Universidad Nacional Federico Villarreal

Vicerrectorado de INVESTIGACIÓN

ESCUELA UNIVERSITARIA DE POSGRADO

"ANÁLISIS EXPLORATORIO DE ATAQUES INFORMÁTICOS APLICANDO HERRAMIENTAS DE MINERIA DE DATOS, PARA LA GESTIÓN DE LA SEGURIDAD DE REDES INALÁMBRICAS EN UNIVERSIDADES DE AREQUIPA"

TESIS PARA OPTAR EL GRADO ACADÉMICO DE:
DOCTOR EN INGENIERÍA DE SISTEMAS

AUTOR:

ROSAS PAREDES KARINA

ASESOR:

DRA. HUAMAN FERNANDEZ JACKELINE ROXANA

JURADO:
DRA. TAFUR ANZUALDO VICENTA IRENE
DR. BOLIVAR JIMENEZ JOSE LUIS
DR. ROMERO ECHEVARRIA LUIS MIGUEL

LIMA – PERÚ

2020

DEDICATORIA

A Dios por ser mi soporte y guía y a mi familia, Carlos, Karla, Emilia y Diogo; a mis padres Antonio y Pina; a mis hermanos Cecilia y José Antonio

los motivos para para concretar

mis metas profesionales.

AGRADECIMIENTOS

Quiero expresar mi más sincero agradecimiento a todas las personas que apoyaron este proyecto y que de alguna manera colaboraron conmigo en la realización de la tesis, a la Dra. Jackeline Huamán, asesora de esta investigación, por las sugerencias, motivación y apoyo constante.

Agradezco la colaboración y soporte de las autoridades y personal administrativo de la UCSM, en especial a Max Rondón, Gonzalo Dávila, Diego Valencia, Toshiro Nagata y Royler Correa.

Y de manera especial agradecer la hospitalidad a Alexander, Mildred, Liam y Daniel. A todos ellos, muchas gracias.

ÍNDICE

Pág.

LISTA DE TABLAS

Pág.

LISTA DE GRAFICOS

Pág.

RESUMEN

El presente trabajo de tesis, realiza la captura de tráfico de redes inalámbricas y cableadas que atraviesan la infraestructura física del campus universitario, el mismo que es analizado de forma exploratoria utilizando herramientas de minería de datos dada la cantidad de registros obtenidos en periodos de tiempo variables que nos permite discriminar entre los diversos ataques, niveles de agresión, duración, procedencia, destino, acciones realizadas para contrarrestar dichos ataques en la misma red interna a través de modelos de clasificación, asociación y segmentación utilizando algoritmos adecuados y basados en la metodología CRISP-DM que describe de forma normalizada, el ciclo de vida de un proyecto de análisis de datos, para predecir futuros eventos y evaluar la relación de gestión de seguridad de las redes en su conjunto con los controles aplicados, sugiriendo a los administradores de TI la implementación de políticas adecuadas y de ésta forma se pueda disminuir las amenazas a las infraestructuras de redes ante los ataques que diariamente se reciben; de la misma forma se fomenta el uso de las herramientas tecnológicas necesarias para el buen funcionamiento de los sistemas académico-administrativos de una entidad de educación superior. Para ello se caracterizó los ataques informáticos utilizando la prueba de Chi cuadrado considerando la "categoría de amenaza" como variable dependiente y las demás características de los ataques informáticos como variables independientes, resultando spyware y code-execution los de mayor incidencia; los algoritmos con el mejor nivel de confiabilidad para la predicción de ataques informáticos REPTree y CHAID, adoptándolos como predictores a la categoría de amenaza; la encuesta aplicada a los administradores de seguridad del departamento de TI, concluye que las universidades pocas veces realizan auditorias y actividades de verificación de los sistemas para minimizar la interrupción de los procesos, por lo que se puede afirmar que el trabajo de investigación logró el objetivo de determinar cómo el análisis exploratorio está estrechamente relacionado con la gestión de seguridad de las redes inalámbricas en universidades locales, permitiendo determinar qué controles de seguridad informática se aplican correctamente o cuales tienen que modificarse para alcanzar un nivel de servicio adecuado; se termina con las discusiones, conclusiones y recomendaciones de la investigación.

Palabras Clave: Ataques, minería de datos, vulnerabilidad, gestión, seguridad, controles.

ABSTRACT

This thesis work, captures the traffic of wireless and wired networks that cross the physical infrastructure of the university campus, which is explored in an exploratory way using data mining tools given the number of records obtained in variable periods of time which allows us to discriminate between the various attacks, levels of aggression, duration, origin, destination, actions taken to counteract these attacks in the same internal network through classification, association and segmentation models using appropriate algorithms and based on the CRISP methodology- DM that describes in a standardized way, the life cycle of a data analysis project, to predict future events and evaluate the relationship of the security management of the networks as a whole with the controls applied, suggesting to IT administrators the implementation of appropriate policies and this form can reduce the network infrastructure threats to the attacks that are received daily, in the same way the use of the technological tools necessary for the proper functioning of the academic-administrative systems of a higher education entity is encouraged. For this, the computer attacks were characterized using the Chi-square test considering the threat category as a dependent variable and the other characteristics of the computer attacks as independent variables, resulting in spyware and code-execution being the ones with the highest incidence; the algorithms with the best level of reliability for predicting REPTree and CHAID computer attacks, adopting them as predictors of the threat category; The survey applied to the security administrators of the IT department, concludes that universities rarely perform audits and systems verification activities to minimize the interruption of processes, so it can be stated that the research work achieved the objective to determine how the exploratory analysis is closely related to the security management of wireless networks in local universities, allowing to determine which computer security controls are applied correctly or which have to be modified to reach an adequate level of service, the discussions are ended , conclusions and recommendations of the investigation.

Keywords: Attacks, data mining, vulnerability, management, security, controls.

I. INTRODUCCIÓN

La seguridad informática hoy en día es un tema muy crítico para el normal desenvolvimiento de las empresas e instituciones dedicadas a cualquier rubro ya que se pretende proteger la infraestructura computacional y la información contenida, almacenada y que se transmite por medio de ellas. Las universidades no ajenas al uso de redes telemáticas deben proveer servicios seguros para proteger a los servidores, computadoras, dispositivos móviles, redes y datos, de los cientos de ataques informáticos que a diario se perciben; de tal forma que los sistemas de información estén siempre disponibles.

Las universidades debido al avance tecnológico, han mejorado sus sistemas de protección a través de firewalls, sistemas de detección de intrusos, sistemas de prevención de intrusos y antivirus con la finalidad de proteger al activo más importante de toda organización como es la "información" de tal forma que la misma esté protegida para garantizar la confiabilidad, integridad y disponibilidad; y de esta forma la comunidad universitaria realice sus transacciones con total confianza.

Las universidades y cualquier entorno en la actualidad se han convertido en blanco atractivo para los ciberdelincuentes, que atacan sin medir las consecuencias muchas veces desde dentro del campus universitario, haciendo la detección de los mismos difíciles y poco creíble que interna a la organización se perpetren los ataques, desde los más simples hasta los más complejos y, lastimosamente algunas políticas de seguridad estandarizadas no se aplican por falta de un buen análisis previo de los ataques más frecuentes y los que producen más daño, poniendo en riesgo los sistemas de información, así como su clasificación de proveniencia, por nivel de riesgo, por tipo de protocolo, por duración en tiempos, por frecuencias y nivel de agresión; lo que no permite realizar una correcta gestión de seguridad informática.

El presente trabajo plantea capturar los diferentes tipos de tráfico, detectar variaciones en los ataques, detectar nuevos orígenes y sus cambios en los perfiles en las capas del 1 al 7 del Modelo OSI, para su posterior análisis exploratorio para predecir las futuras amenazas y definir la relación que tengan en la gestión de seguridad informática con los controles aplicados según estándares ISO, a la vez que

permitan a los administradores de la red desarrollar una mejor gestión de la seguridad informática en base a procedimientos estandarizados, constituyendo un aporte para las comunicaciones en general y una propuesta de implementación de controles de seguridad de los dominios de Cifrado, Seguridad operativa, Seguridad telemática y Gestión de Incidentes que plantea la ISO-27002(2013) y de esta forma se tornen menos vulnerables las redes cableadas e inalámbricas y los sistemas universitarios se encuentren siempre disponibles.

La investigación está dividida en seis capítulos:

En el primer capítulo se presenta el planteamiento del problema, donde se listan los antecedentes similares a la presente investigación, se define el problema y los objetivos del trabajo, se señala la necesidad de realizar el presente trabajo justificando su desarrollo y las limitaciones del mismo, asi mismo de plantean los objetivos a alcanzar y la hipótesis que se pretenden demostrar a través de métodos estadísticos adecuados.

En el segundo capítulo se presenta el marco teórico, en donde se definen aspectos fundamentales para entender la temática, conceptos, modelos, definiciones que soportan el trabajo de investigación, presento conceptos vinculados a seguridad informática, herramientas de captura de datos en redes inalámbricas y cableadas, modelos y técnicas para realizar minería de datos, metodología para contemplar el proceso de análisis de datos con minería de datos, controles de la norma ISO 27002 (2013) y terminología utilizada de gran relevancia para el estudio.

En el tercer capítulo se detalla el método de investigación, se presentan el tipo, el nivel y diseño de la investigación, las hipótesis, variables y su operacionalización, se determina la población y la muestra necesaria para el estudio y finalmente se describen los instrumentos y procedimientos que se siguen en el desarrollo de la tesis.

En el cuarto capítulo se presentan los resultados obtenidos a través de la aplicación de herramientas de minería de datos, que permitieron clasificar, asociar y segmentar los datos para predecir futuros ataques informáticos.

En el quinto capítulo se realiza la interpretación de los resultados, finalizando con la contrastación de la hipótesis.

En el sexto capítulo, se presentan las discusiones, conclusiones y recomendaciones de la investigación, determinando que el análisis exploratorio de ataques informáticos a través de la minería de datos, permite predecir futuras incidencias en las redes universitarias.

Finalmente, se presentan la bibliografía y referencias utilizados en el desarrollo de la tesis, los anexos que contienen las fichas de registro del reporte del firewall y la ficha de recolección de datos que se utilizaron en la investigación.

1.1. PLANTEAMIENTO DEL PROBLEMA

En la actualidad, las redes de datos cableadas y sobre todo inalámbricas se vienen implementando de forma masiva, transformándose en un segmento de aplicación considerable, estas redes son cada vez más utilizadas en aplicaciones a través de dispositivos móviles y su uso se debe a la flexibilidad, escalabilidad y facilidad de implementación.

Exponencialmente, las redes inalámbricas se están introduciendo en el mercado de consumo gracias a precios populares y a un conjunto de posibilidades como la satisfacción de sus usuarios en el tema de la movilidad e integración de los servicios y porque ya estamos viviendo la tecnología M2M (machine to machine) o del IoT (Internet del Todo), así mismo su implementación va de menos a más en todo tipo de ambiente, sea empresarial, gobierno y justamente académico como es nuestro caso.

Pero, estas ventajas, traen asociadas el gran inconveniente en cuanto a sus vulnerabilidades en temas de seguridad y cada día el interés del administrador de las redes es implementar mecanismos de seguridad que muchas veces no son eficientes, pues las habilidades de los atacantes han superado los límites y conocimientos del administrador de la red.

1.2. DESCRIPCIÓN DEL PROBLEMA

Las intrusiones que sufren diariamente nuestras redes de datos universitarias, son intentos que comprometen la confidencialidad, integridad, disponibilidad (CID) y que burlan los mecanismos de seguridad implementados en las redes.

Los IDS (Sistemas de detección de intrusos) existentes si bien detectan los cientos de ataques diarios, muchas veces no pueden gestionar adecuadamente éstos, dado que no se realiza un buen análisis de contenidos, tipos, procedencia, destino, protocolos, etc; para aplicar los controles adecuados en bases al nivel de riesgos que puedan ocasionar; los protocolos soportados por los firewalls, son generalmente a nivel de la capa de aplicación y la capa de red y no están correctamente configurados, resultando ineficientes la implementación de políticas de seguridad y como consecuencia, los sistemas se tornan vulnerables a las intrusiones en el resto de capas.

La vulnerabilidad de las redes de datos está presente, principalmente, en los campus universitarios y en las grandes empresas que poseen más de un punto de acceso a Internet,

que por su tamaño y diversidad de servicios prestados son un blanco atractivo de malwares, phising, ransomwares, DoS, baiting, etc.

Actualmente los administradores de la seguridad de las redes de datos inalámbricas de las universidades de Arequipa, no gestionan adecuadamente los dispositivos de seguridad debido a las grandes cantidades de ataques que reciben a cada minuto (generalmente de DoS y DDoS, virus, phising y spam); grandes volúmenes de información; falta de capacitación y entrenamiento del personal; falta clasificar los mismos en base a frecuencias, tendencias o niveles de riesgo, haciendo muchas veces insegura, poco confiable e inoperable las redes de todo el campus universitario, generando malestar en los usuarios directos como alumnos, docentes, administrativos y en los usuarios indirectos como usuarios remotos, proveedores y e-users. En la Universidad Católica de Santa María (UCSM) se viene utilizando el firewall PaloAlto, WAF (Web Aplication firewall) y antivirus para permitir o denegar el tráfico, de interés de la UCSM, y de esta forma proteger a los diferentes dispositivos y datos sensibles, así como para contrarrestar los virus, pero a pesar de haber invertido miles de dólares, estos no se explotan al máximo.

Entonces se hace necesario capturar los diferentes tipos de tráfico, detectar variaciones en los ataques, detectar nuevos patrones y sus cambios en los perfiles en las diferentes capas del Modelo OSI, para su posterior análisis exploratorio utilizando herramientas de minería de datos porque el volumen de la información que se almacena en pocos minutos es inmensa y software estadístico como las hojas de cálculo cuyo límite son 65535 registros, no lo analizarían; esto permitirá a los administradores de la red implementar mejoras en la gestión de la seguridad informática en base a procedimientos estandarizados como la ISO 27002-2013 (en el Perú NTP-ISO/IEC 27002:2017) y que sus redes se tornen menos vulnerables y los sistemas universitarios se encuentren siempre disponibles.

1.3. FORMULACIÓN DEL PROBLEMA

- **Problema General**

¿Cómo el análisis exploratorio de ataques informáticos aplicando herramientas de minería de datos, está relacionado con la gestión de seguridad de las redes inalámbricas en universidades de Arequipa?

- **Problemas Específicos**

- ¿Cómo realizar la caracterización de ataques informáticos para la gestión de seguridad de las redes inalámbricas?
- ¿Cuáles son las características de los ataques informáticos para determinar los patrones de predicción de acuerdo a la categoría de la amenaza utilizando modelos de clasificación, asociación y segmentación?
- ¿Cómo es la relación de la gestión de seguridad y los controles aplicados según estándares ISO 27002?

1.4. ANTECEDENTES

Se ha consultado varias obras de los siguientes autores a fin de tener una visión completa del problema:

Klenzi y Lopez (2017), en su propuesta "Detección de Ataques DoS con Herramientas de Minería de Datos", trabajaron en modelar y desde allí mitigar, ataques a un servidor de red por denegación de Servicios (Denial of Services) DoS mediante el análisis offline de un flujo de datos simulados y la utilización de algoritmos y herramientas correspondientes a Data StreamMining (Minería de datos - MD- en flujos de datos continuos). La aplicación utiliza módulos y algoritmos específicos de las herramientas de software libre RapidMiner (RM) 5.3.015 y KNIME 3.3. El autor concluye que al utilizar las herramientas de ML se ha observado que KNIME genera archivos de simulación de mayor tamaño que los que genera RM, guardando las visualizaciones de salida, conjuntamente con la estructura del WF. Así, al momento de cargar un WF, se pueden observar las salidas de los diferentes módulos, en tanto en RM cada vez que se carga un WF se debe ejecutar y esperar las visualizaciones de salida hasta que la ejecución termine.

Aguirre (2017) en su tesis de investigación sobre la "Metodología ISO 27000 para optimizar rendimiento de redes corporativas medianas móviles manteniendo estándares de disponibilidad y seguridad", tuvo como objetivo configurar controles de seguridad y optimización de gestión de redes, a los procesos de implementación, gestión y monitoreo de una red corporativa móvil mediano utilizando para su despliegue, gestión y monitoreo del estándar ISO 27001. El resultado final de la investigación plantea optimizar la implementación de infraestructura de red para enlaces de datos e internet dirigidas a redes corporativas medianas que se desarrollen

fuera de su centro habitual de trabajo, mediante el dimensionamiento de hardware y software en base a porcentajes de utilización observados, aplicando las políticas y procesos de seguridad del estándar ISO 27001, se considera también parámetros mínimos de infraestructura de acceso última milla. El autor concluye el trabajo indicando que a fin de garantizar la disponibilidad del servicio ofrecido y de evitar cualquier intrusión de personas no autorizadas dentro de la red o mal uso de la misma, se implementará con el personal de Sistemas, políticas de seguridad de la información, para lo cual se presentará una propuesta en la que se recomendarán las políticas específicas de seguridad para la aprobación del cliente.

En el trabajo de Galán (2015) "Aplicación de la metodología CRISP-DM a un proyecto de minería de datos en el entorno universitario", indica que es importante hacer minería de datos partir de una base de datos o data warehouse (almacén de datos) que contenga la información que se quiere analizar y que ésta información esté correctamente estructurada. La minería de datos trata de sacar toda la información posible de los almacenes de datos, no se conforma sólo con la visualización de estos datos como podría pasar con las consultas simples, si no que trata de obtener resultados en cuanto a la relación que existe entre los mismos y como podrían dar beneficios de algún modo al negocio. Aquí aplica estrictamente cada una de las distintas etapas de la metodología CRISP-DM sobre los datos académicos almacenados por la universidad en sus sistemas informáticos. De esta forma se pretende sacar conclusiones que ayuden a mejorar los servicios que ofrece la universidad a sus alumnos. También demuestra que la metodología CRISP-DM es una metodología que funciona y que además es sencilla de usar, ya que solamente hay que seguir una serie de fases que están claramente delimitadas y está pensada para que cualquier persona con conocimientos de bases de datos y estadística pueda utilizarla.

Mejía (2015) propone desarrollar un plan de seguridad informática basándose en la serie de normas ISO 27000 y de esta forma contemplando el uso de software libre y licenciado, facilitó conocer las fortalezas y debilidades de la institución pública, pero que sirve como referencia a otras instituciones educativas privadas para proveer directrices y comenzar con el análisis de riesgo de sus activos y así fijar una escala del riesgo y proceder a realizar sus planes de seguridad informática. Como resultado del desarrollo de la investigación, el autor indica que el plan de seguridad informática es importante al inicio realizando un análisis de la norma ISO 27001 para encontrar parámetros generales de los documentos y pasos a seguir en el área de sistemas de

gestión de seguridad de la información; usa la ISO 27005 para realizar la valoración de activos como el análisis de riesgos y un listado de amenazas a las cuales están expuestos los equipos informáticos y finalmente para la incorporación de controles y aseguramiento de las diferentes áreas y equipos de las empresas, en el plan de seguridad sugiere tomar en cuenta la ISO 27002. Con todo esto para la implementación del plan de seguridad informática es necesario que el equipo de trabajo se comprometa con la implementación de las políticas, ya que son la base para obtener una mejora significativa en la CID de la información y servicios.

Vallejo y Tenelanda (2012) en su investigación "Minería de datos aplicada en detección de intrusos", indican que es necesario detectar las anomalías que se presenten en los registros de acceso de las redes de datos, lo cual es posible mediante técnicas de detección de intrusos, analizando aquellos accesos que pongan en peligro la confidencialidad, integridad, disponibilidad y no repudio; pero ese registro y almacenamiento para su posterior análisis, crea gran volumen de datos que, a simple vista, no son fáciles de analizar y de correlacionar entre sí, por lo que se utiliza la minería de datos para buscar información no trivial que se encuentre oculta o dispersa en ellos, es decir, se exploran los datos para descubrir la interconexión e interrelación y poder obtener la información oculta, a fin de encontrar y predecir ataques que no son detectados por antivirus, cortafuegos o sistemas de detección de intrusos. Concluye que la minería de datos basada en una metodología adecuada, puede ser muy útil en el proceso de exploración de datos y que mediante tecnologías analíticas y procesos estadísticos permite generar reglas a partir de datos históricos de capturas, para generar reglas y patrones que permiten predecir intrusiones.

Keita (2012), en su investigación "Detección de Intrusos en la Capa de Enlace del Protocolo 802.11", pretende obtener soluciones que detecten intrusiones para redes WiFi, basándose en el análisis de la información y comportamiento de las tramas de control y de gestión de las mismas. Además, estudia y analiza la información de control y de gestión de la trama MAC del 802.11 a fin de obtener algoritmos que detecten intrusiones debido a las vulnerabilidades de los protocolos en esa trama. Como resultados, se obtuvieron algoritmos de detección de intrusiones basándose en las vulnerabilidades de los paquetes de control y de gestión de la 802.11 los cuales condujeron a la implementación de un Script que detecte hasta un 95%, las DoS causadas por los ataques de RTS/CTS falsos, de des-autenticación y de des-asociación, ofreciendo así una nueva perspectiva en el trabajo de control y gestión de

seguridad en dichas redes.

Candia (2019) en su trabajo de investigación, proucra predecir el rendimiento académico de los estudiantes de la UNSAAC en el primer semestre a partir de sus datos de ingreso y determinar el éxito o fracaso de este primer año que puede influir en el desempeño de los años posteriores. Para ello desarrolló la metodología CRISP-DM para generar los modelos predictivos, utilizó la herramienta de análisis exploratorio de libre acceso Weka y diferentes algoritmos supervisados de clasificación para predecir el rendimiento académico. Concluye que el mejor algoritmo para predicción de acuerdo a su data fue RandomForest seguido de Regresión logística.

En la tesis de grado de Romo y Valarezo (2012) "Análisis e implementación de la norma ISO 27002 para el departamento de sistemas de la Universidad Politécnica Salesiana Sede Guayaquil", primeramente se hizo una detección de la falta de políticas y normas de seguridad de la información en la Universidad y frente a esta problemática se propuso la implementación de la norma ISO 27002 para proteger a los activos de información definiendo políticas que serían aplicadas por el personal docente, administrativo y externos.

Según Britos (2010) en su estudio "Detección de Intrusiones en redes de datos con captura distribuida y procesamiento estadístico", propone modelos estadísticos y clasificadores multivariables para detectar perfiles de tráfico anómalos, así también el uso de clasificadores basados en redes neuronales o en agentes cooperantes. El propósito de este trabajo es investigar nuevos métodos, los que utilizados en otras áreas de la ciencia han enfrentado problemáticas similares para aplicarlos con eficacia en la detección de intrusiones en redes de datos. En este sentido, el proceso que sigue la detección de intrusiones involucra etapas tales como, la captura de los datos, la selección estadística de los más relevantes, hasta llegar al mecanismo de decisión que detecta a un ataque. Esta investigación concluye con el uso de 2 herramientas de detección de intrusiones: las redes neuronales y las colonias de hormigas. Con las redes neuronales no se ha logrado un buen desempeño en ataques complejos, por lo que se aplica colonias de hormigas a través de los algoritmos de agrupamiento como K-Means, EM y el agrupamiento Fuzzy c-means, se ha presentado un sistema de autoaprendizaje no supervisado para la detección de intrusiones en redes.

1.5. JUSTIFICACIÓN DE LA INVESTIGACIÓN

La presente investigación se justifica en el análisis del marco teórico, lo que permitirá obtener el dominio teórico para realizar la captura del tráfico que atraviesa el campus universitario a través de software como sniffers y sistemas de detección de intrusos; comprender minuciosamente el comportamiento de las variables que se estudiará en la presente investigación, asimismo, es importante porque se pretende estudiar los tipos de ataques informáticos, su duración, su nivel de efectividad, zona de origen y destino, IP y puerto de origen y destino, aplicaciones atacadas y posibles acciones de defensa aplicadas; así mismo realizar un análisis exploratorio de los mismos a través de herramientas que manejan gran cantidad de información como son las de minería de datos las que nos permitirán asociar los registros de datos, detectar anomalías, clasificarlas, segmentarlas y desarrollar predicciones en base a una metodología muy utilizada como es CRISP-DM; todo esto servirá de apoyo para que los administradores de seguridad de las redes informáticas perciban las falencias de configuración y apliquen políticas se seguridad basada en estándares internacionales, asegurando de esta manera la disponibilidad, confiabilidad y disponibilidad del servicio y de la información, que redundará en una mejor gestión de seguridad de las redes inalámbricas y cableadas del campus universitario.

1.6. LIMITACIONES DE LA INVESTIGACIÓN

Las limitaciones encontradas fueron:

- Falta del tiempo de los administradores de redes para brindar información del funcionamiento de sus redes.
- Falta de autorización para entregar información del tráfico que atraviesa las redes informáticas.
- Falsos positivos encontrados durante la captura de tráfico de Internet
- El factor tiempo, por las actividades laborales del investigador.

1.7. OBJETIVOS

- **Objetivo General**

Determinar cómo el análisis exploratorio de ataques informáticos aplicando herramientas de minería de datos, está relacionado con la gestión de seguridad de las redes inalámbricas en universidades de Arequipa.

- **Objetivos Específicos**

- Realizar la caracterización de ataques informáticos para gestionar la seguridad de las redes inalámbricas.
- Analizar las características de los ataques informáticos para determinar los patrones de predicción de acuerdo a la categoría de la amenaza utilizando modelos de clasificación, asociación y segmentación.
- Evaluar la relación de la gestión de seguridad con los controles de seguridad aplicados según estándares ISO 27002.

1.8. HIPÓTESIS

- Hipótesis General

El análisis exploratorio de ataques informáticos aplicando herramientas de minería de datos, está relacionado con la gestión de seguridad de redes inalámbricas en universidades de Arequipa.

