad Distrital Francisco José De Caldas, Bogotá, Colombia, 2018. Candidato a Ingeniero Telemático de la Universidad Distrital Francisco José De Caldas. Desarrollador de software senior. Áreas de interés: Ingeniería de Software, redes informáticas e Inteligencia Artificial.
ORCID: https://orcid.org/0000-0002-5225-2037.

James Gustavo Martínez Cuenca. Tecnólogo en sistematización de datos de la Universidad Distrital Francisco José De Caldas, Bogotá, Colombia, 2018. Tecnólogo en electrónica de la Corporación Universitaria Minuto de Dios, Bogotá, Colombia, 2015. Candidato a Ingeniero Telemático de la Universidad Distrital Francisco José De Caldas. Desarrollador de software senior. Áreas de interés: Ingeniería de Software, redes informáticas e Inteligencia Artificial.
ORCID: https://orcid.org/0000-0003-2145-7096.

Ismael Antonio Ardila Sánchez. Magister en gestión de proyectos del Instituto Europeo de Posgrado - IEP, Madrid, España, 2017. Ingeniero de Sistemas de la Universidad Antonio Nariño, Bogotá, Colombia, 2000. Docente de la Universidad Distrital Francisco José De Caldas. Áreas de interés: Redes de datos, gestión de proyectos, gestión de redes e ingeniería de software.
ORCID: https://orcid.org/0000-0001-7274-9916.

Darin Jairo Mosquera Palacios. Magister en teleinformática de la Universidad Distrital Francisco José De Caldas, Bogotá, Colombia, 2010. Ingeniero de Sistemas de la Universidad Autónoma De Colombia, Bogotá, Colombia, 1996. Director del grupo de investigación ORION de la Universidad Distrital Francisco José De Caldas. Docente de la Universidad Distrital Francisco José De Caldas. Áreas de interés: Medios de Transmisión, redes Inteligentes, redes Corporativas, gestión de Redes y seguridad en redes.
ORCID: https://orcid.org/0000-0002-4526-2683