- Hipótesis Específicas

- La caracterización de ataques informáticos está asociado a la gestión de seguridad de las redes inalámbricas.
- Las características de los ataques informáticos permite determinar los patrones de predicción de acuerdo a la categoría de la amenaza utilizando modelos de clasificación, asociación y segmentación.
- La gestión de seguridad está relacionada con los controles de seguridad aplicados según estándares ISO 27002.

II. MARCO TEÓRICO

2.1. MARCO CONCEPTUAL

2.1.1 Ataques informáticos

Los ataques informáticos o ciberataques, son aquellos actos que buscan hallar las vulnerabilidades del software y hardware de los sistemas informáticos y de las redes de una organización y aprovecharse de ellos con fines lúdicos como es el caso de los entornos universitarios que estamos analizando y otros con fines económicos, de tal forma que perjudiquen el normal funcionamiento de estos sistemas o redes hasta dejarlos inoperativos, causando diversos tipos de daños desde menores hasta los más severos.

Tipos de ataques informáticos

Según la consultora Kaspersky (2016), sobre los riesgos de seguridad en tecnologías de la información (TI), reveló que el 82% de las empresas a nivel mundial ha sufrido entre uno y cinco incidentes de exposición, filtración o pérdida de datos en los últimos 12 meses. Como resultado de esa clase de incidentes, el 10% de ellas perdió acceso a información crítica durante una semana y el 15% sufrió interrupciones que le impidieron realizar transacciones comerciales durante más de siete días.

Dicha consultora ha clasificado a estos ataques como los más frecuentes, convirtiendo a la seguridad de la información en una de las prioridades de las empresas desde las Pymes hasta las grandes compañías, dado que el activo más importante debe estar protegido y las infraestructuras de redes deben ser menos vulnerables.

Así mismo la descripción de los ataques se comúnmente se han detectado en los campus universitarios son:

a) Malware

Software malicioso, es una categoría diseñado para infiltrarse y dañar un sistema de información sin ser detectado. Aunque el malware se utiliza para referirse de forma general a un software malicioso, existen diversos tipos de malware que responden a

características propias y comportamientos diferentes. Entre los malwares más utilizados, destacan:

- **Virus**: un código maligno que infecta los ficheros del dispositivo en forma de archivo ejecutable (o archivo .exe), y que se vale del desconocimiento de los usuarios para infectar un equipo u ordenador.
- **Gusano:** que es un software un poco más sofisticado que el virus, que crea copias de sí mismo con el objetivo de afectar otros equipos.
- **Troyano:** programas diseñados para ingresar en los sistemas de seguridad y permitir el acceso a otros archivos maliciosos, aparentando ser otra cosa.
- **Spyware**: programas que espían un dispositivo para obtener información privada y que pueden instalar otros softwares maliciosos.
- **Ransomware**: secuestran la información de valor de un dispositivo, con el fin de solicitar una transferencia en criptomoneda o monedas digitales a modo de rescate.

Para minimizar el riesgo de malwares, es importante contar con un potente software antivirus y antimalware. En el caso de las empresas, adicionalmente se deberá capacitar al personal para que este no abra adjuntos de correos electrónicos procedentes de fuentes desconocidas o poco confiables.

b) Ataque DDoS

Es uno de los ataques más frecuentes en Internet. También conocido como "denegación del servicio distribuida" (que proviene del inglés "distributed denial of service"), consiste en el bloqueo al acceso de un sitio web y, en simultáneo, el ataque al servidor mediante el ingreso de un gran volumen de información basura (por ejemplo, el rellenado de formularios con datos falsos o envío de solicitudes). Esto origina una saturación en el flujo del servidor, colapsa el sitio web o determina la pérdida de conectividad en este espacio. Normalmente, estos ataques se hacen a través de ordenadores infectados con troyanos.

Una medida básica a tomar para evitar un ataque DDoS es adicionar la opción de protección contra este tipo de ataques en el firewall o instalar un sistema anti-DDoS. Aunque lo más recomendable es solicitar al proveedor de servicios de Internet que habilite la protección DDoS desde su red, ya que cuenta con mayor capacidad de protección y el ataque es mitigado antes de consumir recursos del Internet contratado.

c) Phishing

Es un método usado por los atacantes para suplantar la identidad de un usuario o de una empresa mediante una comunicación electrónica (correo, mensajería instantánea, etc.), con la finalidad de obtener datos personales y bancarios.

Si bien el phishing no es un ataque directo contra una web o sus servidores, este método busca desviar el flujo de clientes, ingresos o búsquedas hacia un portal falso. Aunque focaliza sus ataques a tiendas o portales de venta online, el phishing es también frecuente en sitios que ofrecen servicios financieros o en aquellas webs que mantienen un flujo de crédito constante. Una forma persistente de forzar la confusión del usuario es que se anuncia la aparición del sitio falso en la red e, incluso, se paga por aparecer primero en los buscadores.

Para evitar caer en este tipo de ataques, es importante verificar que el remitente de cualquier correo electrónico se corresponda con la entidad a la cual dice pertenecer y que no contenga letras o caracteres extraños. Otra forma de identificar estos sitios falsos es observar que en la barra de direcciones aparezca la etiqueta de “sitio seguro” y desconfiar de enlaces insertos en nuestros e-mails.

d) Baiting

Consiste en un ataque dirigido a infectar equipos y redes a partir de dispositivos de almacenamiento extraíbles como pen-drives, tarjetas SD o discos duros externos. A través de estos equipos, los atacantes introducen archivos infectados con malwares. Al ser un software malicioso que ingresa de forma externa al ordenador, la estrategia de ataque suele ser colocar estos dispositivos de almacenamiento externo en las inmediaciones de la empresa, a fin de que sean utilizados y conectados a los equipos corporativos por el personal.

Lógicamente, la mejor forma de evitar un ataque de este tipo será concientizar a sus colaboradores sobre la importancia de no conectar dispositivos de almacenamiento desconocidos y de solo utilizar aquellos inventariados por la empresa.

e) Scripting

Llamado también Cross Site Scripting o XSS, es un tipo de ataque directo o reflejado que consiste en insertar código malicioso en Java Script u otro lenguaje, en la página web (webshell), portal web o en aplicaciones locales. El atacante busca

obtener información del usuario victima una vez que este se conecta a la página web desde la obtención de contraseñas, robo de sesión u otra información a través de algún campo en el que se solicite el ingreso de datos, derivándolo a una página falsa o inclusive instalar un malware en su computador. Este ataque es bloqueado por el WAF de la UCSM.

f) SQL Injection

Consiste en inyectar código malicioso en las bases de datos que usan MySQL u otro gestor pero que interprete las órdenes del servidor en lenguaje SQL, aprovechando las vulnerabilidades de programación (webshell) de los sitios web con mal diseño e infiltrando datos a través de los cuadros de entrada directamente a la base de datos back end, de esta forma los atacantes podrán leer, modificar y borrar los datos de la misma base de datos. Este ataque es bloqueado por el WAF de la UCSM.

g) Fuerza bruta

Este ataque consiste en averiguar contraseñas, probando todas las combinaciones posibles hasta dar con la correcta de forma manual o hasta incluso utilizando grandes diccionarios de datos. Estos ataques son los más utilizados para el robo de contraseñas en Internet dado que no es necesario tener gran conocimiento en seguridad informática ya que hay actualmente programas que lo realizan automáticamente sin mucho esfuerzo, pero que dependiendo de la complejidad y tamaño de la clave puede tomar varias horas de trabajo.

h) Info leaks

O fuga de información, este ataque permite revelar información confidencial de una aplicación, software, usuarios, entornos de datos, redes informáticas completas, etc; los que pueden ser aprovechados por los atacantes para beneficio suyo o de terceros, explotando la aplicación, redes o usuarios.

Herramientas de captura de tráfico

Las herramientas que permiten capturar los diversos tipos de tráfico que atraviesan las redes de datos cableadas o inalámbricas, son conocidas como sniffers y pueden ser de libre acceso o licenciadas. Ambas cumplen funciones similares que es recolectar paquetes de datos desde la capa 2 hasta la capa 7 del Modelo OSI, para ello el dispositivo que contiene instalado este sniffer, deberá tener configurada si tarjeta de red en modo promiscuo. Muchas de ellas capturan tráfico de protocolos que no cifran sus datos y otros que si lo hacen.

Según Informática Digital (2018), los principales usos que se les puede dar a los sniffers son:

- Captura automática de contraseñas enviadas en claro y nombres de usuario de la red. Utilizada por crackers para atacar sistemas a posteriori.
- Conversión del tráfico de red en un formato inteligible por los humanos.
- Análisis de fallos para descubrir problemas en la red.
- Medición del tráfico, mediante el cual es posible descubrir cuellos de botella en algún lugar de la red.
- Detección de intrusos, con el fin de descubrir crackers.
- Creación de registros de red, de modo que los crackers no puedan detectar que están siendo investigados.
- Para los desarrolladores, en aplicaciones cliente-servidor. Les permite analizar la información real que se transmite por la red

Podemos resumir las características de dichas herramientas a propuestas por Informática Digital (2018), para las mejores 7 herramientas de captura de tráfico:

a) Wireshark

Conocido como Ethereal, es un analizador de protocolos utilizado para realizar análisis en vivo o fuera de línea y solucionar problemas en redes de comunicaciones, para desarrollo de software y protocolos, y como una herramienta didáctica. Soporta los protocolos 802.11 a/b/g/n a nivel inalámbrico, soporta los protocolos IPsec, ISAKMP, Kerberos, SNMPv3, SSL/TLS, WEP, and WPA/WPA2. Los resultados pueden ser exportados en formatos como XML, PostScript®, CSV, o texto plano. Está disponible para Windows, Mac OS, FreeBSD, Solaris y Linux y es de libre acceso.

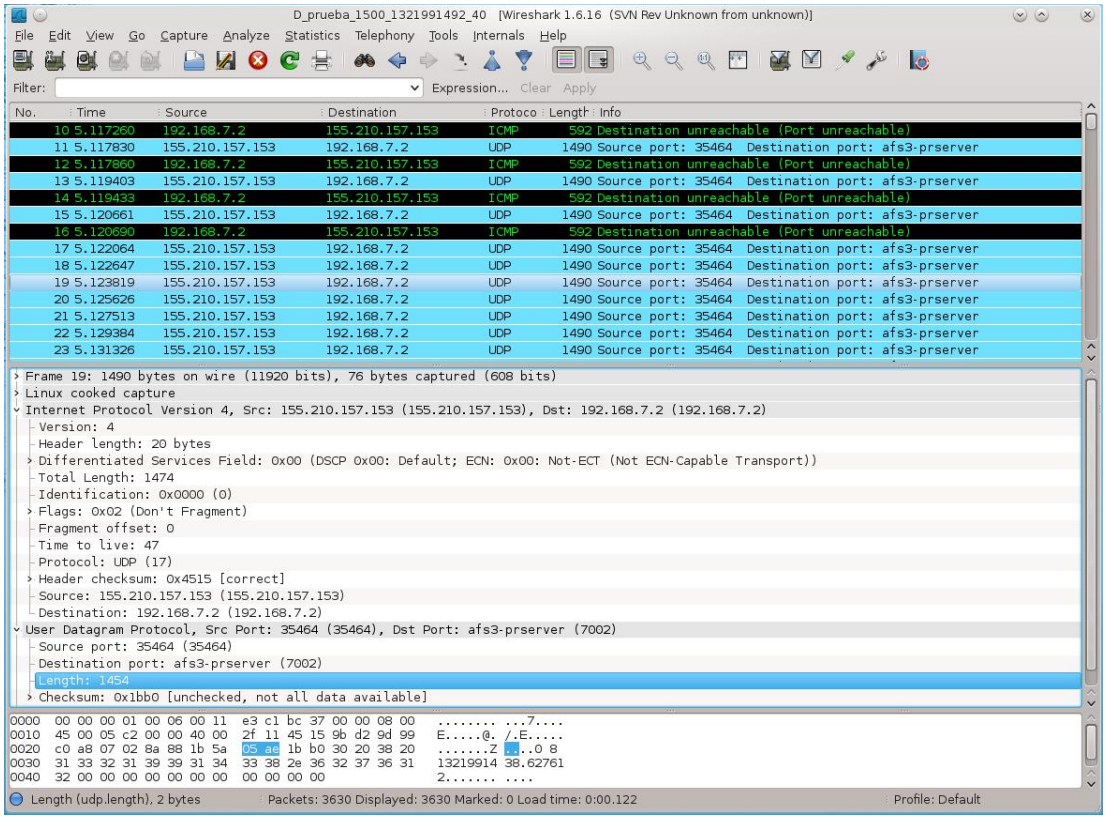

Gráfico 1. *Vista del contenido de un paquete UDP utilizando Wireshark*

Fuente: https://www.telecomsharing.com/es/laboratorio/herramientas/item/30-las-2-mejores-herramientas-para-la-captura-y-la-interpretacion-de-datos-de-red

b) Microsoft Message Analyzer

Esta herramienta gratuita desarrollada por Microsoft nos ofrece un análisis más avanzado en el proceso de análisis de trafico de red gracias a su avanzada técnica de captura local o remota y permite ver la información de la red más de 300 públicos y los protocolos de red, incluyendo los paquetes de red inalámbrica. Permite importar y exportar datos para su análisis, incluye diversos formatos de visualización y verificar el estado de los protocolos y realizar implementaciones de los mismos

Además, puede ser utilizado por los principiantes sólo para analizar su tráfico de red propia, o por los administradores de red para analizar la red completa organización por la inhalación de paquetes de red. Es de libre acceso.

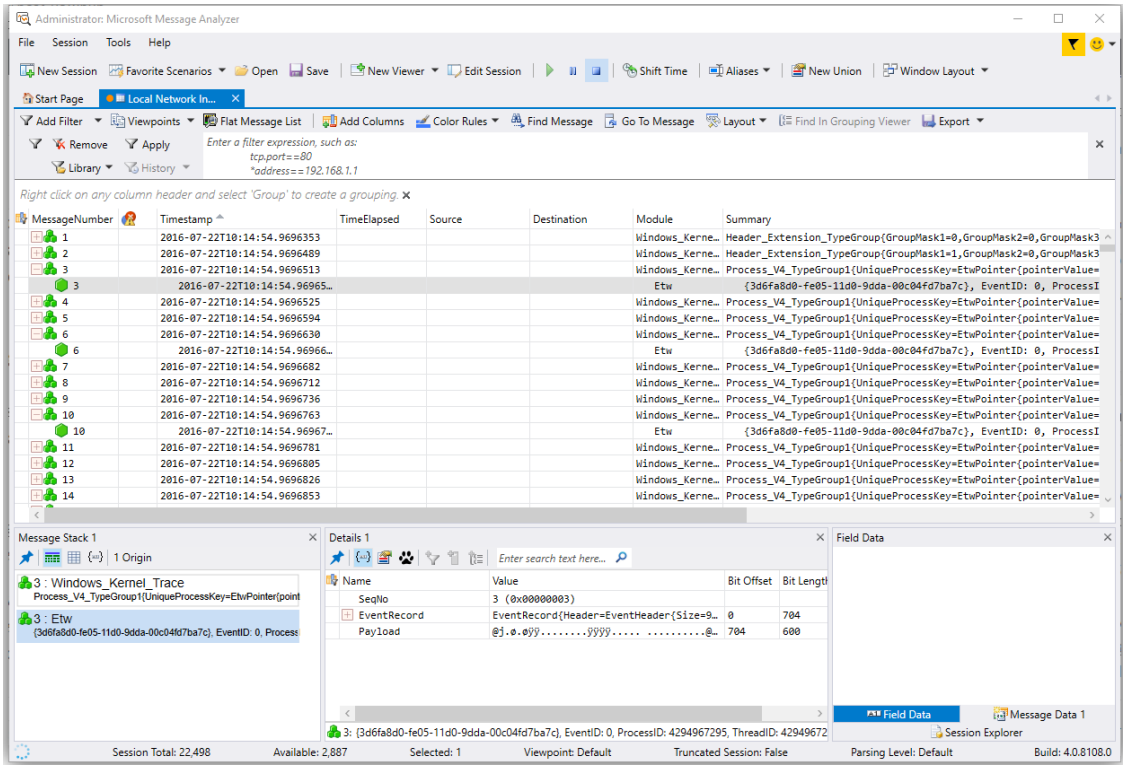

Gráfico 2. *Vista del contenido de tráfico con Microsoft Network Monitor*

Fuente: https://www.locurainformaticadigital.com/2018/03/02/7-mejores-analizadores-de-red-sniffers-windows-y-linux/

c) Tcpdump

Es una herramienta para captura del tráfico que circula por la red en tiempo real, lo que permite la depuración de la salida obtenida por medio de filtros, filtrando capturas de puerto específico, por tipo de protocolo, dirección origen o destino, en una interfaz en específico y otros., además, de ser idónea para la captura de paquetes en forma desatendida, ya que se puede gestionar por línea de comandos. Disponible para Linux y Windows.

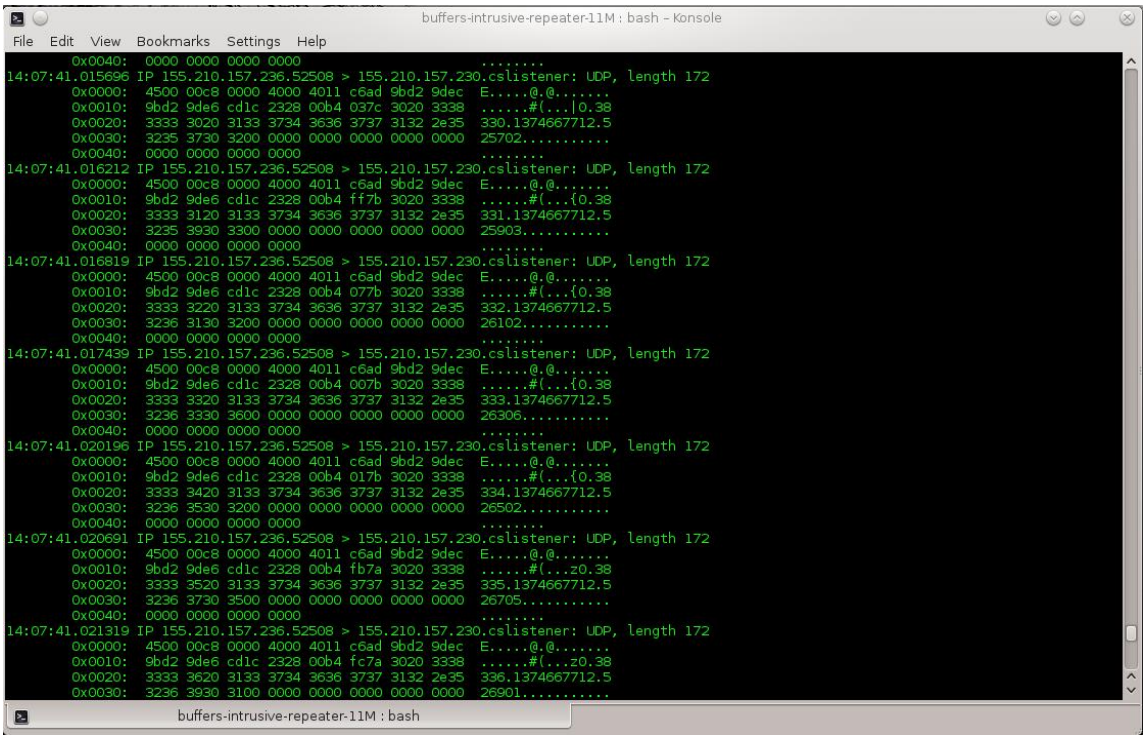

Gráfico 3. *Vista del contenido de varios paquetes por línea de comando utilizando TCPDUMP*

Fuente: https://www.telecomsharing.com/es/laboratorio/herramientas/item/30-las-2-mejores-herramientas-para-la-captura-y-la-interpretacion-de-datos-de-red

d) Escáner WiFi Lizard Systems

Herramienta de solución de problemas de WiFi y escáner WiFi multiusos diseñada para resolver todos los problemas relacionados con WiFi de manera conveniente. Esta herramienta fácil de usar les permite a sus usuarios ubicar fácilmente redes WiFi visibles completas y su información correspondiente de forma sistematizada. Esta herramienta primero obtiene el nombre de la red, la potencia de la señal, la calidad de la señal, la dirección MAC, la información del canal, la velocidad de datos máxima y alcanzable, la información de seguridad y mucho más. Disponible para:Windows

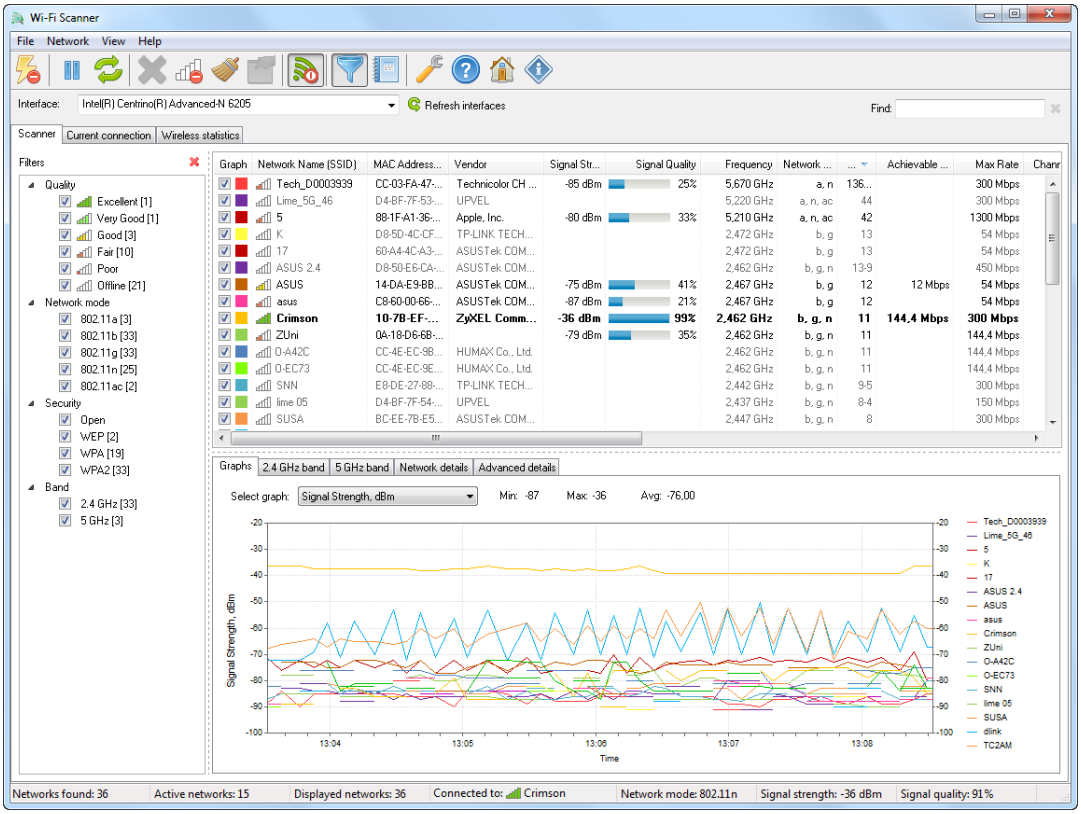

Gráfico 4. *Vista del contenido de paquetes con Scaner Wifi Lizard Systems*

Fuente: https://www.locurainformaticadigital.com/2018/03/02/7-mejores-analizadores-de-red-sniffers-windows-y-linux/2/

e) Capsa Network Analizer

Analizador de red de paquetes de red, útil para los administradores de red para supervisar, diagnosticar y solucionar sus problemas en la red. Realiza un análisis en tiempo real, soluciona problemas de red integrado, su uso es sencillo, es de libre acceso, tiene posibilidad de ejecutar múltiples proyectos de forma simultánea, escanea puertos, genera reportes e incluye más de 300 protocolos.

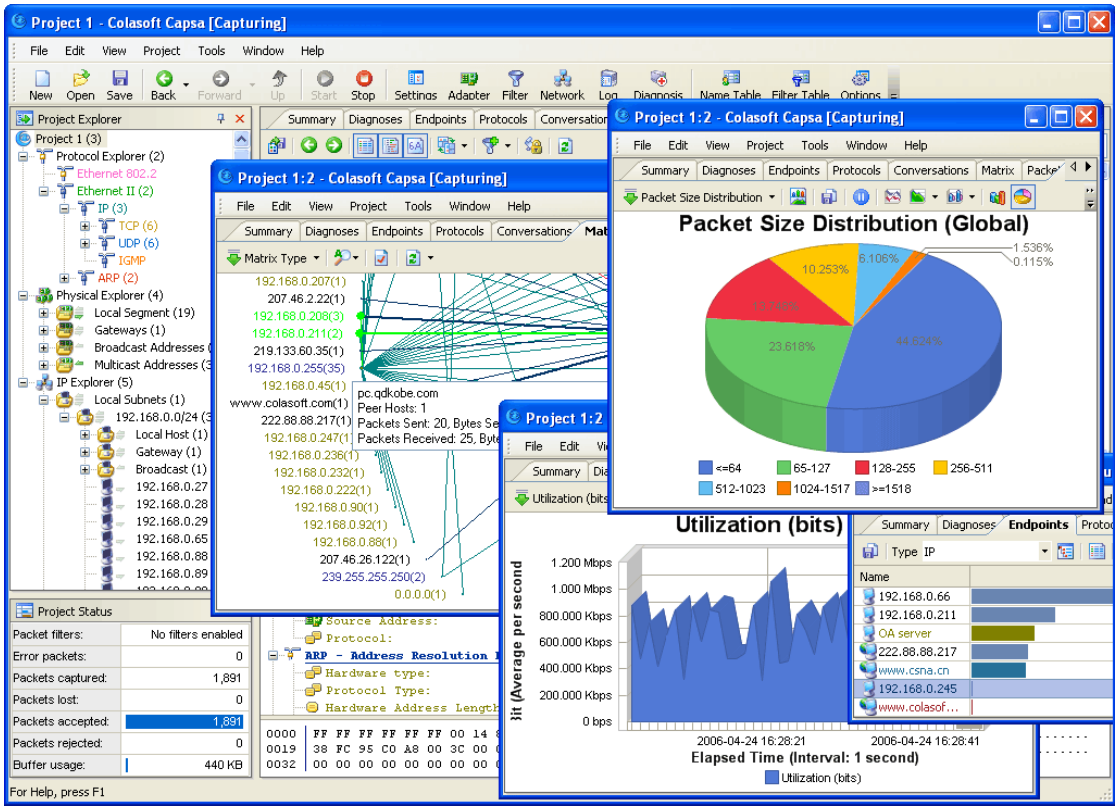

Gráfico 5. *Vista del sniffer Capsa*

Fuente: https://underc0de.org/foro/pentest/las-5-mejores-herramientas-analizadoras-de-red-y-sniffers!/

f) Netcat

Netcat es una herramienta de red que permite a través de intérprete de comandos y con una sintaxis sencilla, abrir puertos TCP/UDP en un host para estar en escucha; asocia una shell a un puerto en concreto (para conectarse por ejemplo a MS-DOS o al intérprete bash de Linux remotamente) y forzar conexiones TCP/UDP útil para realizar rastreos de puertos o realizar transferencias de archivos bit a bit entre dos equipos. Entre sus aplicaciones, se tiene la depuración de aplicaciones de red y apertura de puertas traseras en un sistema. Disponible para: Windows y Linux.

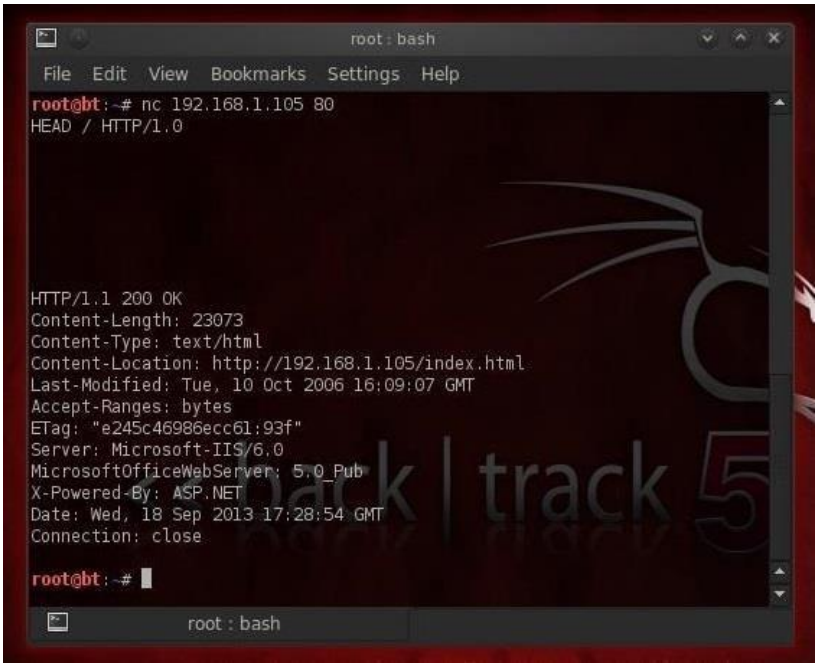

Gráfico 6. Vista del sniffer Netcat

Fuente: https://www.locurainformaticadigital.com/2018/03/02/7-mejores-analizadores-de-red-sniffers-windows-y-linux/2/

g) Snort. Security Onion Live. SGUIL, Squert y Suricata

Según Brisa (2015), se tiene una herramienta para monitorear y vigilar el tráfico a nivel de la capa de red, monitoreando los datos que entran y salen del perímetro de la red así como los datos que circulan dentro de la misma. Esta disciplina se llama NSM (Network Security Monitoring). Para ello se cuenta con las herramientas integradas en una máquina virtual Security Onion que son Sguil, Squert y Suricata.

Sguil, es una consola de análisis de eventos de red diseñada por y para analistas, compuesta de varios sistemas que ayudan a monitorizar la seguridad de una red de ordenadores. Además del acceso a eventos en tiempo real, proporciona datos de sesión y capturas de paquetes de red. Desde su interfaz se visualiza estos datos y se tiene de gestión y clasificación de dichos eventos. Posee una estructura de cliente-servidor en la que los sensores (clientes) monitorizan los enlaces de red (snort, SANCP) y envían los datos al servidor de Sguil donde se almacenan en una base de datos para su posterior tratamiento. La información se visualiza mediante una interfaz gráfica multiplataforma escrito en TCL/TK.

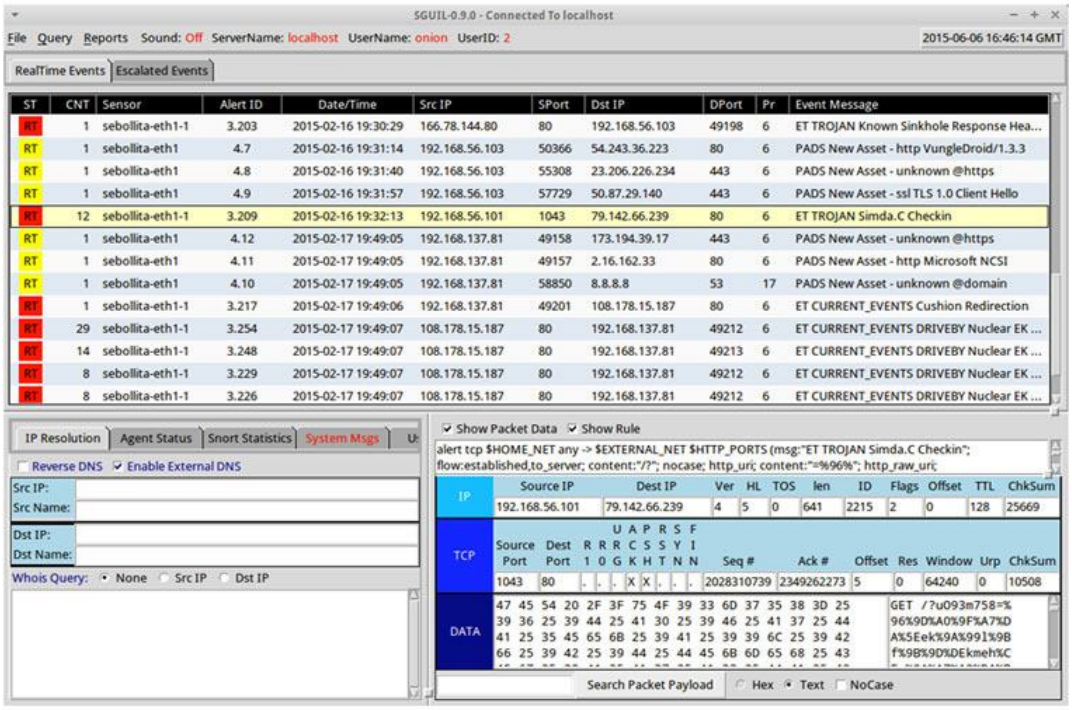

Gráfico 7. *Consola de Sguil*

Fuente: https://www.securityartwork.es/2015/06/22/squert/

SQUERT (Simple QUEry and Report Tool), es una evolución de Sguil, añade una interfaz web para visualizar y consultar los datos almacenados en una base de datos de Sguil. Su interfaz mejora la usabilidad y hace el entorno más amigable, añadiendo a su vez nuevas funcionalidades a una herramienta ya de por sí muy útil.

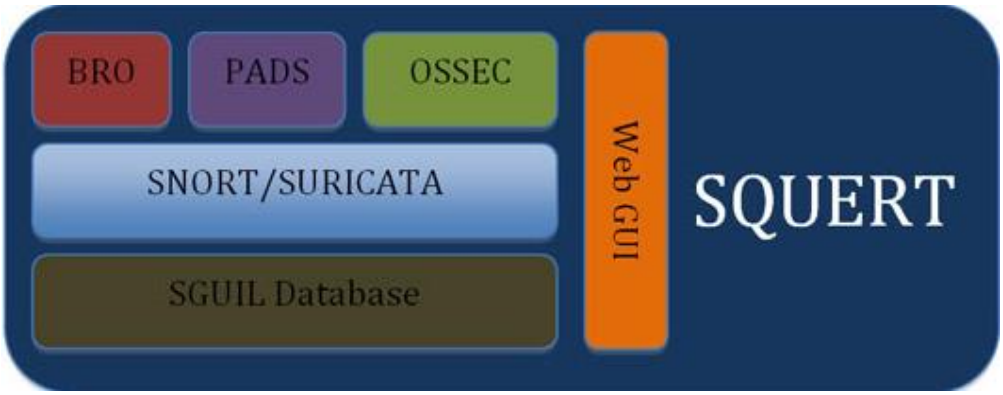

Gráfico 8. *Funcionalidades de Squert*

Fuente: https://www.securityartwork.es/2015/06/22/squert/

Desde el interfaz de SQUERT se puede agrupar y visualizar alertas del IDS (snort/suricata), alertas de host IDS proporcionadas desde OSSEC, PADs (Passive

Asset Detection) y eventos procedentes de logs de Bro. La combinación de toda esta información ayuda a tener una visión más completa de lo que pasa en los sistemas y determinar ante un incidente de seguridad los posibles equipos afectados, conexiones anómalas, etc.

El listado de alertas se agrupan todos los tipos definidos como NIDS + HIDS + PADS + OSSEC + BRO.

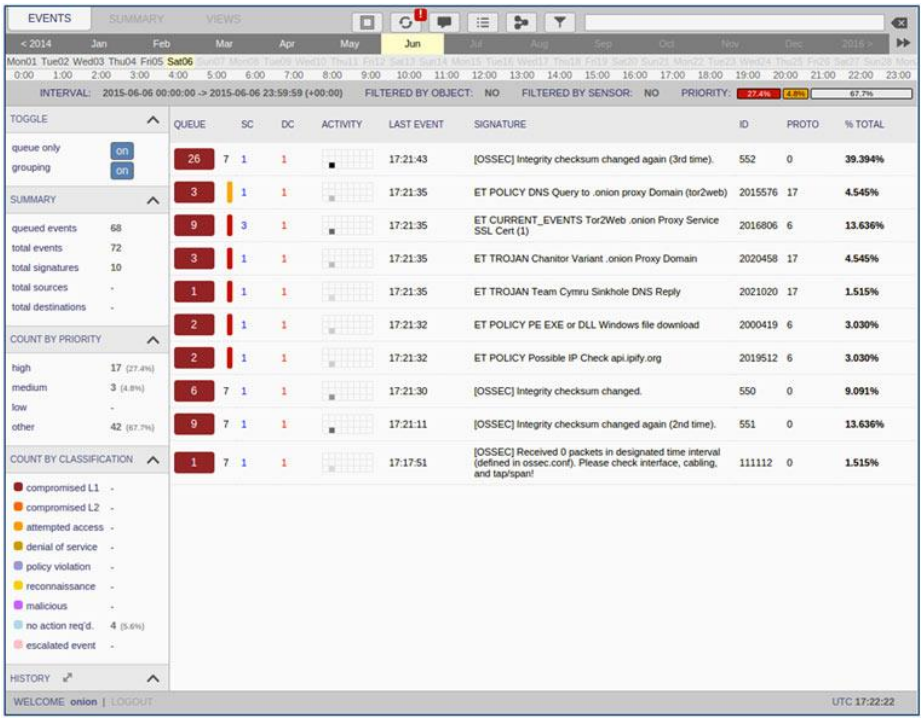

Gráfico 9. *Pantalla principal de Squert*

Fuente: https://www.securityartwork.es/2015/06/22/squert/

Squert sigue la misma clasificación que Sguil, se puede categorizar la alerta mediante el teclado (F1,F2,F3…) o seleccionando la alerta y haciendo click en el ícono de comentarios, que abrirá una ventana donde se selecciona la categoría y se añade un comentario descriptivo si se desea.

C1	Unauthorized admin Access
C2	Unauthorized user Access
C3	Attempted unauthorized Access
C4	Denial of service attack
C5	Policy violation
C6	Reconaissance
C7	Malware
C8	No action required
C9	Escalated event

Gráfico 10. *Clasificación de los tipos de ataque de Squert*

Fuente: https://www.securityartwork.es/2015/06/22/squert/

Una vez categorizados los eventos, estos desaparecen de la cola y aparecen en las estadísticas de la categoría seleccionada.

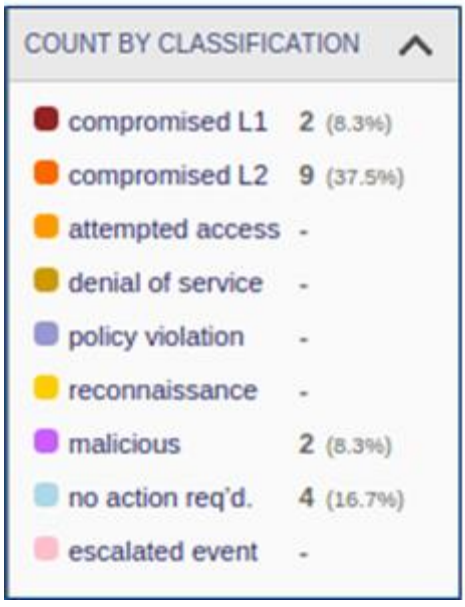

Gráfico 11. *Categoría de eventos de Sguil/Squert*

Fuente: https://www.securityartwork.es/2015/06/22/squert/

Squert también tiene la opción de crear reglas que permiten categorizar automáticamente las alertas que van entrando en el sistema.

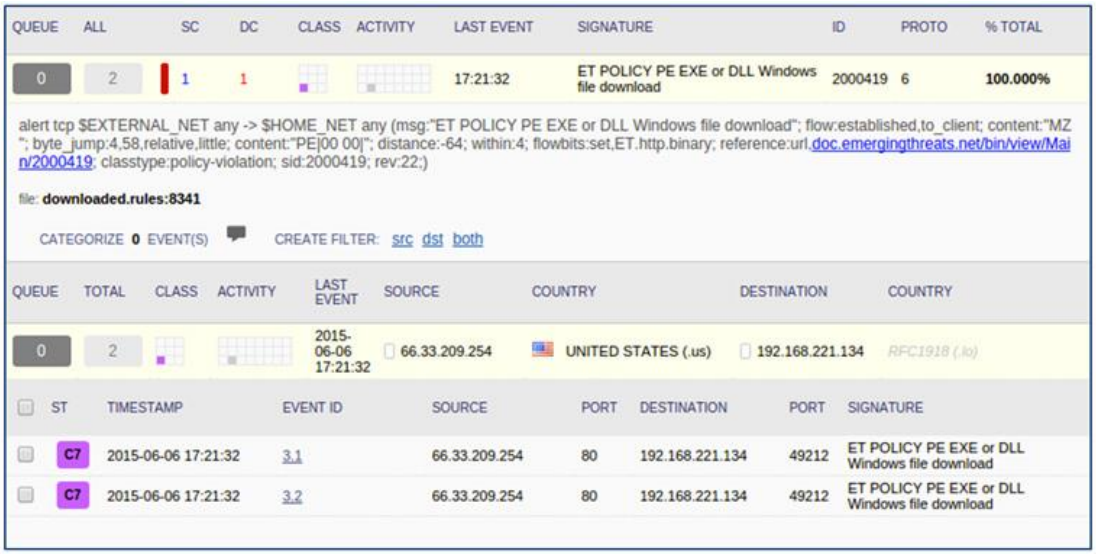

Gráfico 12. *Reglas de eventos de Sguil/Squert*

Fuente: https://www.securityartwork.es/2015/06/22/squert/

Una vez categorizada la alerta y para continuar con la investigación y análisis, se recupera el ejecutable descargado para procesarlo con ELSA. ELSA es un framework de análisis de logs centralizado, que puede integrarse con Squert, para realizar búsquedas directamente sobre los logs almacenados.

Gráfico 13. *Tabla snort para busqueda con ELSA*

Fuente: https://www.securityartwork.es/2015/06/22/squert/

h) Guia del administrador de PAN-OS

El firewall Palo Alto, nos devuelve una base de datos muy útil para el presente estudio, que consta de más de 2 millones de registros de amenazas recibidas por los dispositivos internos de la UCSM, el mismo que es analizado, pero conviene conocer cada uno de los campos obtenidos.

Esta información se ha obtenido de PAN-OS (2019) el mismo que enumera los campos estándar de cada tipo de registro que los firewalls de Palo Alto Networks pueden reenviar a un servidor externo, así como los niveles de gravedad, formatos personalizados y secuencias de escape y acciones a tomar en caso detectar estos ataques. Estos se describen a continuación:

Tabla 1.
Descripción de campos de registro de amenazas

NOMBRE DEL CAMPO DE RESGISTRO DE AMENAZAS	DESCRIPCIÓN
Tiempo de recepción (receive_time)	Hora en que se recibió el registro en el plano de gestión
Número de serie (serial)	Número de serie del firewall que generó el registro
Tipo (type)	Especifica el tipo de registro; los valores son tráfico, amenaza, configuración, sistema y hip-match
Subtipo (subtype)	Subtipo de registro de amenazas. Los valores incluyen: • datos: patrón de datos que coincide con un perfil de filtrado de datos. • archivo: tipo de archivo que coincide con un perfil de bloqueo de archivos. • Inundación: inundación detectada mediante un perfil de Protección de zona. • paquete: protección de ataque basada en paquetes activada por un perfil de Protección de zona. • exploración: exploración detectada mediante un perfil de Protección de zona. • spyware: spyware detectado mediante un perfil Anti-Spyware.

	• url: registro de filtrado de URL. • virus: virus detectado mediante un perfil de antivirus. • vulnerabilidad: vulnerabilidad de vulnerabilidad detectada a través de un perfil de Protección de vulnerabilidad. • wildfire: un veredicto de WildFire generado cuando el firewall envía un archivo a WildFire según un perfil de WildFire Analysis y un veredicto (malicioso, grayware o benigno, dependiendo de lo que esté registrando) se registra en el registro de Envíos de WildFire. • virus wildfire: virus detectado mediante un perfil de antivirus..
Tiempo generado (time_generated)	Hora a la que se generó el registro en el plano de datos
IP de origen (src)	Dirección IP de origen de sesión original
IP de destino (dst)	Dirección IP de destino de la sesión original
IP de origen NAT (natsrc)	Si se realizó NAT de origen, la dirección IP de origen posterior a NAT
IP de destino NAT (natdst)	Si se realizó NAT de destino, la dirección IP de destino posterior a NAT
Nombre de la regla (rule)	Nombre de la regla con la que coincidió la sesión.
Usuario de origen (srcuser)	Nombre de usuario del usuario que inició la sesión
Usuario de destino (dstuser)	Nombre de usuario del usuario al que estaba destinada la sesión
Aplicación (app)	Aplicación asociada a la sesión

Sistema virtual (Porys)	Sistema virtual asociado a la sesión
Zona de origen (from)	Zona de la que se obtuvo la sesión
Zona de destino (to)	Zona a la que estaba destinada la sesión
Interfaz de entrada(inbound_if)	Interfaz de la que se originó la sesión
Interfaz de salida(outbound_if)	Interfaz a la que estaba destinada la sesión
Perfil de reenvío de registros(logset)	Perfil de reenvío de registros que se aplicó a la sesión
ID de sesión (session_id)	Un identificador numérico interno aplicado a cada sesión.
Repetir recuento (Repeat Count)	Número de sesiones con la misma IP de origen, IP de destino, aplicación y subtipo vistas en 5 segundos; utilizado solo para ICMP
Puerto de origen (sport)	Puerto de origen utilizado por la sesión
Puerto de destino (dport)	Puerto de destino utilizado por la sesión
Puerto de origen NAT (natsport)	Puerto de origen post-NAT
Puerto de destino NAT (natdport)	Puerto de destino post-NAT
Banderas (flags)	Campo de 32 bits que proporciona detalles sobre la sesión; este campo puede decodificarse AND-ing los valores con el valor registrado:

	• 0x2000 amenaza de inundación 0.0.0.0 0.0.0.0 • 0x80000000: la sesión tiene una captura de paquetes (PCAP) • 0x02000000: sesión IPv6 • 0x01000000: la sesión SSL se descifró (proxy SSL) • 0x00800000: se denegó la sesión a través del filtrado de URL • 0x00400000: la sesión tiene una traducción NAT realizada (NAT) • 0x00200000: la información del usuario para la sesión se capturó a través del portal cautivo (Portal cautivo) • 0x00080000: el valor X-Forward-For de un proxy está en el campo de usuario de origen • 0x00040000: el registro corresponde a una transacción dentro de una sesión de proxy http (transacción de proxy) • 0x00008000: la sesión es un acceso a la página de contenedor (página de contenedor) • 0x00002000: la sesión tiene una coincidencia temporal en una regla para el manejo implícito de la dependencia de la aplicación. Disponible en PAN-OS 5.0.0 y superior • 0x00000800: se utilizó el retorno simétrico para reenviar el tráfico para esta sesión
Protocolo (proto)	Protocolo IP asociado con la sesión
Acción (action)	Acción tomada para la sesión; los valores son alerta, permitir, denegar, descartar, descartar todos los paquetes, restablecer cliente, restablecer servidor, restablecer ambos, bloquear URL. • alerta: amenaza o URL detectada pero no bloqueada • permitir: alerta de detección de inundaciones • negar: el archivo está bloqueado • caída: la amenaza detectada y la sesión asociada se abandonó

	• reset-client: amenaza detectada y se envía un TCP RST al cliente • reset-server: amenaza detectada y se envía un TCP RST al servidor • reset-both: amenaza detectada y se envía un TCP RST tanto al cliente como al servidor • block-url: la solicitud de URL se bloqueó porque coincidía con una categoría de URL que se configuró para ser bloqueada • block-ip: amenaza detectada y IP del cliente bloqueado • caída aleatoria: se detectó una inundación y el paquete se descartó aleatoriamente • sumidero: sumidero DNS activado • syncookie-enviado: alerta de syncookie • bloquear-continuar (solo subtipo de URL): una solicitud HTTP se bloquea y se redirige a una página Continuar con un botón para confirmar que continúe • continuar (solo subtipo de URL): responda a una página de continuación de bloqueo de continuación de URL que indica que se permitió continuar con una solicitud de bloqueo de continuación • anulación de bloque (solo subtipo de URL): una solicitud HTTP se bloquea y se redirige a una página de anulación de administrador que requiere un código de acceso del administrador del firewall para continuar • override-lockout (solo subtipo de URL): demasiados intentos fallidos de contraseña de anulación de administrador desde la IP de origen y ahora está bloqueado desde la página de redireccionamiento de anulación de bloqueo • anular (solo subtipo de URL): responda a una página de anulación de bloque donde se proporciona un código de acceso correcto y se permite la solicitud
Misceláneo (misceláneo)	Campo con longitud variable con un máximo de 63 caracteres. Un nombre de archivo tiene un máximo de 63 caracteres. Una URL tiene un máximo de 1023 caracteres.

	• El URI real cuando el subtipo es url • Nombre de archivo o tipo de archivo cuando el subtipo es archivo • Nombre de archivo cuando el subtipo es virus • Nombre de archivo cuando el subtipo es wildfire-virus • Nombre de archivo cuando el subtipo es wildfire • URL o nombre de archivo cuando el subtipo es vulnerabilidad, si corresponde
ID de amenaza (amenaza)	Identificador de Palo Alto Networks para la amenaza. Es una cadena de descripción seguida de un identificador numérico de 64 bits entre paréntesis para algunos subtipos: • 8000 - 8099— detección de escaneo • 8500 - 8599— detección de inundaciones • 9999— Registro de filtrado de URL • 10000 - 19999: detección de teléfono en casa • 20000 - 29999: detección de descarga de software • 30000-44999: detección de vulnerabilidades • 52000 - 52999— detección de tipo de archivo • 60000 - 69999: detección de filtrado de datos • 100000 - 2999999 - detección de virus • 3000000 - 3999999 — Feed de firma WildFire • 4000000-4999999 —DNS Firmas de botnet
Categoría (categoría)	Para el subtipo de URL, es la categoría de URL; Para el subtipo WildFire, es el veredicto en el archivo y es 'malicioso', 'grayware' o 'benigno'; Para otros subtipos, el valor es 'cualquiera'.
Severidad (severidad)	Gravedad asociada con la amenaza; los valores son informativos, bajo, medio, alto, crítico
HipotesisDirección (dirección)	Indica la dirección del ataque, de cliente a servidor o de servidor a cliente:

	• 0: la dirección de la amenaza es del cliente al servidor • 1: la dirección de la amenaza es de servidor a cliente
Número de secuencia (seqno)	Un identificador de entrada de registro de 64 bits incrementado secuencialmente. Cada tipo de registro tiene un espacio numérico único. Este campo no es compatible con los firewalls de la serie PA-7000.
Banderas de acción (banderas de acción)	Un campo de bits que indica si el registro se reenvió a Panorama.
Ubicación de origen (srcloc)	País de origen o región interna para direcciones privadas. La longitud máxima es de 32 bytes.
Lugar de destino (dstloc)	País de destino o región interna para direcciones privadas. La longitud máxima es de 32 bytes.
Tipo de contenido (tipo de contenido)	Aplicable solo cuando Subtipo es URL. Tipo de contenido de los datos de respuesta HTTP. Longitud máxima 32 bytes.
ID de PCAP (pcap_id)	El ID de captura de paquetes (pcap) es una integral sin signo de 64 bits que denota un ID para correlacionar los archivos pcap de amenazas con pcaps extendidos tomados como parte de ese flujo. Todos los registros de amenazas contendrán un pcap_id de 0 (sin pcap asociado) o una ID que haga referencia al archivo pcap extendido.
Resumen de archivos (archivado)	Solo para el subtipo WildFire; todos los otros tipos no usan este campo. La cadena archivada muestra el hash binario del archivo enviado para ser analizado por el servicio WildFire.
Nube (nube)	Solo para el subtipo WildFire; todos los demás tipos no usan este campo.La cadena de la nube muestra el FQDN del dispositivo WildFire (privado) o de la nube WildFire (público) desde donde se cargó el archivo para su análisis.
Índice URL (url_idx)	Utilizado en Filtrado de URL y subtipos WildFire.Cuando una aplicación utiliza TCP keepalives para mantener una conexión abierta durante un período de tiempo prolongado, todas las entradas de registro para esa sesión tienen una única ID de sesión. En tales casos, cuando tiene un único

	registro de amenazas (e ID de sesión) que incluye múltiples entradas de URL, url_idx es un contador que le permite correlacionar el orden de cada entrada de registro dentro de la única sesión.Por ejemplo, para conocer la URL de un archivo que el firewall reenvió a WildFire para su análisis, ubique el ID de sesión y el url_idx en el registro de Envíos de WildFire y busque el mismo ID de sesión y url_idx en sus registros de filtrado de URL. La entrada de registro que coincide con el ID de sesión y url_idx contendrá la URL del archivo que se envió a WildFire.
Agente de usuario (user_agent)	Solo para el subtipo de filtrado de URL; todos los demás tipos no usan este campo.El campo Agente de usuario especifica el navegador web que el usuario utilizó para acceder a la URL, por ejemplo, Internet Explorer. Esta información se envía en la solicitud HTTP al servidor.
Tipo de archivo (tipo de archivo)	Solo para el subtipo WildFire; todos los demás tipos no usan este campo.Especifica el tipo de archivo que el firewall reenvió para el análisis WildFire.
X-Forward-For (xff)	Solo para el subtipo de filtrado de URL; todos los demás tipos no usan este campo.El campo X-Fordered-For en el encabezado HTTP contiene la dirección IP del usuario que solicitó la página web. Le permite identificar la dirección IP del usuario, lo cual es útil particularmente si tiene un servidor proxy en su red que reemplaza la dirección IP del usuario con su propia dirección en el campo de la dirección IP de origen del encabezado del paquete.
Árbitro (árbitro)	Solo para el subtipo de filtrado de URL; todos los demás tipos no usan este campo.El campo Referer en el encabezado HTTP contiene la URL de la página web que vincula al usuario a otra página web; es la fuente que redirigió (remitió) al usuario a la página web que se solicita.
Remitente (remitente)	Solo para el subtipo WildFire; todos los demás tipos no usan este campo.Especifica el nombre del remitente de un correo electrónico que WildFire determinó que era malicioso al analizar un enlace de correo electrónico reenviado por el firewall.
Asunto (sujeto)	Solo para el subtipo WildFire; todos los demás tipos no usan este campo.Especifica el asunto de un correo electrónico que WildFire determinó que era malicioso al analizar un enlace de correo electrónico reenviado por el firewall.

Destinatario (destinatario)	Solo para el subtipo WildFire; todos los demás tipos no usan este campo.Especifica el nombre del receptor de un correo electrónico que WildFire determinó que era malicioso al analizar un enlace de correo electrónico reenviado por el firewall.
ID de informe (reportid)	Solo para el subtipo WildFire; todos los demás tipos no usan este campo.Identifica la solicitud de análisis en la nube WildFire o el dispositivo WildFire.
Jerarquía de grupo de dispositivos (dg_hier_level_1 a dg_hier_level_4)	Una secuencia de números de identificación que indican la ubicación del grupo de dispositivos dentro de una jerarquía de grupos de dispositivos. El firewall (o sistema virtual) que genera el registro incluye el número de identificación de cada antepasado en su jerarquía de grupo de dispositivos. El grupo de dispositivos compartidos (nivel 0) no está incluido en esta estructura.Si los valores de registro son 12, 34, 45, 0, significa que el registro fue generado por un firewall (o sistema virtual) que pertenece al grupo de dispositivos 45, y sus antepasados son 34 y 12. Para ver los nombres de los grupos de dispositivos que corresponden al valor 12, 34 o 45, utilice uno de los siguientes métodos: Comando CLI en modo de configuración: muestre solo datos de metadatos de lectura API consulta: /api/?type=op&cmd=<show><dg-hierarchy></dg-hierarchy></show>
Nombre del sistema virtual (Porys_name)	El nombre del sistema virtual asociado con la sesión; solo válido en firewalls habilitados para múltiples sistemas virtuales.
Nombre del dispositivo (nombre_dispositivo)	El nombre de host del firewall en el que se registró la sesión.
Contentver	Version del constenido. Malware detectado en apps descargadas.

Fuente: https://docs.paloaltonetworks.com/pan-os/7-1/pan-os-admin/monitoring/syslog-field-descriptions.html#67983

2.1.2. Análisis exploratorio con herramientas de Minería de datos

Las empresas en la actualidad almacenan gran cantidad de información producto de sus transacciones y de acuerdo a sus necesidades para un posterior análisis y toma de decisiones a nivel gerencial, para el normal desarrollo de sus actividades.

Se crean entonces técnicas de almacenamiento y análisis de la información como la Minería de Datos, que facilita trabajar con grandes volúmenes de datos, procesarlos, analizarlos, asociarlos, agruparlos, describirlos y saber el comportamiento de las variables que la componen.

Según García y Acevedo (2010), la minería de datos se fundamenta en dos áreas del conocimiento: la estadística clásica para conocer la distribución estándar, la varianza, análisis de clustering pero no permiten identificar relaciones cualitativas entre los datos y de otro lado la inteligencia artificial, que procura aplicar procesamiento lógico a través de algoritmos genéticos, redes neuronales, árboles de decisión, entre otros, a diversos problemas estadísticos con el fin de descubrir información que se encuentra oculta en las bases de datos de la organizaciones.

Técnicas de modelado de minería de datos

Específicamente para nuestro caso, tomaremos las definidas en la investigación de Vallejo y Tenelanda (2012) que son:

a) Modelos de clasificación:

Utilizan el valor de uno o más campos de entrada para predecir el valor de uno o más resultados o campos de destino. Entre ellos tenemos a:

- **Modelo de árbol de clasificación y regresión C&R:** genera un árbol de decisión que pronostica o clasifica observaciones futuras. Un nodo se considera "puro" si el 100% de los casos del nodo corresponden a una categoría específica del campo objetivo. Los campos de entrada y objetivo pueden ser continuos (numéricos) o categóricos (nominal, ordinal o marca). Todas las divisiones son binarias (sólo crea dos subgrupos).
- **Modelo QUEST:** método de clasificación binario para generar árboles de decisión; está diseñado para reducir el tiempo de procesamiento de análisis de C&RT y reducir la tendencia de los métodos de clasificación de árboles para

favorecer a las entradas que permitan realizar más divisiones. Los campos de entrada pueden ser continuos (numéricos), pero el campo objetivo debe ser categórico. Todas las divisiones son binarias.

- **El modelo CHAID:** genera árboles de decisión utilizando estadísticos de chi-cuadrado para identificar las divisiones óptimas. A diferencia de C&RT y QUEST, este modelo puede generar árboles no binarios, significando que algunas divisiones generarán más de dos ramas. Los campos de entrada y objetivo pueden ser continuos (numéricos) o categóricos.
- **El modelo C5.0:** genera un árbol de decisión o un conjunto de reglas. Se divide la muestra basada en el campo con la máxima ganancia de información en cada nivel. El campo objetivo debe ser categórico. Se permiten varias divisiones en más de dos subgrupos.
- **Modelo lineal**: predicen un destino continuo tomando como base las relaciones lineales entre el destino y uno o más predictores.
- **La regresión lineal**: técnica estadística para resumir datos y realizar pronósticos ajustando una superficie o línea recta que minimice las discrepancias existentes entre los valores de salida reales y los pronosticados.
- **La regresión logística**: técnica estadística para clasificar los registros en función a los valores de los campos de entrada. Se diferencia de la anterior en tomar un campo objetivo categórico en lugar de uno numérico.
- **El modelo Regresión de Cox**: crea un modelo de supervivencia para datos de tiempo hasta el evento en presencia de registros censurados. Este modelo produce una función de supervivencia que pronostica la probabilidad de que el evento de interés se haya producido en el momento dado (t) para valores determinados de las variables de entrada.
- **Modelo Red bayesiana**: crea un modelo de probabilidad combinando pruebas observadas y registradas con conocimiento del mundo real para establecer la probabilidad de instancias. El nodo se centra en las redes Naïve Bayes aumentado a árbol (TAN) y de cadena de Markov que se utilizan para la clasificación.

b) Modelos de asociación

Estos modelos encuentran patrones en los datos en los que una o más entidades, se asocian con una o más entidades. Los modelos construyen conjuntos

de reglas que definen estas relaciones. Aquí los campos de los datos pueden funcionar como entradas y destinos. Los algoritmos de reglas de asociaciones realizan muy rápido el trabajo y pueden explorar patrones más complejos. Los modelos inducción de reglas generalizado son:

- **Inducción de reglas generalizado GRI:** capaz de encontrar las reglas de asociación existentes en los datos.
- **El modelo A priori:** extrae un conjunto de reglas de los datos y destaca aquellas reglas con un mayor contenido de información. A priori ofrece cinco métodos diferentes para la selección de reglas y utiliza un sofisticado esquema de indización para procesar eficientemente grandes conjuntos de datos.
- **El modelo CARMA**: extrae un conjunto de reglas de los datos sin necesidad de especificar campos de entrada ni de objetivo, ofrece configuraciones de generación basadas en el soporte de las reglas (soporte para el antecedente y el consecuente).
- **El modelo Secuencia:** encuentra reglas de asociación en datos secuenciales o en datos ordenados en el tiempo. Se basa en el algoritmo de reglas de asociación de CARMA, que utiliza un método de dos pasos para encontrar las secuencias.

c) Modelos de segmentación

Dividen los datos en segmentos o conglomerados de registros que tienen patrones similares de campos de entrada, entonces no contemplan el concepto de campos de salida o destino. Aquí podemos citar a:

- **El modelo K-medias**: agrupa conjuntos de datos en grupos distintos (o conglomerados). El método define un número fijo de conglomerados a los que asigna los registros de forma iterativa y ajusta los centros hasta que no se pueda mejorar el modelo. En lugar de intentar pronosticar un resultado, los modelos k-medias utilizan el proceso de aprendizaje no supervisado para revelar los patrones del conjunto de campos de entrada.
- **El modelo Kohonen**: genera un tipo de red neuronal que se puede usar para conglomerar un conjunto de datos en grupos distintos. Cuando la red se termina de entrenar, los registros que son similares se deberían presentar juntos en el mapa de resultados, mientras si los registros son diferentes aparecerán aparte.

- **El modelo Bietápico**: es un método de conglomerado de dos pasos. Primero, se hace una única pasada por los datos para comprimir los datos de entrada de la fila en un conjunto de sub-conglomerados administrable. El segundo paso utiliza un método de conglomerado jerárquico para fundir progresivamente los subconglomerados en conglomerados cada vez más grandes. El bietápico tiene la ventaja de estimar automáticamente el número óptimo de conglomerados para los datos de entrenamiento. Puede tratar tipos de campos mixtos y grandes conjuntos de datos de manera eficaz.
- **El modelo Detección de anomalías:** identifica casos extraños, o valores atípicos, que no se ajustan a patrones de datos "normales".

Metodología CRIPS-DM

Según CRIPS-DM (2016), la metodología contempla el proceso de análisis de datos como un proyecto profesional, estableciendo así un contexto mucho más rico que influye en la elaboración de los modelos. Está conformada por 6 fases:

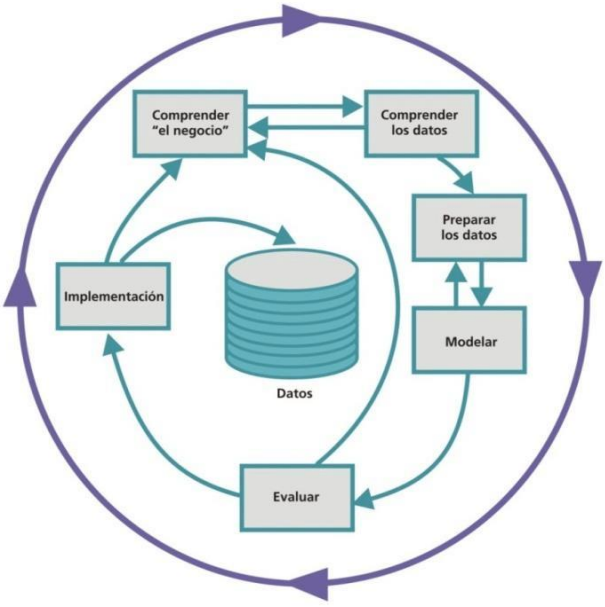

Gráfico 14. *Fases de la metodología CRISP-DM*

Fuente: CRIPS-DM (2016)

La secuencia de las fases no es rígida: se permite movimiento hacia adelante y hacia atrás entre diferentes fases. El resultado de cada fase determina qué fase, o qué tarea particular de una fase, hay que hacer después. Las flechas indican las dependencias más importantes y frecuentes.

El círculo externo en la figura simboliza la naturaleza cíclica de los proyectos de análisis de datos. El proyecto no se termina una vez que la solución se despliega. La información descubierta durante el proceso y la solución desplegada pueden producir nuevas iteraciones del modelo. Los procesos de análisis subsecuentes se beneficiarán de las experiencias previas.

Fase I. Definición de necesidades del cliente (comprensión del negocio)

Esta fase se enfoca en la comprensión de los objetivos de proyecto. Después se convierte este conocimiento de los datos en la definición de un problema de minería de datos y en un plan preliminar diseñado para alcanzar los objetivos.

Fase II. Estudio y comprensión de los datos

La fase de entendimiento de datos comienza con la colección de datos inicial y continúa con las actividades que permiten familiarizarse con los datos, identificar los problemas de calidad, descubrir conocimiento preliminar sobre los datos, y/o descubrir subconjuntos interesantes para formar hipótesis en cuanto a la información oculta.

Fase III. Análisis de los datos y selección de características

La fase de preparación de datos cubre todas las actividades necesarias para construir el conjunto final de datos (los datos que se utilizarán en las herramientas de modelado) a partir de los datos en bruto iniciales. Las tareas incluyen la selección de tablas, registros y atributos, así como la transformación y la limpieza de datos para las herramientas que modelan.

Fase IV. Modelado

En esta fase, se seleccionan y aplican las técnicas de modelado que sean pertinentes al problema (cuantas más mejor), y se calibran sus parámetros a valores óptimos. Típicamente hay varias técnicas para el mismo tipo de problema de minería de datos. Algunas técnicas tienen requerimientos específicos sobre la forma de los datos. Por lo tanto, casi siempre en cualquier proyecto se acaba volviendo a la fase de preparación de datos.

Fase V. Evaluación (obtención de resultados)

En esta etapa en el proyecto, se han construido uno o varios modelos que parecen alcanzar calidad suficiente desde la perspectiva de análisis de datos.

Antes de proceder al despliegue final del modelo, es importante evaluarlo a fondo y revisar los pasos ejecutados para crearlo, comparar el modelo obtenido con los objetivos de negocio. Un objetivo clave es determinar si hay alguna cuestión importante de negocio que no haya sido considerada suficientemente. Al final de esta

fase, se debería obtener una decisión sobre la aplicación de los resultados del proceso de análisis de datos.

Fase VI. Despliegue (puesta en producción)

Generalmente, la creación del modelo no es el final del proyecto. Incluso si el objetivo del modelo es de aumentar el conocimiento de los datos, el conocimiento obtenido tendrá que organizarse y presentarse para que el cliente pueda usarlo. Dependiendo de los requisitos, la fase de desarrollo puede ser tan simple como la generación de un informe o tan compleja como la realización periódica y quizás automatizada de un proceso de análisis de datos en la organización.

La siguiente figura presenta una guía visual de todas las fases, listando las tareas a realizar en cada fase, así como las conexiones entre ellas y las iteraciones que pueden llevarse a cabo:

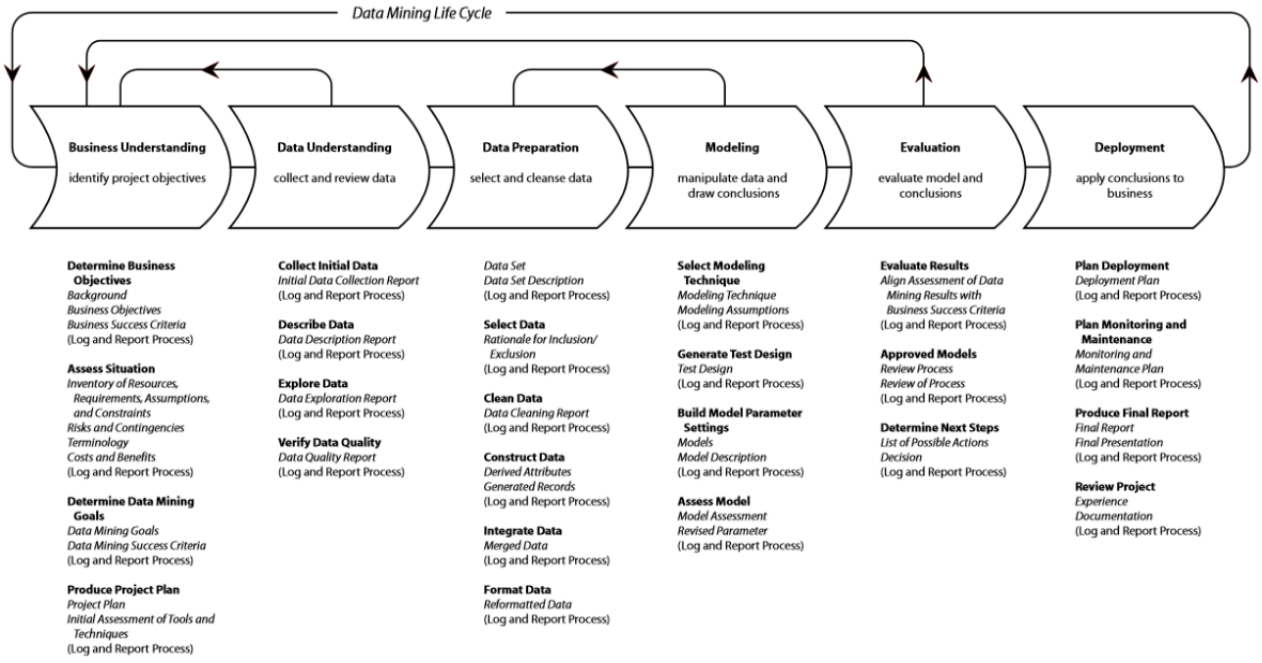

Gráfico 15. *Guía visual de la Metodología CRISP-DM*

Fuente: CRIPS-DM (2016)

2.1.3. Red de área local inalámbrica (WLAN)

Las redes de área local inalámbricas tienen alcance geográfico limitado, el IEEE define el estándar 802.11 para regular las WLANs. Se pueden utilizar junto a las redes cableadas, haciendo ésta ultima de backbone, también como red de interconexión entre diferentes redes cableadas o como extensión de éstas. Las características de

flexibilidad, escalabilidad y facilidad de implementación de las redes WLAN, están ganando espacio en el mercado de consumo, junto a sus precios populares y a una serie de posibilidades especialmente en movilidad e integración de los servicios, que viene logrando la satisfacción de los clientes, gracias a los avances de la microelectrónica y la miniaturización de los circuitos integrados (Castillo, 2002).

Sin embargo, estas ventajas, asociadas a muchas otras, traen consigo el gran inconveniente en cuanto a sus vulnerabilidades de seguridad y privacidad.

A pesar de las nuevas tendencias de comunicaciones que estamos empezando a utilizar con el llamado IoT (internet de las cosas), según Harald (2010) indica que "entre los grandes retos en esta área se encuentran: el desarrollo de tecnología que garantice la privacidad de conjuntos de dispositivos heterogéneos y la seguridad y confiabilidad en la "nube"; la creación de modelos para la autenticación descentralizada; el desarrollo de tecnologías de protección de datos y de encriptado eficiente energéticamente, así como de tecnologías para la autenticación de redes y objetos; y la creación de mecanismos de anonimato"; podemos observar esta falencia.

Seguridad en las WLAN

Para Pascual (2007), Cruz (2004) y TSB (2006), los departamentos encargados de la seguridad de las redes, se enfrentan a diario al reto de mantener las redes inalámbricas de forma rentable y con niveles de rendimiento óptimo, garantizando la confidencialidad, integridad, disponibilidad, protegiendo los activos digitales frente a posibles robos o abusos, siempre dentro del marco de las normas gubernamentales y económicas correspondientes.

Por los años 90's el uso de las herramientas de ataque requería mucha experticia y conocimiento por parte de los intrusos, actualmente las amenazas son cada vez más sofisticadas a medida que disminuye el conocimiento técnico necesario para implementar ataques, existe un mayor acceso a las herramientas para determinar las debilidades de los sistemas y explotarlas y obtener los privilegios necesarios para realizar cualquier acción dañina. A todo esto, hay que sumar los problemas de configuración, y la falta de recursos para instalar los parches de seguridad necesarios.

En definitiva, se hace patente la necesidad de concientizar y enseñar temas relativos a la seguridad informática.

Por lo tanto, nos encontramos con el uso de políticas y procedimientos de seguridad, como las técnicas de encriptación y los cortafuegos (firewalls). Si bien estos elementos,

aunque proveen una primera línea de defensa para asegurar los recursos de la red, deben ser complementados con herramientas que permitan monitorizar el comportamiento del tráfico y las actividades de los usuarios de la red. Es así que surgen los sistemas de detección de intrusos, como uno de las opciones para implementar seguridad.

27

2.1.4. Gestión de seguridad de la información

La información es el conjunto de datos que se originan, transmiten, se encuentran organizados, almacenados y le pertenecen a una organización.

La seguridad de la información, según la ISO 27001(2013), consiste en "garantizar la confidencialidad, integridad y disponibilidad de la información, así como los sistemas de tratamiento, dentro de una organización".

Hacer mención que la protección total no existe, por más de tener un presupuesto ilimitado; el objetivo es garantizar que los riesgos de la seguridad de la información sean conocidos, asumidos, gestionados y minimizados.

Para ello es conveniente seguir un plan o buenas prácticas especificadas en estándares internacionales; para nuestro caso específico, definiremos los controles especificados en la ISO 27002-2013 aprobado en la NTP-ISO/IEC 27002:2017, organizado en 14 dominios, 35 objetivos de control y 114 controles.

Para nuestro estudio solo se tomó los conceptos referidos al Control de Seguridad de Cifrado, Seguridad en la operativa y Seguridad en las telecomunicaciones (ISO 27000.2013)

Control 10: CIFRADO

El objetivo es garantizar un uso adecuado y eficaz de la criptografía para proteger la confidencialidad, la autenticidad y/o la integridad de la información.

10.1 Controles criptográficos

- Utilizar controles criptográficos para la protección de claves de acceso a sistemas, datos y servicios, para la transmisión de información clasificada y/o para el resguardo de aquella información relevante en atención a los resultados de la evaluación de riesgos realizada por la organización.
- Desarrollo de procedimientos y asignación de funciones respecto de la administración de claves, de la recuperación de información cifrada en caso

de pérdida, compromiso o daño de las claves y en cuanto al reemplazo de las claves de cifrado.

- El uso de algoritmos de cifrado (simétricos y/o asimétricos) y las longitudes de clave deberían ser revisadas periódicamente para aplicar las actualizaciones necesarias en atención a la seguridad requerida y los avances en técnicas de descifrado.
- Las firmas digitales proporcionan un medio de protección de la autenticidad e integridad de los documentos electrónicos y, en algunas ocasiones, podría ser necesario asesoramiento legal para establecer acuerdos especiales que respalden su uso.

Actividades de control

10.1.1 Política de uso de los controles criptográficos: desarrollar e implementar una política que regule el uso de controles criptográficos para la protección de la información.

10.1.2 Gestión de claves: desarrollar e implementar una política sobre el uso, la protección y el ciclo de vida de las claves criptográficas a través de todo su ciclo de vida.

Control 12: SEGURIDAD EN LA OPERATIVA

El objetivo es controlar la existencia de los procedimientos de operaciones y el desarrollo y mantenimiento de documentación actualizada relacionada.

12.1 Responsabilidades y procedimientos de operación

El objetivo es evitar el acceso físico no autorizado, los daños e interferencias a la información de la organización y las instalaciones de procesamiento de la información.

Actividades de control

12.1.1 Documentación de procedimientos de operación: Se deberían documentar los procedimientos operativos y dejar a disposición de todos los usuarios que los necesiten.

12.1.2 Gestión de cambios: Se deberían controlar los cambios que afectan a la seguridad de la información en la organización y procesos de negocio, las instalaciones y sistemas de procesamiento de información.

12.1.3 Gestión de capacidades: Se debería monitorear y ajustar el uso de los recursos junto a proyecciones necesarias de requisitos de capacidad en el futuro con el objetivo de garantizar el rendimiento adecuado en los sistemas.

12.1.4 Separación de entornos de desarrollo, prueba y producción: Los entornos de desarrollo, pruebas y operacionales deberían permanecer separados para reducir los riesgos de acceso o de cambios no autorizados en el entorno operacional.

12.2 Protección contra código malicioso

El objetivo es garantizar que la información y las instalaciones de procesamiento de información estén protegidas contra el malware.

Actividades de control

12.2.1 Controles contra el código malicioso: Se deberían implementar controles para la detección, prevención y recuperación ante afectaciones de malware en combinación con la concientización adecuada de los usuarios.

12.3 Copias de seguridad

El objetivo es alcanzar un grado de protección deseado contra la pérdida de datos.

Actividades de control

12.3.1 Copias de seguridad de la información: Se deberían realizar y pruebas regulares de las copias de la información, del software y de las imágenes del sistema en relación a una política de respaldo (Backup) convenida.

12.4 Registro de actividad y supervisión

El objetivo es registrar los eventos relacionados con la seguridad de la información y generar evidencias.

Actividades de control

12.4.1 Registro y gestión de eventos de actividad: Se deberían producir, mantener y revisar periódicamente los registros relacionados con eventos de actividad del usuario, excepciones, fallas y eventos de seguridad de la información.

12.4.2 Protección de los registros de información: Se debería proteger contra posibles alteraciones y accesos no autorizados la información de los registros.
12.4.3 Registros de actividad del administrador y operador del sistema: Se deberían registrar las actividades del administrador y del operador del sistema y los registros asociados se deberían proteger y revisar de manera regular.
12.4.4 Sincronización de relojes: Se deberían sincronizar los relojes de todos los sistemas de procesamiento de información pertinentes dentro de una organización o de un dominio de seguridad y en relación a una fuente de sincronización única de referencia.

12.5 Control del software en explotación

El objetivo es garantizar la integridad de los sistemas operacionales para la organización.

Actividades de control

12.5.1 Instalación del software en sistemas en producción: Se deberían implementar procedimientos para controlar la instalación de software en sistemas operacionales.

12.6 Gestión de la vulnerabilidad técnica

El objetivo es evitar la explotación de vulnerabilidades técnicas.

Actividades de control

12.6.1 Gestión de las vulnerabilidades técnicas: Se debería obtener información sobre las vulnerabilidades técnicas de los sistemas de información de manera oportuna para evaluar el grado de exposición de la organización y tomar las medidas necesarias para abordar los riesgos asociados.
12.6.2 Restricciones en la instalación de software: Se deberían establecer e implementar las reglas que rigen la instalación de software por parte de los usuarios.

12.7 Consideraciones de las auditorías de los sistemas de información

El objetivo es minimizar el impacto de actividades de auditoría en los sistemas operacionales.

Actividades de control

12.7.1 Controles de auditoría de los sistemas de información: Se deberían planificar y acordar los requisitos y las actividades de auditoría que involucran la verificación de los sistemas operacionales con el objetivo de minimizar las interrupciones en los procesos relacionados con el negocio.

Control 13. SEGURIDAD EN LAS TELECOMUNICACIONES

El objetivo es asegurar la protección de la información que se comunica por redes telemáticas y la protección de la infraestructura de soporte.

La gestión segura de las redes, la cual puede abarcar los límites organizacionales, requiere de la cuidadosa consideración del flujo de datos, implicaciones legales, monitoreo y protección.

La información confidencial que pasa a través de redes públicas suele requerir de controles adicionales de protección.

Los intercambios de información por parte de las organizaciones se deberían basar en una política formal de intercambio y en línea con los acuerdos de intercambio, y debiera cumplir con cualquier legislación relevante.

13.1 Gestión de la seguridad en las redes

El objetivo es evitar el acceso físico no autorizado, los daños e interferencias a la información de la organización y las instalaciones de procesamiento de la información.

Actividades de control

13.1.1 Controles de red: Se deberían administrar y controlar las redes para proteger la información en sistemas y aplicaciones.

13.1.2 Mecanismos de seguridad asociados a servicios en red: Se deberían identificar e incluir en los acuerdos de servicio (SLA) los mecanismos de seguridad, los niveles de servicio y los requisitos de administración de todos los servicios de red, independientemente de si estos servicios se entregan de manera interna o están externalizados.

13.1.3 Segregación de redes: Se deberían segregar las redes en función de los grupos de servicios, usuarios y sistemas de información.

13.2 Intercambio de información con partes externas

El objetivo es mantener la seguridad de la información que transfiere una organización internamente o con entidades externas.

Políticas y procedimientos de intercambio de información: Deberían existir políticas, procedimientos y controles formales de transferencia para proteger la información que viaja a través del uso de todo tipo de instalaciones de comunicación.

Actividades de control

13.2.2 Acuerdos de intercambio: Los acuerdos deberían abordar la transferencia segura de información comercial entre la organización y las partes externas. Tener canales de comunicaciones alternativos y "pre-autorizados", en especial direcciones de e-mail secundarias por si fallan las primarias o el servidor de correo, y comunicaciones offline por si caen las redes.

13.2.3 Mensajería electrónica: Se debería proteger adecuadamente la información referida en la mensajería electrónica.

13.2.4 Acuerdos de confidencialidad y secreto: se deberían identificar, revisar y documentar de manera regular los requisitos para los acuerdos de confidencialidad y "no divulgación" que reflejan las necesidades de la organización para la protección de información.

Tabla 2.
Resumen de Controles de Seguridad ISO 27002-2013

DOMINIO	**OBJETIVOS DE CONTROL**	
	OBJETIVO DE CONTROL	**ACTIVIDAD DE CONTROL DE RIESGO**
CIFRADO	**10**	
	10.1	**Controles criptográficos**
	10.1.1	Política de uso de los controles criptográficos
	10.1.2	Gestión de claves
SEGURIDAD EN LA OPERATIVA	**12**	
	12.1	**Controles criptográficos**
	12.1.1	Documentación de procedimientos de operación
	12.1.2	Gestión de cambios.
	12.1.3	Gestión de capacidades.
	12.1.4	Separación de entornos de desarrollo, prueba y producción.
	12.2	**Protección contra código malicioso.**
	12.1.1	Controles contra el código malicioso.

DOMINIO	OBJETIVOS DE CONTROL	
	OBJETIVO DE CONTROL	**ACTIVIDAD DE CONTROL DE RIESGO**
	12.3	**Copias de seguridad**
	12.3.1	Copias de seguridad de la información
	12.4	**Registro de actividad y supervisión**
	12.4.1	Registro y gestión de eventos de actividad.
	12.4.2	Protección de los registros de información.
	12.4.3	Registros de actividad del administrador y operador del sistema.
	12.4.4	Sincronización de relojes
	12.5	**Control del software en explotación**
	12.5.1	Instalación del software en sistemas en producción
	12.6	**Gestión de la vulnerabilidad técnica**
	12.6.1	Gestión de las vulnerabilidades técnicas
	12.6.2	Restricciones en la instalación de software
	12.7	**Consideraciones de las auditorías de los sistemas de información**
	12.7.1	Controles de auditoría de los sistemas de información
SEGURIDAD EN LAS TELECOMUNICACIONES	**13**	
	13.1	**Gestión de la seguridad en las redes**
	13.1.1	Controles de red
	13.1.2	Mecanismos de seguridad asociados a servicios en red
	13.1.3	Segregación de redes
	13.2	**Protección contra código malicioso.**
	13.2.1	Políticas y procedimientos de intercambio de información
	13.2.2	Acuerdos de intercambio.
	13.2.3	Mensajería electrónica
	13.2.4	Acuerdos de confidencialidad y secreto

2.1.5 Definición de términos básicos

Según Ebel et al (2015), definen algunos términos muy utilizados en seguridad informática, que se detallan a continuación:

- **Ataque informático**

  O ciberataque es un método por el cual un individuo, mediante un sistema informático, intenta tomar el control, desestabilizar o dañar otro sistema informático (ordenador, red privada, etc).

- **Confidencialidad**

  Es la propiedad de la información, por la que se garantiza que está accesible únicamente a personal autorizado a acceder a dicha información

- **Integridad**

  Se refiere a la correcta y completa información en una base de datos.

- **Disponibilidad.**

  Una vez que la información ha sido capturada en un sistema de cómputo, debe ser almacenada de manera segura y estar disponible para los usuarios cuando la necesiten

- **No repudio**

  Un estado de negocios donde el supuesto autor de una declaración no es capaz de desafiar con éxito la validez de declaración o contrato

- **DHCP**

  Protocolo de control de host dinámico, permite la obtención de una dirección IP de forma dinámica, muy usado en redes inalámbricas

- **Gestión**

  Tareas de "despliegue, integración y coordinación del hardware, software y los elementos humanos para monitorizar, probar, sondear, configurar, analizar, evaluar y controlar los recursos "de una red.

- **Hackers Universitarios**

  Grupo de estudiantes que comparten el conocimiento sobre el funcionamiento de los sistemas informáticos, datos y redes que cuentan con programadores con altos conocimientos, especialistas en tecnología y redes.

- **IDS**

  Sistema de detección de intrusos. Es un programa de detección de accesos no autorizados a un computador o a una red. El IDS suele tener sensores virtuales con los que el núcleo del IDS puede obtener datos externos

- **IPS**

  Sistema de prevención de intrusos. es un software que ejerce el control de acceso en una red informática para proteger a los sistemas computacionales de ataques y abusos

- **ISO**

  International Organization for Standardization

- **Minería de datos**

  La minería de datos o exploración de datos es un campo de la estadística y las ciencias de la computación referido al proceso que intenta descubrir patrones en grandes volúmenes de conjuntos de datos. Utiliza los métodos de la inteligencia artificial, aprendizaje automático, estadística y sistemas de bases de datos

- **Pentesting**

  Son una práctica para poner a prueba un sistema informático, red o aplicación web para encontrar vulnerabilidades que un atacante podría explotar.

- **Sistema informático**

  Conjunto de programas, instrucciones y reglas informáticas para ejecutar ciertas tareas en una computadora.

- **TCP/IP**

  Modelo de referencia para entender el proceso de viaje de la información desde el nodo origen hasta el destino, que lleva el nombre de los protocolos Transfer Control Protocol e Internet Protocol.

- **WLAN**

  Significa Wireless Local Area Network. Sistema de comunicación inalámbrico para minimizar las conexiones cableadas.

- **Wardriving**

  Técnica que consiste en buscar redes inalámbricas Wi-Fi desde un vehículo en movimiento.

- **Warchalking**

  Búsqueda de redes inalámbricas WIFI, de forma que puedan ser utilizadas por aquellos que 'pasen por allí´.

2.2 ASPECTOS DE RESPONSABILIDAD SOCIAL

El rendimiento del WiFi continúa mejorando y hoy en día es una de las tecnologías de comunicación inalámbrica más omnipresentes. Su instalación es fácil, el uso es sencillo y, además, es barato. Hoy en día, hay puntos de acceso WiFi en casas, oficinas, universidades y en puntos públicos de acceso, lo que permite conectarse cómodamente a Internet desde portátiles y con más frecuencia de teléfonos móviles.

Pero su uso masificado, ha traído consigo numerosos problemas de seguridad en las organizaciones, al ser un vector de ataque muy importante y aprovechado por los ciberdelincuentes. Esta necesidad ha obligado a los administradores de las redes de datos a mejorar la protección de sus redes aplicando diversas técnicas y métodos para hacerlos menos vulnerables, pero de otro lado está también la conciencia que debe tener el usuario para disfrutar de este servicio sin perpetrar ataques, muchas veces realizados solo por curiosidad o juego. Para nuestra investigación implica completar a través de la tecnología y los servicios que se ofrecen, las medidas de protección adecuadas para que los usuarios estén verdaderamente protegidos cuando utilicen la infraestructura del campus universitario y al mismo tiempo estén prevenidos sobre los riesgos existentes en Internet.

III. MÉTODO

3.1. TIPO INVESTIGACIÓN

3.1.1. Tipo de Investigación

Aplicada

La investigación es aplicada ya que, con las bases teóricas, se pretende capturar el tráfico con posibles amenazas y ataques de las redes en funcionamiento de las universidades para posteriormente con las herramientas de minería de datos desarrollar un análisis exploratorio de estos tráficos con la finalidad de mejorar la gestión de seguridad de las redes universitarias basado en estándares internacionales de gestión de seguridad de la información.

3.1.2. Nivel de Investigación

Descriptiva, porque se capturarán y caracterizarán las propiedades de los datos que ingresan a las redes en forma de tráfico malicioso y nos permite hacer predicciones en base al análisis exploratorio al que es sometido; es correlacional porque al clasificar, asociar y segmentar los diferentes tráficos se puede predecir futuros ataques y percibir las falencias en la gestión de seguridad de las redes WLAN universitarias.

3.1.3. Diseño de la investigación

Es no experimental, porque se observa tipos de ataques presentes en el tráfico en las diferentes capas del modelo OSI tal cual se obtuvieron, para después analizarlos de forma exploratoria, cómo se segmentan, como se asocian y como se clasifican; para determinar qué modelo permite hacer predicciones y mejorar la mejor gestión de seguridad informática.

3.2. POBLACIÓN Y MUESTRA

3.2.1. Población

La población de la investigación está conformada por:

- ✓ El tráfico capturado con diferentes vulnerabilidades y tipos de ataques.
- ✓ Todo el personal del departamento de Seguridad de Tecnologías de la Información a quienes se les presentará el resultado del análisis exploratorio a fin de proponer medidas correctivas de mejora para la gestión de seguridad de las redes informáticas de la Universidad.

3.2.2. Muestra

No se realizó un muestreo porque se trabajó con la totalidad de la población:

- ✓ 8441 registros de ataques
- ✓ 11 especialistas.

3.3. OPERACIONALIZACIÓN DE VARIABLES

Variable Independiente:

Análisis exploratorio de ataques informáticos aplicando herramientas de minería de datos.

Los ataques informáticos son acciones que vulneran las medidas de protección de la infraestructura computacional de las organizaciones y ponen el riesgo la confidencialidad, integridad y disponibilidad de la información contenida, la misma que es registrada para su análisis off line, creando grandes volúmenes de información que son procesadas y analizadas por herramientas de minería de datos para encontrar información oculta o implícita estructurada o no estructuradas, que no es posible procesarla con métodos estadísticos comunes.

Variable Dependiente:

Gestión de seguridad de las redes inalámbricas.

Políticas definidas y controles de seguridad implementados para garantizar la confidencialidad, integridad y disponibilidad de la información que está contenida o viaja a través de las redes de comunicaciones, basada en estándares internacionales

Operacionalización de variables

VARIABLES	INDICADORES
Variable Independiente *Análisis exploratorio de ataques informáticos aplicando herramientas de minería de datos*	- Número de ataques por categoría - Temporalidad del ataque - Número de ataques por dispositivo - Frecuencia de ataques por país de procedencia - Frecuencia de ataques por país de destino - Número de incidencias por mes - Incidencia de zona de ataque - Severidad de la amenaza - Numero de repeticiones de acceso - Ataques por zona de origen - Ataques por IP origen - Ataques por puerto origen - Aplicaciones atacadas - Tipo de acción de defensa ejecutada
Variable Dependiente *Gestión de seguridad de las redes inalámbricas*	- Uso de protocolos de cifrados autorizados. - Uso de contraseñas para acceso remoto. - Identificación de acceso de usuarios de la red. - Política de cambio de contraseñas. - Desactivación de credenciales de trabajadores retirados. - Existencia de procedimientos frente a problemas informáticos. - Periodo de tiempo y cifrado de respaldo de la información. - Aplicación de parches de seguridad. - Registro, análisis y seguimiento de logs. - Plan de instalaciones de software. - Uso de servicios informáticos - Control de detección y prevención de códigos maliciosos. - Análisis de tráfico, auditoría y monitoreo de sistemas. - Plan de seguridad informática utilizando ISO 27002.

3.4. INSTRUMENTOS

Técnicas

- **Análisis documental**: Se usará esta técnica, para recoger la información de los diversos tráficos que atraviesan las redes de datos a través de herramientas de captura, los mismos que serán analizados por herramientas de minería de datos con el fin de detectar las características que ponen en riesgo la confidencialidad, integridad, y disponibilidad de los datos.
- **Encuesta**: Se usará esta técnica, para evaluar la gestión que realizan los Administradores y Oficiales de Seguridad de TI de las universidades en relación a los controles de seguridad aplicados en base a estándares ISO 27002(2013).

Instrumentos

- **Fichas de registro de reporte obtenido del firewall**

  Que permitirá obtener los registros de las capturas de tráfico de ataques, logs y DMZ (zona desmilitarizada), orígenes, destinos, puertos, acciones de defensa, nivel de riesgo y otros datos de la universidad.

Ficha de registro de reporte obtenido del firewall

Autor	
Nombre	
Software de computadora	
Versión	
Lugar	
Fecha de registro	
Fecha recepción sesión	
Hora recepción sesión	
Tipo de amenaza	
Sub tipo amenaza	
Hora_ataque	
IP_origen	
IP_destino	
Regla_de_sesion	

Aplicación_asociada a_sesion	
Zona_origen_sesion	
Zona_destino_sesion	
Fecha_hora_sesion	
Num_sesiones_misma_IP	
Puerto_origen	
Puerto_destino	
Detalle_ataque_hexadec	
Protocol IP	
Accion_para_sesion	
Identificador/descripción_amenaza	
Categoría_URL	
Severidad_amenaza	
Dirección_ataque	
País origen	
Pais_destino	
Categoría_amenaza	
Id_App_amenaza	

Fuente: Elaboración propia

- **Ficha de recolección de datos**

  Utilizado para conocer la percepción sobre la gestión de seguridad informática en aspectos de cifrado, seguridad en la operativa y gestión de seguridad definidos en la norma ISO 27002(2013), el mismo que fue aplicado a todo el personal del departamento de TI, entre administradores especialistas en seguridad informática, oficiales y jefes del área, en un numero de 11 personas. El instrumento se encuentra en el Anexo 2.

  Para la **validación del instrumento** se sometió al juicio de seis especialistas en Informática, Sistemas, Seguridad Informática y Ciencias de la Computación con amplia experiencia y grado de Doctor y Magister, quienes revisaron y evaluaron la ficha de recolección de datos en aspectos de claridad, objetividad, actualización, organización, suficiencia, intencionalidad, consistencia, coherencia, metodología y pertinencia, y consideraron en promedio el puntaje de

83.3, que recae en la valoración de Muy Bueno el instrumento utilizado. Los resultados del instrumento se encuentran en el Anexo 3

3.5. PROCEDIMIENTOS

La recopilación de datos se inició con la captura de tráfico de las redes informáticas de la Universidad, para lo cual se utilizó herramientas de captura de libre acceso como: Wireshark que básicamente permitió analizar los protocolos que circulan por las redes de la universidad; Sguil instalada en una máquina virtual con Security Onion, el que nos permitió auditar la seguridad de la red; así mismo se obtuvo la data del tráfico del mismo firewall Palo Alto con detalles más significantes para el presente estudio, data que permitió realizar el análisis.

Posteriormente para el análisis exploratorio con la data recabada, se siguió los pasos de la metodología CRISP-DM con el uso de las herramientas de minería de datos Weka y SPSS Modeler debido al gran volumen de información, el cual de detalla a continuación:

3.5.1 FASE I: Análisis o entendimiento del negocio

3.5.1.1 Determinar el contexto y objetivos del negocio

Las universidades tienen dentro de sus perspectivas, la implementación de políticas de aseguramiento de la calidad, que contemple la preparación científica y profesional con buenos servicios ofrecidos tanto a sus estudiantes, personal docente, administrativo, de servicios y demás stakeholders. Entre los servicios que se ofrece, se puede identificar al soporte informático a través de las redes de comunicaciones inalámbricas que se vienen implementando de forma masiva, estas redes son cada vez más utilizadas en aplicaciones a través de dispositivos móviles y su uso se debe a la flexibilidad, escalabilidad y facilidad de implementación.

3.5.1.2 Evaluación de la situación

El servicio de comunicación inalámbrica, muchas veces se ha visto disminuido por el inconveniente de la vulnerabilidad en temas de seguridad y cada día el interés del administrador de las redes es implementar mecanismos de seguridad que muchas veces no son eficientes, pues las habilidades de los atacantes han superado los límites y conocimientos del administrador de la red. Las intrusiones

que sufren diariamente nuestras redes de datos universitarias, son intentos que comprometen la confidencialidad, integridad, disponibilidad (CID) y que burlan los mecanismos de seguridad implementados en las redes.

3.5.1.3 Determinar objetivos de la minería de datos

Con el presente estudio, se busca

- Realizar la limpieza de los datos a través de técnicas estadísticas para determinar los campos significativos para realizar la minería de datos.
- Caracterizar los algoritmos de clasificación de datos para obtener niveles de confianza óptimos para utilizarlos en la predicción.
- Elegir la técnica predictiva de clasificación para procesarla con datos relevantes.

3.5.1.4 Realizar el plan de proyecto

El estudio se seguirá las siguientes tareas:

Tarea 1: Análisis del contenido de los datos. (1 semana)

Tarea 2: Realizar consultas para obtener muestras representativas de los datos (1 semana).

Tarea 3: Seleccionar campos, limpiar registros, formatear y convertir datos para evitar errores en la minería de datos (1 semana).

Tarea 4: Elegir las técnicas de modelado y probarlas sobre los datos (2 semanas)

Tarea 5: Análisis e interpretación de los resultados obtenidos (1 semana).

Tarea 6: Elaboración de informes con los resultados obtenidos en función a los objetivos planteados (1 semana).

Tarea 7: Presentación de resultados finales (1 semana).

3.5.2 FASE II: Comprensión de los datos

3.5.2.1 Recolectar los datos iniciales

Los datos utilizados para esta investigación fueron la captura de la data real desde el firewall Palo Alto de la Universidad, de la Zona desmilitarizada, logs de eventos y threats de todo el campus, agrupados en 8441 registros de ataques registrados desde el 26 de junio 2018 hasta el 1 de agosto 2019, que tiene los siguientes campos:

Fecha recepción sesión
Hora recepción sesión
Tipo de amenaza
Sub tipo amenaza
Hora_ataque
IP_origen
IP_destino
Regla_de_sesion
Aplicación_asociada a_sesion
Zona_origen_sesion
Zona_destino_sesion
Fecha_hora_sesion
Num_sesiones_misma_IP
Puerto_origen
Puerto_destino
Detalle_ataque_hexadec
Protocol IP
Accion_para_sesion
Identificador/descripcion_amenaza
Categoría_URL
Severidad_amenaza
Dirección_ataque
País origen
Pais_destino
Categoría_amenaza
Id_App_amenaza

Se realizará la predicción, en relación a la Categoría_amenaza.

3.5.2.2 Descripción de los datos

Fecha recepción sesión: Tipo fecha. Registro de la fecha de recepción de la sesión.
Hora recepción sesión: Tipo hora. Registro de la hora de recepción de la sesión.

Tipo de amenaza: Tipo alfanumérico. Es la clase de amenaza detectada, los valores pueden ser tráfico, amenaza, configuración, sistema y hip-match

Sub tipo amenaza: Tipo alfanumérico. Es el subtipo de amenaza detectada, los valores pueden ser datos, archivo, inundación, paquete, exploración, spyware, url, virus, vulnerabilidad, wildfire, virus wildfire.

Hora_ataque: Tipo hora. Hora a la que se generó el registro de datos del ataque.

IP_origen: Tipo alfanumérico. Dirección IP de origen de ataque.

IP_destino: Tipo alfanumérico. Dirección IP de destino del ataque.

Regla_de_sesion: Tipo alfanumérico. Nombre de la regla con la que coincidió la sesión

Aplicación_asociada_a_sesion: Tipo alfanumérico. Aplicación que se utilizó para generar el ataque.

Zona_origen_sesion: Tipo alfanumérico. Zona donde se originó el ataque.

Zona_destino_sesion: Tipo alfanumérico. Zona a la que estuvo dirigido el ataque.

Fecha_hora_sesion: Tipo fecha. Fecha y hora del registro del ataque.

Num_sesiones_misma_IP: Tipo numérico. Cantidad de intentos de conexión con la misma IP de origen, IP de destino, aplicación y subtipo vistas en 5 segundos.

Puerto_origen: Tipo numérico. Puerto de origen utilizado para la sesión.

Puerto_destino: Tipo numérico. Puerto destino utilizado para la sesión

Detalle_ataque_hexadec: Tipo numérico. Campo en formato hexadecimal, que devuelve un valor descriptivo del ataque propio del firewall, los valores pueden ser los indicados en la Tabla No. 1

Protocol IP: Tipo alfanumérico. Protocolo IP asociado con la sesión

Accion_para_sesión: Tipo alfanumérico. Es la acción que el firewall tomó para proteger la red interna, los valores pueden ser alerta, permitir, denegar, descartar, descartar todos los paquetes, restablecer cliente, restablecer servidor, restablecer ambos, bloquear URL.

Identificador/descripción_de_amenaza: Tipo alfanumérico. Contiene la descripción de la amenaza seguido entre paréntesis del ID de la amenaza, los valores van comprendidos entre 8000 y 4999999.

Categoría_URL: Tipo alfanumérico**.** Descripción de la página utilizada.

Severidad_amenaza: Tipo alfanumérico. Es la gravedad de la amenaza, los valores pueden ser informativos, bajo, medio, alto y crítico.

Dirección_ataque: Tipo alfanumérico: Que dirección tomó el ataque, los valores son 0 si fue de cliente a servidor y 1 si fue de servidor a cliente.

País origen: Tipo alfanumérico: País o región interna de origen del tráfico para direcciones privadas.

Pais_destino: Tipo alfanumérico. País o región interna de destino del tráfico para direcciones privadas.

Categoría_amenaza: Tipo alfanumérico. Nombre de la amenaza detectada, los valores pueden ser forcé-brute, code-execution, webshell, spyware, info-leak, net-worm, overflow, sql-injection, webshell y otras.

Id_App_amenaza: Tipo alfanumérico. Id del app utilizado en la amenaza.

3.5.2.3 Exploración de los datos

Para la exploración de los datos utilizaremos pruebas estadísticas básicas que muestran propiedades de los datos, sea crean tablas de frecuencia y gráficos de distribución de los datos. Este informe sirve particularmente para determinar la consistencia y completitud de los datos. A continuación, se muestran los mismos de acuerdo a los campos de la base de datos:

a) Hora de recepción de sesión

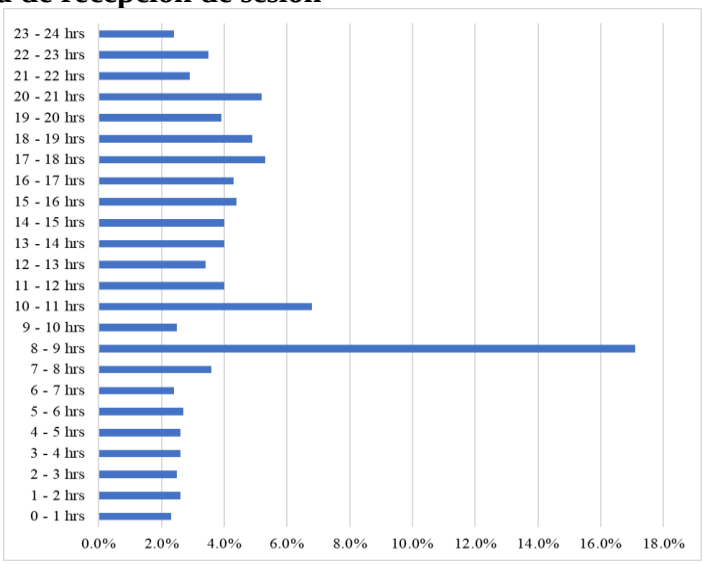

Gráfico 16. *Frecuencia de ataques informáticos por hora*
Fuente: Elaboración propia

La figura 16. Muestra la frecuencia de ataques informáticos clasificados por hora (considerando un rango de 24 horas), la misma que indica que durante el día de 8 a 9am se producen más interacciones, seguido de las horas de 10 a 11am. Esto puede concluir que a esa hora se tiene mayor cantidad de accesos por parte de la comunidad universitaria.

b) Fecha de recepción de sesión

Tabla 3.
Frecuencia de ataques informáticos clasificados por mes

	Frecuencia	Porcentaje (%)
Enero	70	0.8
Febrero	78	0.9
Marzo	298	3.5
Abril	1582	18.7
Mayo	398	4.7
Junio	948	11.2
Julio	1152	13.6
Agosto	1923	22.8
Setiembre	531	6.3
Octubre	486	5.8
Noviembre	213	2.5
Diciembre	762	9.0
Total	8441	100.0

Fuente: Elaboración propia

Se puede apreciar que la mayor cantidad de ataques se dieron durante los meses de julio y agosto, seguidos de abril. Esta cantidad puede indicar que durante esos meses se registra las matriculas a la universidad.

c) Subtipo de amenaza

Tabla 4.
Clasificación de ataques por el subtipo de amenaza

	Frecuencia	Porcentaje (%)
Vulnerabilidad	5650	66.9
Spyware	2791	33.1
Total	8441	100.0

Fuente: Elaboración propia

La clasificación por subtipo de amenaza indica que las capturas detectadas en el tráfico, indican que el 66.9% son vulnerabilidades.

d) Regla de sesión

Tabla 5.
Frecuencia de ataques según tipo de regla

	Frecuencia	Porcentaje (%)
Actualizaciones-1-1	2	0.0
RULE-LAN-WAN-LABORATOR	44	0.5
RULE-WIFI-CAMPUS	1492	17.7
RULE-WIFI-CAMPUS-PRUEBA	135	1.6
RULE-WIFI-CAMPUS-PRUEBA-1	903	10.7
VIP_Servidor_Biblioteca_HTTP	13	0.2
VIP_Servidor_TIC_MOQUEGUA	865	10.2
VIP_Servidor_Web_HTTP-1	25	0.3
VIP_Servidor_Web_HTTPS	3529	41.8
VIP_Servidor_Web_HTTPS-1	74	0.9
VIP_Servidor_Web_RADIO	642	7.6
VIP_Svr_Biblioteca_HTTP	717	8.5
Total	8441	100.0

Fuente: Elaboración propia

Las políticas de seguridad implementadas en el firewall universitario, están identificadas por reglas hacia servidores o dominios. Se aprecia que la regla que hace mayor cantidad de comparaciones recae en el Servidor_Web_HTTPS, seguido de las políticas hacia las redes inalámbricas.

e) Aplicación asociada a la sesión

Tabla 6.
Frecuencia de ataques según tipo de aplicación

	Frecuencia	Porcentaje (%)
ms-rdp	2	0.0
unknow-tcp	2319	27.5
web-browsing	5860	69.4
webdav	260	3.1
Total	8441	100.0

Fuente: Elaboración propia

Las aplicaciones que fueron utilizadas para iniciar la sesión de ataque fueron en 69.4% por web-browsing seguida de aplicaciones desconocidas, pero con conexión TCP, seguido de la Creación y control de versiones distribuidos en web (webdav); lo curioso es la baja interacción de remote-desktop-protocol propio de Microsoft.

f) Zona de origen de la sesión

Tabla 7.
Frecuencia de ataques según la zona de origen

	Frecuencia	Porcentaje (%)
LAN	2547	30.2
WAN	5867	69.5
WAN-WIFI	27	0.3
Total	8441	100.0

Fuente: Elaboración propia

Se puede indicar que las conexiones casi en un 70% vienen de fuera del campus, desde la zona de WAN, seguido de ataques desde la LAN interna, muy poco desde las redes inalámbricas externas.

g) Zona de destino de la sesión

Tabla 8.
Frecuencia de ataques según la zona de destino

	Frecuencia	Porcentaje (%)
LAN	29	0.3
WAN	1536	18.2
WAN-WIFI	1011	12.0
ZONA_DMZ1	5223	61.9
ZONA_DMZ2	642	7.6
Total	8441	100.0

Fuente: Elaboración propia

Podemos apreciar el que el objetivo es llegar a la DMZ1, zona de servidores públicos.

h) Número de sesiones de la misma IP

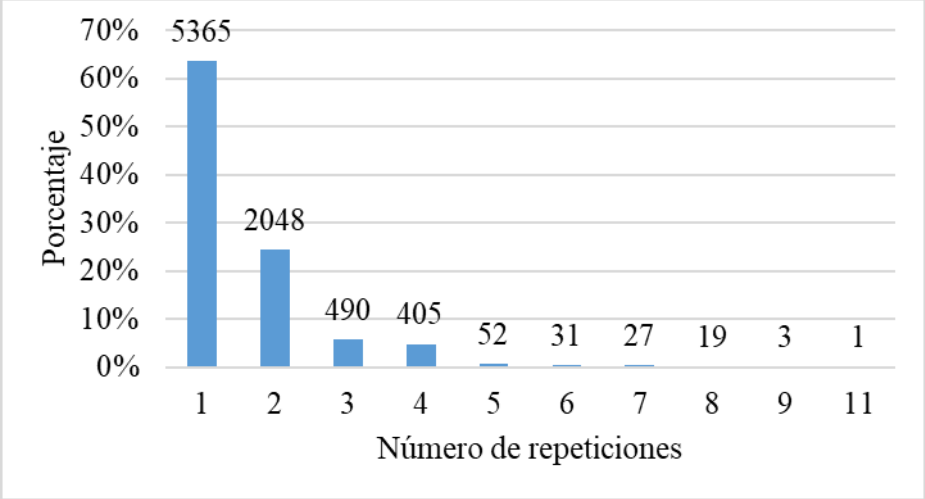

Gráfico 17. *Porcentaje de ataques desde la mima IP*

Fuente: Elaboración propia

Del total de ataques recibidos (8841), el intento de conexión desde la misma IP es única 5365 por primera vez, esto nos hace suponer que son robots programados para enviar ataques.

i) Detalle del ataque en hexadecimal (flags)

Tabla 9.
Frecuencia de ataques según detalle de la sesión

	Frecuencia	Porcentaje (%)
0x10200	616	7.3
0x1102000	46	0.5
0x2000	6741	79.9
0x422000	1038	12.3
Total	8441	100.0

Fuente: Elaboración propia

Los ataques de mayor frecuencia según la clasificación que hace el firewall son los 0x2000 que indican amenaza de inundación 0.0.0.0 0.0.0.0, típico ataque de denegación de servicio DoS o DDoS, es decir se recibió del total, casi el 80% de este tipo, seguido de inyección de malware 0x422000 con 12.3% el mismo que no es identificado por el firewall, significando un riesgo.

j) Acción ejecutada

Tabla 10.
Frecuencia de acciones ante ataques

	Frecuencia	Porcentaje (%)
Alert	2	0.0
reset-both	7976	94.5
reset-server	463	5.5
Total	8441	100.0

Fuente: Elaboración propia

La acción configurada en el firewall ante la detección de una amenaza en 94.5% es enviar un TCP RST tanto al cliente como al servidor; en

menor porcentaje 5.5% se envía un TCP RST sólo al servidor, el resto solo se alerta.

k) Frecuencia de amenazas según país de origen

Tabla 11.
Frecuencia de ataques según el país de origen

	Frecuencia	Porcentaje
Válido	3	0
Argentina	25	0.3
Armenia	1	0
Australia	2	0
Austria	3	0
Azerbaijan	1	0
Bangladesh	7	0.1
Botswana	1	0
Brazil	20	0.2
Bulgaria	4	0
Canada	19	0.2
Chile	14	0.2
China	1332	15.8
Colombia	13	0.2
Costa Rica	1	0
Denmark	2	0
Ecuador	44	0.5
Egypt	350	4.1
El Salvador	9	0.1
Ethiopia	2	0
France	45	0.5
French Guiana	2	0
Germany	26	0.3
Greece	10	0.1
Guinea	1	0
Haiti	1	0
Hong Kong	222	2.6
Hungary	5	0.1
Iceland	3	,0
India	20	,2
Indonesia	42	,5
INTERNO	4304	51,0
Iran Islamic Republic Of	2	,0
Ireland	3	,0
Israel	1	,0
Italy	88	1,0
Japan	56	,7
Jordan	1	,0
Kazakhstan	14	,2
Korea Republic	73	,9

	Frecuencia	Porcentaje
Of Malaysia	16	,2
Martinique	1	,0
Mexico	13	,2
Mongolia	1	,0
Nepal	3	,0
Netherlands	39	,5
Nigeria	10	,1
Norway	7	,1
Pakistan	3	,0
Peru	7	,1
Philippines	6	,1
Poland	16	,2
Romania	13	,2
RUnion	1	,0
Russian Federation	617	7,3
Saudi Arabia	1	,0
Serbia	3	,0
Seychelles	30	,4
Singapore	30	,4
Slovakia	1	,0
South Africa	9	,1
Spain	13	,2
Sri Lanka	1	,0
Sweden	4	,0
Taiwan ROC	68	,8
Thailand	36	,4
Trinidad And Tobago	1	,0
Tunisia	4	,0
Turkey	28	,3
Ukraine	12	,1
United Kingdom	160	1,9
United States	220	2,6
Venezuela Bolivarian Republic Of	1	,0
Viet Nam	292	3,5
Virgin Islands U	1	,0
Yemen	1	,0
Total	8441	100,0

Fuente: Elaboración propia

Se puede concluir que los ataques más frecuentes vienen desde la red interna de la Universidad, seguido de ataques externos, principalmente de China y seguidos de Rusia, Vietnam, Egipto, Estados Unidos y Hong Kong.

l) Frecuencia de amenazas por categoría

Tabla 12.
Frecuencia de acciones ante ataques

	Frecuencia	Porcentaje
Válido	3	,0
brute-force	28	,3
code-execution	5344	63,3
info-leak	10	,1
net-worm	193	2,3
overflow	262	3,1
spyware	2592	30,7
sql-injection	3	,0
webshell	6	,1
Total	8441	100,0

Fuente: Elaboración propia

Las amenazas más frecuentes, son la vinculadas al code-execution, seguido de spyware, lo que presume que se debe tomar alguna acción de defensa.

3.5.2.4 Verificar la calidad de los datos

Luego de la etapa de exploración de los datos, se determina que los datos están completos y no contienen errores ya que fueron generados por el firewall y exportados en formato *.pcap y *.xls. Por, requisito para introducirlo al Weka y SPSS Modeler y realizar un análisis más detallado para demostrar los objetivos que pretendo alcanzar.

Los datos no contienen valores fuera de rango, solo en un campo (Fecha_hora_recepción_sesión) juntos en el mismo campo, no hay riesgo de trabajar con datos basura durante la minería de datos.

Se encontró valores nulos en la captura del tráfico, lo que indica que la herramienta no registró ningún dato asociado a este parámetro, los

cuales fueron eliminados del archivo, así mismo se ignoró el campo por ser irrelevantes para el estudio final y otro no tomado en cuenta por ser repetitivo al momento de hacer el análisis exploratorio.

3.5.3 FASE III: Preparación de los datos

3.5.3.1 Seleccionar los datos más relevantes

De la tabla con datos capturados, se utilizó todos los registros (8441) de amenazas capturadas en el periodo de junio 2018-agosto 2019, puesto que todos los datos tienen importancia para el estudio. Sin embargo, hay campos que no tenían datos y por lo tanto no son necesarios para nuestros objetivos de minería de datos, los mismos que fueron eliminados. Además, estos campos no responden a los indicadores propuestos.

3.5.3.2 Limpiar los datos

La base de datos obtenida para la investigación contiene la información necesaria para medir los indicadores propuestos y cumplir los objetivos de la exploración de los datos, son datos que no contienen errores normalizados a través del software implementado en el firewall, por lo tanto, no es necesario hacer una limpieza exhaustiva de los mismos.

Más bien si se tuvo campos sin valores (blank), considerándose como datos faltantes, los mismos que fueron eliminados porque no aportan información adicional al estudio. Para realizar la asociación con Weka, se eliminó campos como fecha y hora, descripción de amenaza, categoría URL, severidad, dirección de ataque, acción tomada, detalle ataque en hex.

3.5.3.3 Construir los datos

Inicialmente se separó el Fecha_hora_recepción_sesión, que contenía la fecha y hora del registro de la amenaza en dos campos: Fecha y Hora.

Se codificó los datos, asignando un valor numérico sólo para efectos del análisis estadístico exploratorio, en los siguientes campos:

- Subtipo_amenaza: 1 (vulnerability), 2 (spyware).

- Aplicación_asociada_sesion:
    - 1 ms-rdp (Ms Remote Desktop Protocol)
    - 2 unknown-tcp
    - 3 web-browsing
    - 4 webdav "Creación y control de versiones distribuidos en web
- Zona origen de sesión:
    - 1 LAN
    - 2 WAN
    - 3 WAN WIFI
- Zona de destino de sesión
    - 1 LAN
    - 2 WAN
    - 3 WAN-WIFI
    - 4 Zona DMZ1
    - 5 Zona-DMZ
- Flags: o Detalle_ataque_hex
    - 1 0x10200
    - 2 0x1102000
    - 3 0x2000
    - 4 0x422000
- Acción ejecutada:
    - 1 alert
    - 2 reset-both
    - 3 reset-server
- Direccion_ataque:
    - 1 client-to-server
    - 2 server-to-client

3.5.3.4 Integrar los datos

Los nuevos valores descritos en el punto anterior, se integraron a la tabla de ataques a procesar.

3.5.4 FASE IV: Modelado

3.5.4.1 Escoger la técnica de modelado

Para realizar el modelado utilizaremos la herramienta de minería de datos WEKA 3.8, en donde se realiza la clasificación de los datos utilizando 7 de los principales algoritmos implementados en esta herramienta. Para la predicción se utiliza el SPSS Modeler 18.0.

3.5.4.2 Generar el plan de prueba

Se utilizó la herramienta Weka para determinar el algoritmo que mejor se adecúe a la predicción de futuros ataques informáticos en el campus universitario tomando como predictor la "categoría de amenaza". Se define el modelo que se adecúa, son los Arboles de decisión.

- **Algoritmo Arbol de decisión J48**

Tabla 13.
Clasificación de instancias Algoritmo Árbol de decisión J-48

	Cantidad	%
Instancias correctamente clasificadas	8393	99.4667 %
Instancias incorrectamente clasificadas	45	0.5333 %
TOTAL	**8438**	**100%**

Fuente: Elaboración propia

Tabla 14.
Matriz de confusión J48

code-execution	**spyware**	**sql-injection**	**overflow**	**info-leak**	**webshell**	**brute-force**	**net-worm**	**TOTAL**
5344	0	0	0	0	0	0	0	5344
0	2591	0	0	0	0	0	1	2592
3	0	0	0	0	0	0	0	3
2	0	0	260	0	0	0	0	262
10	0	0	0	0	0	0	0	10
0	1	0	0	0	5	0	0	6
28	0	0	0	0	0	0	0	28
0	0	0	0	0	0	0	193	193

Fuente: Elaboración propia

Interpretación: Como se puede apreciar en la Tabla 12, el algoritmo clasificó de manera correcta hasta un 99.4667% las instancias para la condición de tipos de amenazas y clasifica de manera incorrecta 0.5333% del total de las instancias. Así mismo, la Tabla 13 en la matriz de consistencia muestra que, de 8438 registros, 5344 fueron clasificados como "code-execution", representando la mayor cantidad de ataques dentro de la categoría amenazas de la base de datos capturada.

- **Algoritmo: Árbol de decisión Random Forest**

Tabla 15.
Clasificación de instancias Algoritmo Árbol de decisión Random Forest

	Cantidad	%
Instancias correctamente clasificadas	8432	99.9289 %
Instancias incorrectamente clasificadas	6	0.0711 %
TOTAL	**8438**	**100%**

Fuente: Elaboración propia

Tabla 16.
Matriz de confusión Random Forest

code-execution	**spyware**	**sql-injection**	**overflow**	**info-leak**	**webshell**	**brute-force**	**net-worm**	**TOTAL**
5344	0	0	0	0	0	0	0	5344
3	2588	0	0	0	0	0	1	2592
0	0	3	0	0	0	0	0	3
0	0	0	262	0	0	0	0	262
0	0	0	0	10	0	0	0	10
2	0	0	0	0	4	0	0	6
0	0	0	0	0	0	28	0	28
0	0	0	0	0	0	0	193	193

Fuente: Elaboración propia

Interpretación: En la Tabla 14, el algoritmo clasificó de manera correcta hasta un 99.9289 % las instancias para la condición de tipos de amenazas y clasifica de manera incorrecta 0.0711 % del total de las instancias. Así mismo, la matriz de consistencia muestra que, de 8438 registros, 5344 fueron clasificados solamente como "code-execution", representando la mayor cantidad de ataques dentro de la categoría amenazas de la base de datos capturada.

- **Algoritmo: Árbol de decisión DecisionStump**

Tabla 17.
Clasificación de *instancias Algoritmo Árbol de decisión DecisionStump*

	Cantidad	%
Instancias correctamente clasificadas	7936	94.0507 %
Instancias incorrectamente clasificadas	502	5.9493 %
TOTAL	**8438**	**100%**

Fuente: Elaboración propia

Tabla 18.
Matriz de confusión DecesionStump

code-execution	**spyware**	**sql-injection**	**overflow**	**info-leak**	**webshell**	**brute-force**	**net-worm**	**TOTAL**
5344	0	0	0	0	0	0	0	5344
0	2592	0	0	0	0	0	0	2592
3	0	0	0	0	0	0	0	3
262	0	0	0	0	0	0	0	262
10	0	0	0	0	0	0	0	10
0	6	0	0	0	0	0	0	6
28	0	0	0	0	0	0	0	28
0	193	0	0	0	0	0	0	193

Fuente: Elaboración propia

Interpretación: En las Tablas 16 y 17, el algoritmo clasificó de manera correcta hasta un 94.0507% las instancias para la condición de

tipos de amenazas y clasifica de manera incorrecta 5.9493 % del total de las instancias. Así mismo, la matriz de consistencia muestra que, de 8438 registros, 5344 fueron clasificados solo como "code-execution", representando nuevamente la mayor cantidad de ataques dentro de la categoría amenazas de la base de datos capturada.

- **Algoritmo: Árbol de decisión Hoeffding Tree**

Tabla 19.
Clasificación de instancias Algoritmo Árbol de decisión Hoeffding Tree

	Cantidad	%
Instancias correctamente clasificadas	6832	80.9671 %
Instancias incorrectamente clasificadas	1606	19.0329 %
TOTAL	**8438**	**100%**

Fuente: Elaboración propia

Tabla 20.
Matriz de confusión HoeffdingTree

code-execution	**spyware**	**sql-injection**	**overflow**	**info-leak**	**webshell**	**brute-force**	**net-worm**	**TOTAL**
3815	1	0	27	66	12	1423	0	5344
0	2542	4	26	0	14	4	2	2592
0	0	0	0	0	0	3	0	3
0	1	0	260	0	1	0	0	262
0	0	0	1	0	0	9	0	10
0	1	0	0	0	0	5	0	6
0	0	0	0	0	0	28	0	28
0	0	0	0	0	1	5	187	193

Fuente: Elaboración propia

Interpretación: En las Tablas 18 y 19, el algoritmo clasificó de manera correcta hasta un 80.9671% las instancias para la condición de 8 tipos de amenazas y clasifica de manera incorrecta 19.0329% del total de las instancias. Así mismo, la matriz de consistencia muestra que, de 8438 registros, 5344 fueron clasificados como "code-

execution" dispersas con brute-force, info-leak y overflow, representando nuevamente la mayor cantidad de ataques dentro de la categoría amenazas de la base de datos capturada.

- **Algoritmo: Árbol de decisión Random Tree**

Tabla 21.
Clasificación de instancias Algoritmo Árbol de decisión RandomTree

	Cantidad	%
Instancias correctamente clasificadas	8248	97.7483 %
Instancias incorrectamente clasificadas	190	2.2517 %
TOTAL	**8438**	**100%**

Fuente: Elaboración propia

Tabla 22.
Matriz de confusión RandomTree

code-execution	**spyware**	**sql-injection**	**overflow**	**info-leak**	**webshell**	**brute-force**	**net-worm**	**TOTAL**
5330	0	2	7	0	0	2	3	5344
14	2578	0	0	0	0	0	0	2592
2	0	1	0	0	0	0	0	3
91	0	0	163	0	0	0	8	262
4	0	0	0	6	0	0	0	10
2	0	0	0	0	4	0	0	6
5	0	0	0	0	0	23	0	28
46	0	0	4	0	0	0	143	193

Fuente: Elaboración propia

Interpretación: En las Tablas 20 y 21, el algoritmo clasificó de manera correcta hasta un 97.7483% las instancias para la condición de 8 tipos de amenazas y clasifica de manera incorrecta 2.2517% del total de las instancias. Así mismo, la matriz de consistencia muestra que, de 8438 registros, 5344 fueron clasificados como "code-execution" dispersas con overflow, net-worm, sql-injection y webshell; representa la mayor cantidad de ataques dentro de la categoría amenazas de la base de datos capturada, seguido de spyware.

- **Algoritmo: Árbol de decisión NaiveBayes**

Tabla 23.
Clasificación de instancias Algoritmo Árbol de decisión NaiveBayes

	Cantidad	%
Instancias correctamente clasificadas	8322	98.6253 %
Instancias incorrectamente clasificadas	116	1.3747 %
TOTAL	**8438**	**100%**

Fuente: Elaboración propia

Tabla 24.
Matriz de confusión NaiveBayes

code-execution	**spyware**	**sql-injection**	**overflow**	**info-leak**	**webshell**	**brute-force**	**net-worm**	**TOTAL**
5305	1	0	23	2	12	1	0	5344
0	2544	0	26	0	13	0	9	2592
3	0	0	0	0	0	0	0	3
0	1	0	260	0	1	0	0	262
8	0	0	1	1	0	0	0	10
0	1	0	0	0	0	0	5	6
9	0	0	0	0	0	19	0	28
0	0	0	0	0	0	0	193	193

Fuente: Elaboración propia

Interpretación: En las Tablas 22 y 23, el algoritmo clasificó de manera correcta hasta un 98.6253 % las instancias para la condición de 8 tipos de amenazas y clasifica de manera incorrecta 1.3747% del total de las instancias. Así mismo, la matriz de consistencia muestra que, de 8438 registros, 5344 fueron clasificados como “code-execution” dispersas con overflow y webshell principalmnete; representa la mayor cantidad de ataques dentro de la categoría amenazas de la base de datos capturada, seguido de spyware.

- **Algoritmo: Arbol de decisión REPTree**

Tabla 25.
Clasificación de instancias Algoritmo Árbol de decisión REPTree

	Cantidad	%
Instancias correctamente clasificadas	8433	99.9407 %
Instancias incorrectamente clasificadas	5	0.0593 %
TOTAL	**8438**	**100%**

Fuente: Elaboración propia

Tabla 26.
Matriz de confusión REPTree

code-execution	**Spyware**	**sql-injection**	**overflow**	**info-leak**	**webshell**	**brute-force**	**net-worm**	**TOTAL**
5344	0	0	0	0	0	0	0	5344
4	2588	0	0	0	0	0	0	2592
0	0	3	0	0	0	0	0	3
0	0	0	262	0	0	0	0	262
0	0	0	0	10	0	0	0	10
1	0	0	0	0	5	0	0	6
0	0	0	0	0	0	28	0	28
0	0	0	0	0	0	0	193	193

Fuente: Elaboración propia

Interpretación: En las Tablas 24 y 25, el algoritmo clasificó de manera correcta hasta un 99.9407% las instancias para la condición de 8 tipos de amenazas y clasifica de manera incorrecta 0.0593% del total de las instancias. Así mismo, la matriz de consistencia muestra que, de 8438 registros, 5344 fueron clasificados solamente como "code-execution"; representa la mayor cantidad de ataques dentro de la categoría amenazas de la base de datos capturada, seguido de spyware disperso con code-execution.

Podemos resumir el resultado de la clasificación para el campo de categoría de amenazas en la siguiente tabla.

Tabla 27.
Resumen de clasificación de los algoritmos para la Categoría de Amenazas.

	CATEGORIA DE AMENAZA		
	% predicción acertada	Instancias correctamente clasificadas	Instancias incorrectamente clasificadas
REPTree	99.94%	8433	5
RandomForest	99.93%	8432	6
J48	99.47%	8393	45
NaiveBayes	98.63%	8322	116
RandomTree	97.75%	8248	190
DecisionStump	94.05%	7936	502
HoeffdingTree	80.97%	6832	1606

Fuente: Elaboración propia

Entonces como se aprecia el mejor desempeño lo realiza el algoritmo de clasificación REPTree, por lo tanto, con este algoritmo se realiza la predicción de los futuros ataques informáticos.

3.5.4.3 Construir el modelo

Para construir el modelo utilizamos la herramienta Weka y ejecutamos el algoritmo REPTree, para cada campo de la base de datos, solamente para las amenazas

- code-execution
- spyware

3.5.4.4 Evaluar el modelo

Para obtener la precisión de la evaluación del modelo se utilizó el algoritmo REPTree para las amenazas indicadas, siendo el resultado el mostrado a continuación:

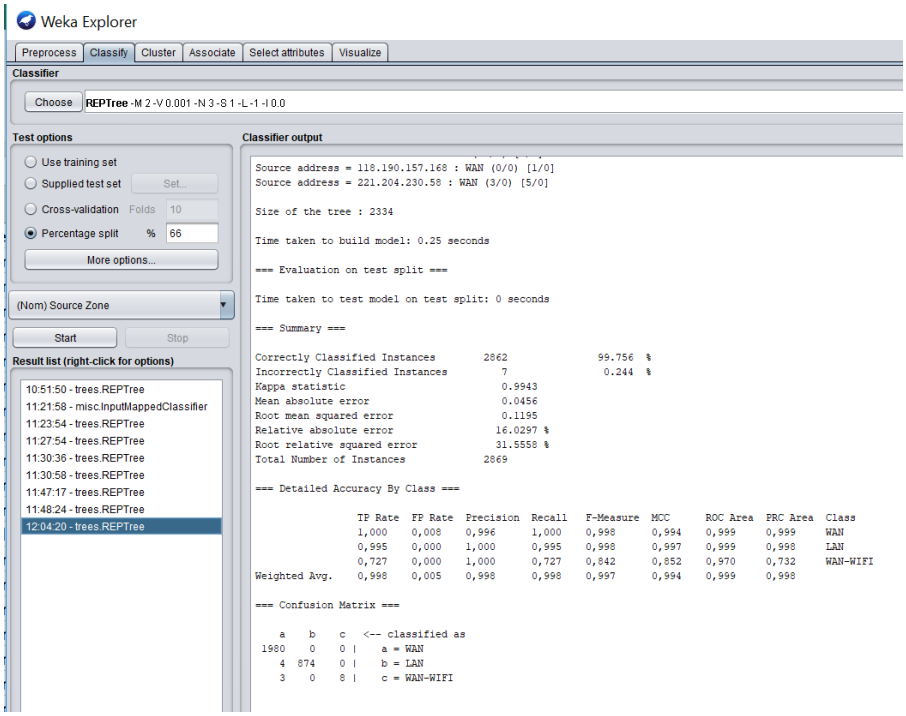

Gráfico 18. *Entrenamiento para Zona de orígen*

Fuente: Elaboración propia

Podemos concluir que según el entrenamiento realizado con el algoritmo REPTree, de un total de 2869 incidencias, 1980 ataques vienen de fuentes externas al campus universitario, seguido de ataques internos; de los cuales 5344 son de “code-execution” y 48 del tipo “spyware”.

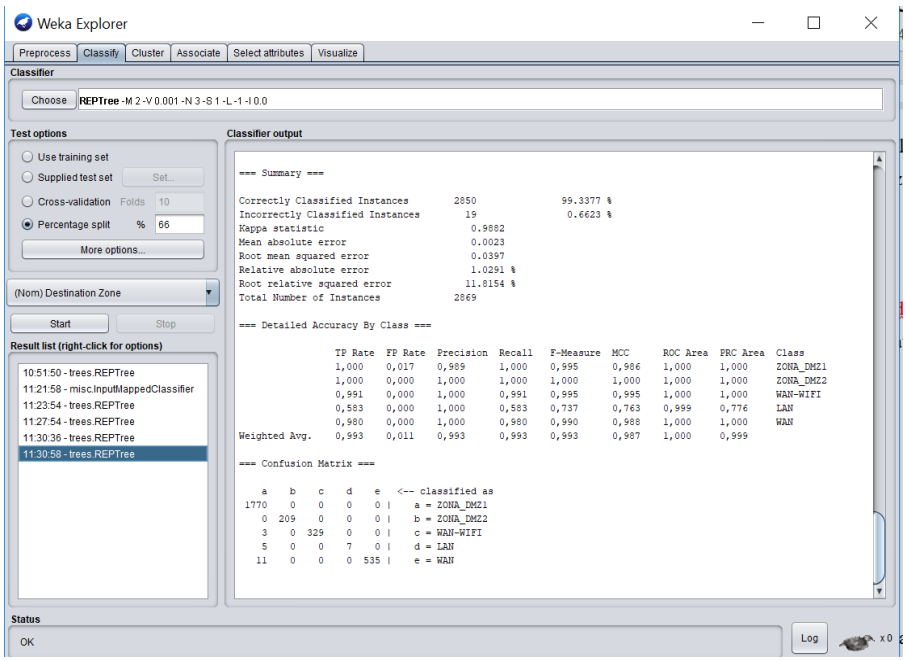

Gráfico 19. *Predicción para Zona de destino*

Fuente: Elaboración propia

Se puede indicar que según el entrenamiento realizado con el algoritmo REPTree, la ZONA_DMZ1 será la que tenga mayor incidencia de ataques (5223) del total, de los cuales 4992 serán de “code-execution” y 8 del tipo “spyware”.

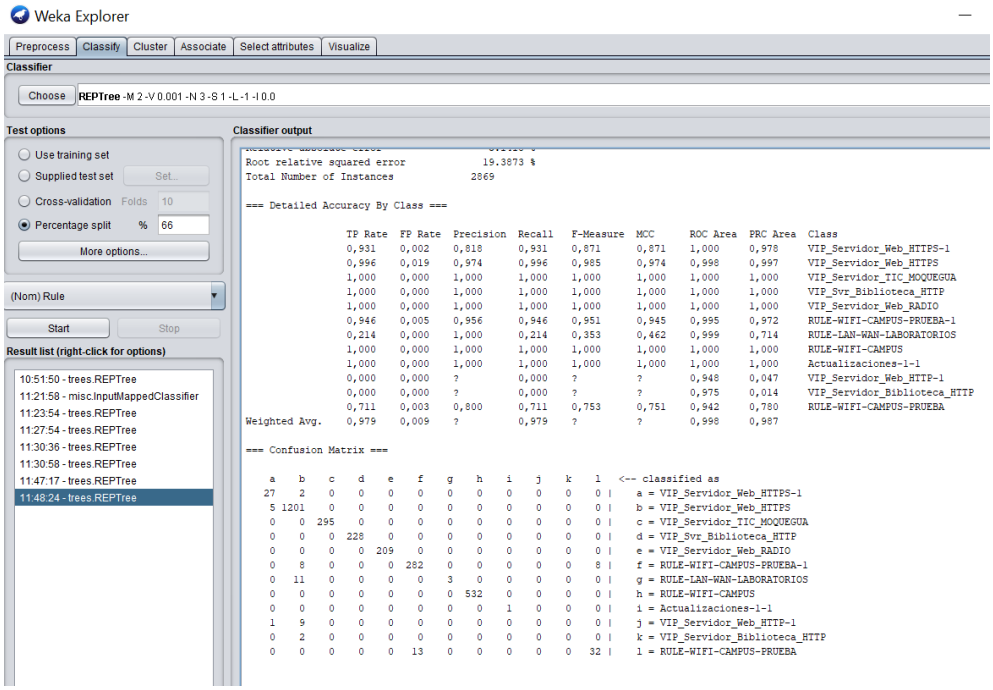

Gráfico 20. *Predicción para Regla de sesión*

Fuente: Elaboración propia

Se aprecia que según el entrenamiento realizado con el algoritmo REPTree, las reglas configuradas al Servidor_Web_HTTPS de la universidad es el que tienen más incidencia de ataques (3539) del total, de los cuales 3364 son de “code-execution” y 4 del tipo “spyware”. Sigue en el orden, los ataques que se predicen llegarán a las redes inalámbricas en conjunto WIFI_* con 2530 ataques, lo que nos indica que estas redes pueden ser de fácil acceso.

3.5.5 FASE V: Evaluación

3.5.5.1 Evaluar los resultados

Para indicar que son aceptables o no las predicciones indicadas por los modelos ejecutados, de hace necesario verificar los indicadores estadísticos que se han obtenido.

Predicción para Regla de sesión

Predicción acertada	97.9435 %
Predicción errada	2.0565 %
Error absolute medio	0.004
Error cuadrático medio	0.0487
Error absolute relative	3.1418 %
Error relative de raiz cuadrada	19.3873 %
Numero total de instancias	2869

Predicción para Zona origen

Predicción acertada	99.756%
Predicción errada	0.244%
Error absolute medio	0.0456
Error cuadrático medio	0.1195
Error absolute relative	16.0297 %
Error relative de raiz cuadrada	31.5558 %
Numero total de instancias	2869

Predicción para zona de destino

Predicción acertada	99.34%
Predicción errada	0.66%
Error absolute medio	0.0023
Error cuadrático medio	0.0397
Error absolute relative	1.03%
Error relative de raiz cuadrada	11.82%
Numero total de instancias	2869

Por lo tanto, en base a los indicadores de predicción acertadas podemos indicar que estos modelos nos ofrecen un nivel de confiabilidad aceptable.

3.5.6 FASE VI: Implantación

3.5.6.1 Planificación de la implantación

El resultado del análisis exploratorio, se pretende poner a consideración de los Jefes de Seguridad Informática y Administradores de TI de la Universidad, a quienes además se les propuso implementar políticas de control en bases a las normas de la ISO 27002(2013), ellos respondieron el instrumento (cuestionario) para validar la variable 2 del presente estudio global, el mismo que previo a la aplicación de la encuesta se realizó una prueba de validez por juicio de expertos.

3.5.6.2 Creación del informe final

Es el detallado a través de la metodología CRISP-DM y el resultado de la presente tesis

3.5.6.3 Revisión final del proyecto

En esta etapa se hace una evaluación de los pasos que se hizo correctamente y en las que se tuvo posibles fallas para hacer una corrección, el modelo permite retroalimentarse, para que en futuras pruebas con minería de datos se obtengan mejores resultados.

3.6. ANÁLISIS DE DATOS

3.6.1. Validez y fiabilidad de los instrumentos

La Ficha de registro de reporte obtenido del firewall no requiere validarse, ya que contiene la información recabada de la misma herramienta de captura de tráfico y del firewall, el cual tiene información en campos validados y normalizados con una estructura de base de datos completa.

La Ficha de recolección de datos para la percepción de la gestión de seguridad implementada actualmente en las redes universitarias, tampoco requiere validarse, por ser un instrumento que sólo permite recopilar información sobre la aplicación de políticas, normas y procedimientos de seguridad informática utilizados en el campus universitario, organizado en base a 3 controles de Seguridad de la Norma ISO 27002(2013). El contenido de las preguntas fue validado por juicio de 6 expertos de la especialidad con grado de Doctor y Magister, encontrando un valor promedio de 83.3, que recae en la valoración de Muy Bueno el instrumento utilizado.

3.6.2. Análisis de datos

Para el primer análisis de datos, se utilizó el tráfico obtenido de la red universitaria en los campos establecidos por el firewall, el mismo que fue consolidado en una matriz de datos utilizando hojas de cálculo en Ms. Excel formato .csv y en el software SPSS Stadistics. A continuación, se utilizó la metodología CRISP-DM, el que fue descrito en la sección 3.5, con actividades de análisis del contenido de los datos, tomar muestras representativas, seleccionar campos, limpiar registros, formatear y convertir datos utilizando la estadística básica para mostrar las propiedades de los datos a través de tablas

de frecuencia y gráficos de distribución con el fin de determinar la consistencia de los datos a trabajar en la siguiente etapa.

Seguidamente en base a este análisis previo, utilizando técnicas de modelado de la herramienta Weka, se clasificó los datos utilizando 7 algoritmos de clasificación que se muestra en el Plan de pruebas en la sección 3.5.4.2 que devuelve el algoritmo adecuado para realizar la predicción de ataques futuros, con características como porcentajes de instancias correctamente clasificadas, estadísticas Kappa, error absoluto medio, etc., resultados propios de las herramientas de minería de datos. Se analizó los indicadores estadísticos obtenidos para definir el nivel de confiabilidad en la predicción acertada del algoritmo elegido, alcanzando un 99.34%, 99.77% y 97.94% para 3 pruebas realizadas, por lo que se acepta este algoritmo para el paso siguiente.

Para el análisis e interpretación de los datos de la primera hipótesis específica, se utilizó el estadístico del Chi-cuadrado con una significancia del 0.001, lo que indica que cada uno de los tipos de ataques hallados, tienen sus propios valores, mantienen sus propias características, así mismo se utilizó la estadística descriptiva para representar las frecuencias y porcentajes de los indicadores de las variables.

Para la segunda hipótesis específica, se analizó los datos con las herramientas de minería de datos, que para el caso de clasificación los modelos utilizan estadísticos del Chi-cuadrado para identificar divisiones optimas en la generación de árboles de decisión y de esta forma predecir los futuros ataques informáticos. Para el caso de asociación, se utilizó el método A priori con niveles de confianza superior a 98% y soporte mínimo de regla, el modelo construyó 30 conjuntos de reglas que definen las relaciones en base al patrón común encontrado que sería el atributo consecuente: “categoría de la amenaza: code-execution”. Para la segmentación se encontraron 2 grupos homólogos, el primero de 2375 registros y el segundo de 6066 registros, que no se ajustan a patrones de datos normales, lo que permitió identificar valores atípicos en el primer grupo en los campos: destination address, país de destino y categoría. En el segundo grupo las anomalías fueron de 32 registros encontrados de los campos: destination address, país de destino, categoría, regla, flags, nombre de amenaza, acción y zona de origen, lo que nos permitió detectar los campos comunes (destination address, país de destino y categoría) como los casos extraños.

Para la tercera hipótesis específica, se aplicó la encuesta vía on-line a través del google-forms a todos los Jefes de TI y Oficiales de seguridad informática de las universidades locales y luego se procedió al análisis estadístico descriptivo para representar las frecuencias y porcentajes de los indicadores de las variables, se utilizó la prueba de Chi Cuadrado para evaluar la relación entre la aplicación del plan de seguridad de la información basado en estándares ISO 27002 y las características de las gestión de la seguridad de la información aplicadas en la universidad.

IV. RESULTADOS

CONTRASTACIÓN DE PRIMERA HIPÓTESIS

Para la primera prueba de hipótesis se utilizó la prueba de Chi Cuadrado, considerando a la categoría de amenaza como variable dependiente y a las demás características de los ataques informáticos del tráfico capturado, como variables independientes. A continuación, se muestra los resultados para cada prueba, teniendo como dato de entrada a partir de la minería de datos, que los ataques más frecuentes son: code-execution y spyware.

Prueba de Categoría de la Amenaza por Mes

Tabla 28.
Categoría de amenaza por Fecha de recepción del ataque

	Categoría de la amenaza			
	Spyware	code-execution	Otros	Total
Enero	1 (0.0%)	34 (0.6%)	35 (7%)	70 (0.8%)
Febrero	7 (0.3%)	39 (0.7%)	32 (6.4%)	78 (0.9%)
Marzo	41 (1.6%)	232 (4.3%)	25 (5%)	298 (3.5%)
Abril	1387 (53.5%)	190 (3.6%)	5 (1%)	1582 (18.7%)
Mayo	140 (5.4%)	246 (4.6%)	12 (2.4%)	398 (4.7%)
Junio	285 (11%)	654 (12.2%)	9 (1.8%)	948 (11.2%)
Julio	71 (2.7%)	1037 (19.4%)	41 (8.2%)	1149 (13.6%)
Agosto	10 (0.4%)	1894 (35.4%)	19 (3.8%)	1923 (22.8%)
Setiembre	269 (10.4%)	226 (4.2%)	36 (7.2%)	531 (6.3%)
Octubre	360 (13.9)	101 (1.9%)	25 (5%)	486 (5.8%)
Noviembre	20 (0.8%)	141 (2.6%)	52 (10.4%)	213 (2.5%)
Diciembre	1 (0%)	550 (10.3%)	211 (42%)	762 (9)
Total	2592 (100%)	5344 (100%)	502 (100%)	8438 (100%)

Fuente: Elaboración propia

Interpretación: Se ha analizado si la categoría de la amenaza tiene variaciones de acuerdo con el mes del año evaluado, se ha encontrado que diferencia en cada mes de acuerdo a los tipos de ataques recibidos (χ^2: 5869.4; $p<0.001$), los ataques de spyware

fueron considerablemente más frecuentes durante el mes de abril (53.5%). La mayoría de los ataques code-execution ocurrieron significativamente más en agosto. 42% de los otros tipos de ataques ocurrieron en diciembre. Esa variación nos hace entender que existen diferencias en cada mes.

Prueba de Categoría de la Amenaza por Hora (formato de 24 horas)

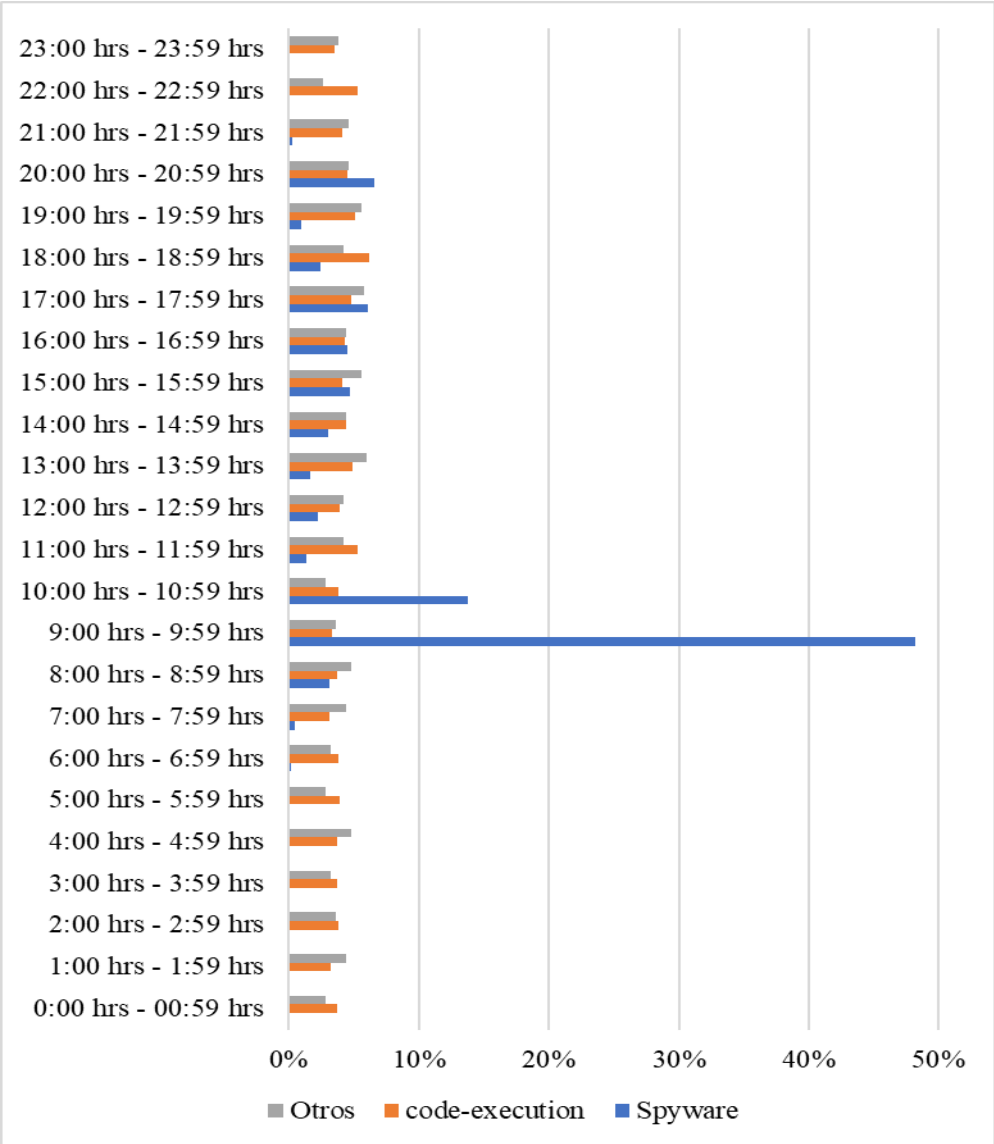

Gráfico 21. *Tabla cruzada Categoría de amenaza Por Hora de recepción del ataque*

Fuente: Elaboración propia

Interpretación: Al evaluar la frecuencia de ataques durante las distintas horas del día de acuerdo al tipo de ataque se ha encontrado que casi la mitad de ataques de Spyware ocurren entre las 9:00 y 9:59 horas. La frecuencia de los ataques de code-execution se encuentran distribuidos durante todo el día. Las diferencias entre las horas y los tipos de amenazas son diferentes, comprobado estadísticamente (χ^2: 3696.2; $p<0.001$).

Prueba de Categoría de la Amenaza Por Hora (Periodo 4 horas)

Tabla 29.
Tabla cruzada Categoría de la amenaza Por Hora

	Categoría de la amenaza			
	Spyware	code-execution	Otros	Total
0:00 hrs - 5:59 hrs	7 (0.3%)	1171 (21.9%)	108 (21.5%)	1286 (15.2%)
6:00 hrs - 11:59 hrs	1741 (67.2%)	1237 (23.1%)	115 (22.9%)	3093 (36.7%)
12:00 hrs - 17:59 hrs	573 (22.1%)	1409 (26.4%)	152 (30.3%)	2134 (25.3%)
18:00 hrs - 23:59 hrs	271 (10.5%)	1527 (28.6%)	127 (25.3%)	1925 (22.8%)
Total	2592 (100%)	5344 (100%)	502 (100%)	8438 (100%)

Fuente: Elaboración propia

Interpretación: Al categorizar las horas del día en cuatro tiempos, las diferencias entre los tipos de ataques y los grupos horarios mantienen sus diferencias, estadísticamente significativas (χ^2: 5869.4; $p<0.001$).

Prueba de Categoría de la amenaza Por Zona de Destino

Tabla 30.
Tabla cruzada Categoría de la amenaza Por Zona de destino

	Categoría de la amenaza			
	Spyware	code-execution	Otros	Total
Interno	35 (1.4%)	5333 (99.8%)	496 (98.8%)	5864 (69.5%)
Externo	2557 (98.6%)	11 (0.2%)	6 (1.2%)	2574 (30.5%)
Total	2592 (100%)	5344 (100%)	502 (100%)	8438 (100%)

Fuente: Elaboración propia

Interpretación: La amenaza puede clasificarse de acuerdo con el origen en las que provienen del exterior de la universidad (que en su gran mayoría son de otros países) y las que provienen del interior de la universidad. Spyware proviene casi en su totalidad del medio externo (98.6%), de igual manera, code-execution proviene del interior de la Universidad en 99.8%. Las otras amenazas también provienen del interior (98.8%). Por lo que el origen y la categoría de las amenazas tienen diferencias estadísticamente significativas (χ^2: 8195.4; $p<0.001$).

Prueba de Categoría de la amenaza Por Regla de sesión

Tabla 31.
Tabla cruzada Categoría de la amenaza Por Regla de sesión

	Categoría de la amenaza			
	Spyware	code-execution	Otros	Total
Actualizaciones-1-1	0 (0%)	2 (0%)	0 (0%)	2 (0%)
RULE-LAN-WAN-LABORATORIOS	44 (1.7%)	0 (0%)	0 (0%)	44 (0.5%)
RULE-WIFI-CAMPUS	1492 (57.6%)	0 (0%)	0 (0%)	1492 (17.7%)
RULE-WIFI-CAMPUS-PRUEBA	133 (5.1%)	0 (0%)	2 (0.4%)	135 (1.6%)
RULE-WIFI-CAMPUS-PRUEBA-1	888 (34.3%)	11 (0.2%)	4 (0.8%)	903 (10.7%)
VIP_Servidor_Biblioteca_HTTP	0 (0%)	13 (0.2%)	0 (0%)	13 (0.2%)
VIP_Servidor_TIC_MOQUEGUA	1 (0%)	821 (15.4%)	43 (8.6%)	865 (10.3%)
VIP_Servidor_Web_HTTP-1	0 (0%)	25 (0.5%)	0 (0%)	25 (0.3%)
VIP_Servidor_Web_HTTPS	4 (0.2%)	3364 (62.9%)	158 (31.5%)	3526 (41.8%)
VIP_Servidor_Web_HTTPS-1	0 (0%)	69 (1.3%)	5 (1%)	74 (0.9%)
VIP_Servidor_Web_RADIO	27 (1%)	339 (6.3%)	276 (55%)	642 (7.6%)
VIP_Svr_Biblioteca_HTTP	3 (0.1%)	700 (13.1%)	14 (2.8%)	717 (8.5%)
Total	2592 (100%)	5344 (100%)	502 (100%)	8438 (100%)

Fuente: Elaboración propia

Interpretación: La regla (rule) más frecuente en ataques de Spyware fue la denominada "RULE-WIFI-CAMPUS", con una frecuencia del 57.6%, seguida del RULE-WIFI-"CAMPUS-PRUEBA-1" (34.3%). Diferente a las frecuencias de spyware, las que en su mayoría (62.9%) estaban categorizadas en "VIP_Servidor_Web_HTTPS" (62.9%), en segundo lugar se encontraba el "VIP_Svr_Biblioteca_HTTP" (13.1%). 55% de los otros tipos de amenazas provenían de "VIP_Servidor_Web_RADIO". Se ha considerado encontrado diferencias estadísticamente (χ^2: 8195.4; $p<0.001$).

Prueba de Categoría de la Amenaza Por Aplicación

Tabla 32.
Tabla cruzada Categoría de la amenaza Por Aplicación

	Categoría de la amenaza			Total
	Spyware	code-execution	Otros	
ms-rdp	0 (0%)	2 (0%)	0 (0%)	2 (0%)
unknow-tcp	2319 (89.5%)	0 (0%)	0 (0%)	2319 (27.5%)
web-browsing	273 (10.5%)	5342 (100%)	242 (48.2%)	5857 (69.4%)
webdav	0 (0%)	0 (0%)	260 (51.8%)	260 (3.1%)
Total	2592 (100%)	5344 (100%)	502 (100%)	8438 (100%)

Fuente: Elaboración propia

Interpretación: Se ha encontrado que la mayoría de las amenazas recibidas de spyware pertenecen a la aplicación categorizada "unknow-tcp", conformando 89.5%. el 100% de los ataques code-execution provienen de "web-browsing". Esto denota una diferencia estadísticamente significativa entre la categoría del ataque y la aplicación (χ^2: 11387.4; $p<0.001$).

Prueba de Categoría de la Amenaza Por Zona de orígen

Tabla 33.
Tabla cruzada Categoría de la amenaza Por Zona de orígen

	Categoría de la amenaza			
	Spyware	code-execution	Otros	Total
LAN	2544 (98.1%)	0 (0%)	3 (0.6%)	2547 (30.2%)
WAN	35 (1.4%)	5333 (99.8%)	496 (98.8%)	5864 (69.5%)
WAN-WIFI	13 (0.5%)	11 (0.2%)	3 (0.6%)	27 (0.3%)
Total	2592 (100%)	5344 (100%)	502 (100%)	8438 (100%)

Fuente: Elaboración propia

Interpretación: La zona de origen también varía notablemente entre las diferentes categorías de amenaza, encontrándose que en 98.1% de las amenazas spyware provienen de LAN, y que el 99.8% de las amenazas code-execution provienen del WAN. Estas diferencias de frecuencias nos dan a entender las notorias variaciones estadísticamente significativas de la zona de estudio entre las categorías evaluadas (χ^2: 8231.1; $p<0.001$).

Prueba de Categoría de la amenaza POR Zona de destino del ataque

Tabla 34.
Tabla cruzada Categoría de la amenaza Por Zona destino de ataque

	Categoría de la amenaza			
	Spyware	code-execution	Otros	Total
	13(0,5%)	13(0,2%)	3(0,6%)	29(0,3%)
WAN	1536(59,3%)	0(0,0%)	0(0,0%)	1536(18,2%)
WAN-WIFI	1008(38,9%)	0(0,0%)	3(0,6%)	1011(12,0%)
ZONA_DMZ1	8(0,3%)	4992(93,4%)	220(43,8%)	5220(61,9%)
ZONA_DMZ2	27(1,0%)	339(6,3%)	276(55,0%)	642(7,6%)
	2592(100,0%)	5344(100,0%)	502(100,0%)	8438(100,0%)

Fuente: Elaboración propia

Interpretación: La evaluación de la categoría de la amenaza y la zona de destino nos indican diferencias notables entre las diferentes categorías, fundamentalmente entre spyware y code-execution. Al evaluar spyware se ha identificado que más de la mitad (59.3%) tuvo como destino a la zona WAN, 38.9% tuvo como destino a la zona WAN-WIFI. En lo que respecta a la categoría zone-execution, la gran proporción de este tipo de amenaza tiene como destino a la Zona DMZ1 (93.4%). Las otras amenazas se encuentran distribuidas entre las Zonas DMZ1 (43.8%) y DMZ2 (55%). Al hacer los análisis estadísticos, se han encontrado diferencias significativas (χ^2: 9842.1; $p<0.001$).

Prueba de Categoría de la amenaza por Número de sesiones con la misma IP (sin agrupar)

Tabla 35.
Tabla cruzada Categoría de la amenaza Por Número de repeticiones sin agrupar

		Categoría de la amenaza			
		Spyware	**code-execution**	**Otros**	**TOTAL**
Repeat Count	**1,00**	1128(43,5%)	3894(72,9%)	340(67,7%)	5362(63,5%)
	2,00	1368(52,8%)	655(12,3%)	25(5,0%)	2048(24,3%)
	3,00	60(2,3%)	308(5,8%)	122(24,3%)	490(5,8%)
	4,00	36(1,4%)	362(6,8%)	7(1,4%)	405(4,8%)
	5,00	0(0,0%)	52(1,0%)	0(0,0%)	52(0,6)
	6,00	0(0,0%)	30(0,6%)	1(0,2%)	31(0,4%)
	7,00	0(0,0%)	25(0,5%)	2(0,4%)	27(0,3%)
	8,00	0(0,0%)	16(0,3%)	3(0,6%)	19(0,2%)
	9,00	0(0,0%)	180,0%)	2(0,4%)	3(0,0%)
	11,00	0(0,0%)	1(0,0%)	0(0,0%)	1(0,0%)
TOTAL		2592(100,0%)	5344(100,0%)	502(100,0%)	8438(100,0%)

Fuente: Elaboración propia

Interpretación: Al evaluar el número de sesiones iniciadas con la misma IP se ha encontrado que entre una (01) a dos (02) veces son las más frecuentes. Tenemos que la amenaza spyware ha iniciado una vez en 43.5%, dos veces en 52.8%. La categoría code-execution ha iniciado una vez en 72.9%, dos veces en 12.3%, tres veces en 5.8%, cuatro veces en 6.8%. Las amenazas categorizadas en “otro”, tienen repeticiones principalmente en una vez de inicio de sesión (67.7%) y tres veces (24.3%).

Prueba de Categoría de la amenaza por Numero de sesiones con la misma IP (agrupadas)

Tabla 36.
Tabla cruzada Categoría de la amenaza Por Numero de repeticiones agrupadas

		Categoría de la amenaza			
		Spyware	**code-execution**	**Otros**	**Total**
Sesiones misma IP (categorizadas)	**1 repetición**	1128(43,5%)	3894(72,9%)	340(67,7%)	5362(63,5%)
	2 repeticiones	1368(52,8%)	655(12,3%)	25(5,0%)	2048(24,3%)
	3 a más repeticiones	96(3,7%)	795(14,9%)	137(27,3%)	1028(12,2%)
Total		2592(100,0%)	5344(100,0%)	502(100,0%)	8438(100,0%)

Fuente: Elaboración propia

Interpretación: Al agrupar el número de repeticiones en tres categorías (1 repetición, 2 repeticiones y 3 a más repeticiones) encontramos que la más frecuente en el spyware fue la de 2 repeticiones (52.8%), casi complementada con las de 1 repetición (43.5%). En el caso de la amenaza code-execution la más frecuente fue la de 1 repetición (72.9%), seguida de las 3 a más repeticiones (14.9%) y por último en frecuencia se encontraban las que tenían dos repeticiones (12.3%). En la categoría de otras amenazas, al igual que la anterior, las más frecuentes fueron las que tenían 1 repetición (67.7%), seguido de las que tenían 3 a más repeticiones (27.3%) y por último complementando las que tenían 2 repeticiones (5%). Al realizar el análisis estadístico, se ha encontrado que existen diferencias estadísticamente significativas entre las diferentes categorías de la amenaza y las repeticiones, es decir sesiones iniciadas con el mismo IP (χ^2: 1779.5.4; $p<0.001$).

Prueba de Categoría de la amenaza por Puertos de origen del ataque

Tabla 37.
Tabla cruzada Categoría de la amenaza Por Puerto de origen

		Categoría de la amenaza			
		Spyware	code-execution	Otros	Total
Puerto de destino	Entre 0 y 1012	13(0,5%)	13(0,2%)	3(0,6%)	29(0,3%)
	Entre 1024 y 49151	864(33,3%)	3875(72,5%)	369(73,5%)	5108(60,5%)
	Entre 49152 y 65535	1715(66,2%)	1456(27,2%)	130(25,9%)	3301(39,1%)
Total		2592(100,0%)	5344(100,0%)	502(100,0%)	8438(100,0%)

Fuente: Elaboración propia

Interpretación: Al evaluar las frecuencias de las categorías de las amenazas y los puertos de destino también se han encontrado diferencias estadísticamente significativas (χ^2: 1161.5; $p<0.001$). Esto queda demostrado que en el caso de las amenazas por spyware fueron más frecuentemente en los puertos que se encontraban entre 49152 y 65535 (66.2%), a diferencia de la amenaza por code-execution, en donde se ha encontrado que los ataques más frecuentes tuvieron como puerto de destino a los que se encontraban entre 1024 y 49151.

Prueba de Categoría de la amenaza Por Puertos de destino del ataque

Tabla 38.
Tabla cruzada Categoría de la amenaza Por Puerto de destino

		Categoría de la amenaza			
		Spyware	**code-execution**	**Otros**	**TOTAL**
Puerto de destino	**Menores de 1024**	1751(67,6%)	4990(93,4%)	223(44,4%)	6964(82,5%)
	De 1024 a más	841(32,4%)	354(6,6%)	279(55,6%)	1474(17,5%)
TOTAL		2592(100,0%)	5344(100,0%)	502(100,0%)	8438(100,0%)

Fuente: Elaboración propia

Interpretación: Al recategorizar los puertos de destino en dos categorías: menores de 1024 y de 1024 a más, se ha visto que las diferencias estadísticamente significativas perduran, pese a que la tendencia de las frecuencias guardan similaridad. Se encontró que en ambas categorías la frecuencia de ataques en los puertos menores de 1024 es la más alta, pero las frecuencias tienen variaciones considerables, en las amenazas por spyware la frecuencia fue de 67.6% mientras que en las de code-execution fue del 93.4%. Contrariamente a estos, en los otros tipos de ataques se ha encontrado que los puertos de 1024 a más son los más frecuentes (55.6%). (χ^2: 1344.9; $p<0.01$).

Prueba de Categoría de la amenaza Por Puertos de destino web

Tabla 39.
Tabla cruzada Categoría de la amenaza Por Puerto de destino Web

		Categoría de la amenaza			
		Spyware	**code-execution**	**Otros**	**Total**
Puertos Web de destino	**80 y 8000 (web internos)**	242(9,3%)	5329(99,7%)	499(99,4%)	6070(71,9%)
	Otros	2350(90,7%)	15(0,3%)	3(0,6%)	2368(28,1%)
Total		2592(100,0%)	5344(100,0%)	502(100,0%)	8438(100,0%)

Fuente: Elaboración propia

Interpretación: Al evaluar si existen diferencias entre los puertos web de destino y las distintas categorías de las amenazas se han encontrado diferencias estadísticamente significativas (χ^2: 7262.3; $p<0.001$). Las frecuencias difieren notoriamente en el caso de las dos amenazas principales, se ha encontrado que los puertos de destino favoritos para las amenazas code-execution son los 80 y 8000 (99.7%), a diferencia de las amenazas por spyware, en donde los otros puertos son los favoritos, encontrándose una frecuencia de 90.7%. Así mismo, los que tienen otras categorías de amenazas tienen su mayoría de puertos web de destino a los que están entre 80 y 8000 (99.4%).

Prueba de Categoría de la amenaza Por Detalle del ataque en hexadecimal

Tabla 40.
Tabla cruzada Categoría de la amenaza Por Detalle del ataque en hexadecimal

		Categoría de la amenaza			
		Spyware	code-execution	Otros	Total
Flags	0x10200(credenc.empr)	1(0,0%)	339(6,3%)	276(55,0%)	616(7,3%)
	0x1102000(tx. en sesion proxy)	44(1,7%)	2(0,0%)	0(0,0%)	460,5%
	0x2000 (inundación)	1526(58,9%)	4992(93,4%)	220(43,8%)	6738(79,9%)
	0x422000 (NAT)	1021(39,4%)	11(0,2%)	6(1,2%)	1038(12,3%)
Total		2592(100,0%)	5344(100,0%)	502(100,0%)	8438(100,0%)

Fuente: Elaboración propia

Interpretación: Los detalles del ataque en hexadecimal guardan diferencias estadísticamente significativas entre las distintas categorías de la amenaza (χ^2: 4428; $p<0.001$). Se ha encontrado que el flag 0x2000 es uno de los más frecuentes, en las amenazas de tipo spyware se ha visto que representa el 58.9%, en las amenazas de tipo code-execution conforman casi la totalidad de éstas (93.4%), y en los otros tipos de amenazas conforman el 43.8%. En este último grupo la flag de mayor frecuencia ha sido 0x10200.

Prueba de Categoría de la amenaza por Acción tomada para la sesión

Tabla 41.
Tabla cruzada Categoría de la amenaza Por Acción tomada para la sesión

		Categoría de la amenaza			
		Spyware	**code-execution**	**Otros**	**Total**
Action	**Alert**	0(0,0%)	2(0,0%)	0(0,0%)	2(0,0%)
	reset-both	2332(90,0%)	5337(99,9%)	304(60,6%)	7973(94,5%)
	reset-server	260(10,0%)	5(0,1%)	198(39,4%)	463(5,5%)
Total		2592(100,0%)	5344(100,0%)	502(100,0%)	8438(100,0%)

Fuente: Elaboración propia

Interpretación: Las acciones tomadas para la sesión de acuerdo a la amenaza ha sido similar en las tres categorías de amenazas, variando en sus frecuencias. En las amenazas de tipo spyware se ha encontrado que el 90% tienen la acción *reset-both*, esta misma acción se da en el 99.9% de las amenazas code-execution, y la frecuencia de esta acción en los otros tipos de ataques fue del 60.6%. Se debe considerar que en solo dos casos se ha dado la acción *alert*, los cuales pertenecieron a amenazas del tipo code-execution.

Prueba de Categoría de la amenaza POR Dirección del ataque

Tabla 42.
Tabla cruzada Categoría de la amenaza Por Dirección del ataque

		Categoría de la amenaza			
		Spyware	**code-execution**	**Otros**	**Total**
Dirección	**client-to-server**	2579(99,5%)	5331(99,8%)	499(99,4%)	8409(99,7%)
	server-to-client	13(0,5%)	13(0,2%)	3(0,6%)	29(0,3%)
Total		2592(100.0%)	5344(100.0%)	502(100.0%)	8438(100.0%)

Fuente: Elaboración propia

Interpretación: Se analizó la dirección del ataque, se ha encontrado que esta dirección no tiene variación de acuerdo con la categoría de la amenaza (χ^2: 4.404; p=0.111). Se observa que la dirección de ataque fundamentalmente se da de cliente a servidor, 99.5% en spyware, 99.8% en code-exection y 99.4% en los otros tipos de amenazas.

Prueba de Categoría de la amenaza por Lugar de origen

Tabla 43.
Tabla cruzada Categoría de la amenaza Por Lugar de origen

		Categoría de la amenaza			
		Spyware	**code-execution**	**Otros**	**Total**
Lugar de Origen	**Interno**	2545(98,2%)	1684(31,5%)	75(14,9%)	4304(51,0%)
	Externo	47(1,8%)	3660(68,5%)	427(85,1%)	4134(49,0%)
	Total	2592(100,0%	5344(100,0%)	502(100,0%)	8438(100,0%)

Fuente: Elaboración propia

Interpretación: Considerando el lugar de origen y las diferentes categorías de amenazas se ha visto que existen diferencias estadísticamente significativas (χ^2: 3382.8; p<0.001). Para las amenazas de tipo spyware se ha encontrado que el origen es interno en 98.2% y para las amenazas de tipo code-execution el origen es externo es de 68.5%. Los otros tipos de amenazas también provienen de origen externo en 85.1% de los casos.

Prueba de Categoría de la amenaza POR Continente

Tabla 44.
Tabla cruzada Categoría de la amenaza Por Continente

	brute-force	code-execution	info-leak	net-worm	overflow	spyware	sql-injection	webshell	TOTAL
América	0(0,0%)	294(8,9%)	5(83,3%)	21(14,8%)	29(11,6%)	39(83,0%)	0(0,0%)	1(50,0%)	389(10,4%)
Europa	11(100,0%)	1061(32,2%)	1(16,7%)	5(3,5%)	8(3,2%)	8(17,0%)	0(0,0%)	0(0,0%)	1094(29,2%0)
Asia	0(0,0%)	1565(47,6%)	0(0,0%)	115(81,0%)	213(85,2%)	0(0,0%)	1(100,0%)	1(50,0%)	1895(50,5%)
África	0(0,0%)	370(11,2%)	0(0,0%)	1(0,7%)	0(0,0%)	0(0,0%)	0(0,0%)	0(0,0%)	371(9,9%)
TOTAL	11(100,0%)	3290(100,0%)	6(100,0%)	142(100,0%)	250(100,0%)	4(100,0%)7	1(100,0%)	2(100,0%)	3749(100,0%)

Fuente: Elaboración propia

Interpretación: Al evaluar cada uno de los tipos de amenazas en relación con los continentes de origen del ataque se han encontrado diferencias entre sí, no teniendo una distribución similar por continente y tipo de amenaza (χ^2: 579.9; $p<0.001$). Se observa que el 100% de las amenazas brute-force proviene de Europa, las amenazas code-execution provienen en su mayoría de Asia (47.6%) y Europa (32.2%). Info-leak tiene su origen en América en 83.3% de estas amenazas. Net-worm proviene principalmente de Asia (81%). El 85.2% de Overflow proviene de Asia. 83% de Spyware proviene de América. El único ataque (100%) de sql-injection provino de Asia. Uno de los dos ataques (50%) de Webshell provino de América y el otro (50%) de Asia.

CONTRASTACIÓN DE SEGUNDA HIPÓTESIS ESPECÍFICA

La prueba de hipótesis indica que “las características de los ataques informáticos permiten predecir el tipo de la amenaza a través de modelos de minería de datos”, se utilizó las herramientas Weka y SPSS Modeler para clasificar, asociar y segmentar.

Se puede concluir que con un nivel de confianza de 99.88% y soporte minimo de regla de 63.91%, se predice la categoria de amenaza "code-execution" si la aplicacion utilizada es "Web-browsing" y la acción de proteccion a enviar es "reset-both".

1. CLASIFICACIÓN

Se utilizó el modelo de clasificación basado en árboles de decisión CHAID y C5.0 del SPSS Modeler para clasificar y pronosticar valores de la variable dependiente (categoría de la amenaza) en contrastación de los demás valores considerados como variables independientes (predictores). Del Weka se utilizó el algoritmo REPTree.

CHAID

Este modelo genera árboles de decisión de múltiples niveles utilizando estadísticos de chi-cuadrado para identificar las divisiones óptimas.

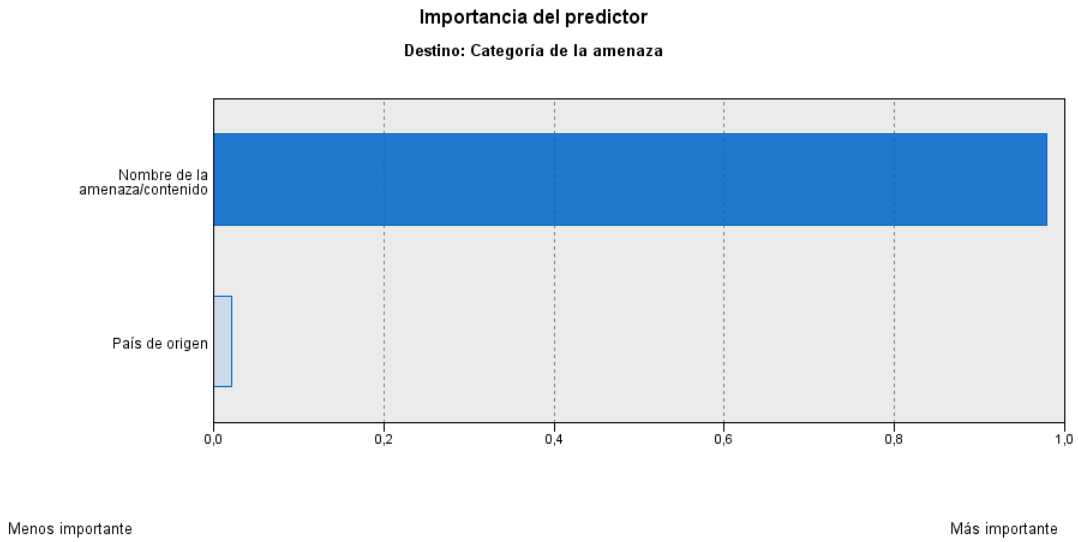

Gráfico 22. *Importancia del predictor considerando como destino a la Categoría de la amenaza*

Fuente: Elaboración propia

Se ha encontrado que la categoría con más alta probabilidad de predictor es el nombre de la amenaza/contenido (98%), seguida del país de origen (2%). Es decir, podemos indicar que ésta sería una posible clasificación para tomarla en cuenta en la predicción.

Resumen del Análisis:

Análisis

Profundidad del árbol: 2

Campos

Destino

Categoría de la amenaza

Entradas

Nombre de la amenaza/contenido
País de origen
Configuración de creación
Utilizar los datos en particiones: falso
Calcular importancia de predictor: true
Calcular puntuaciones de propensión en bruto: falso
Calcular puntuaciones de propensión ajustada: falso
Continuar entrenando modelo existente: falso
Utilizar frecuencia: falso
Utilizar ponderación: falso
Método: CHAID
Niveles por debajo del raíz: 5
Alfa para división: 0,05
Alfa para fusión: 0,05
Épsilon para convergencia: 0,001
Número máximo de iteraciones para la convergencia: 100
Utilizar corrección de Bonferroni: true
Permitir división de categorías fusionadas: falso
Método de chi-cuadrado: Pearson
Criterios de parada: Utilizar porcentaje
Mínimo de registros en rama padre (%): 2
Mínimo de registros en rama hijo (%): 1
Utilizar costes de clasificación errónea: falso

C5.0

Este modelo divide la muestra basándose en el campo que tiene mayor ganancia de información en cada nivel. Resultando entonces el siguiente gráfico:

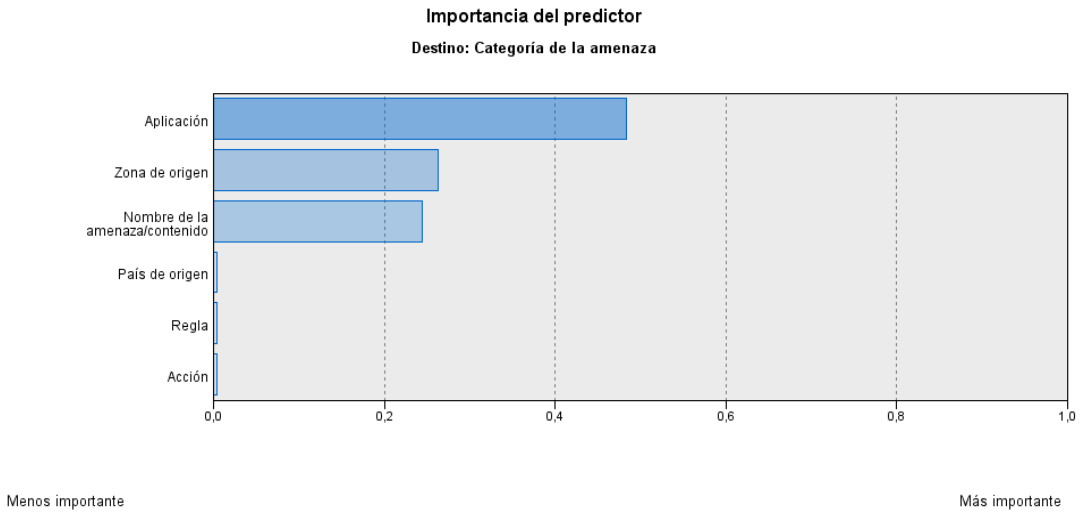

Gráfico 23. *Importancia del predictor considerando como destino a la Categoría de la amenaza*

Fuente: Elaboración propia

Se ha encontrado que la importancia del predictor “Aplicación” es de 48%, de “Zona de origen” es de 26%, de nombre de la amenaza/contenido es de 24%. Y con menor importancia, acumulando un 2% hacen “País de origen”, “Regla” y “Acción”.

Resumen del Análisis:
Análisis
Profundidad del árbol: 5
Campos
Destino
Categoría de la amenaza
Entradas
Aplicación
Acción
Zona de origen
Nombre de la amenaza/contenido
Regla
País de origen
Configuración de creación
Utilizar los datos en particiones: falso
Calcular importancia de predictor: true
Calcular puntuaciones de propensión en bruto: falso
Calcular puntuaciones de propensión ajustada: falso
Utilizar ponderación: falso
Tipo de resultados: Árbol de decisión
Agrupar simbólicos: falso
Utilizar aumento: falso
Efectuar validación cruzada: falso
Modo: Simple
Favorecer: Precisión
Ruido esperado (%): 0
Utilizar costes de clasificación errónea: falso

REPTree

El algoritmo implementado en Weka, utiliza la información de varianza para construir el árbol y lo va podando para reducir los errores en base a la información de la base de datos pre-procesada.

Tabla 45.
Clasificación de instancias Algoritmo Árbol de decisión REPTree

	Cantidad	%
Instancias correctamente clasificadas	8433	99.9407 %
Instancias incorrectamente clasificadas	5	0.0593 %
TOTAL	**8438**	**100%**

Fuente: Elaboración propia

En la tabla anterior se observa que el algoritmo clasificó de manera correcta hasta un 99.9407% las instancias para la condición de 8 tipos de amenazas y clasifica de manera incorrecta 0.0593% del total de las instancias.

Tabla 46.
Matriz de confusión REPTree

code-execution	**spyware**	**sql-injection**	**overflow**	**info-leak**	**webshell**	**brute-force**	**net-worm**	**TOTAL**
5344	0	0	0	0	0	0	0	5344
4	2588	0	0	0	0	0	0	2592
0	0	3	0	0	0	0	0	3
0	0	0	262	0	0	0	0	262
0	0	0	0	10	0	0	0	10
1	0	0	0	0	5	0	0	6
0	0	0	0	0	0	28	0	28
0	0	0	0	0	0	0	193	193

Fuente: Elaboración propia

La matriz de consistencia muestra que, de 8438 registros, 5344 fueron clasificados solamente como “code-execution”; representa la mayor cantidad de ataques dentro de la categoría amenazas de la base de datos capturada, seguido de spyware disperso con code-execution, esto nos permitió definir el predictor para usarlo con C5.0 y CHAID.

No solo se modeló los datos en Weka con REPTree, sino que en el apartado 3.5.4.2 se trabajó para la clasificación con 7 algoritmos que detallamos el resumen en la tabla siguiente.

Tabla 47.
Resumen del modelado de algoritmos de clasificación para la Categoría de Amenazas.

	CATEGORIA DE AMENAZA		
	% predicción acertada	Instancias correctamente clasificadas	Instancias incorrectamente clasificadas
REPTree	99.94%	8433	5
RandomForest	99.93%	8432	6
J48	99.47%	8393	45
NaiveBayes	98.63%	8322	116
RandomTree	97.75%	8248	190
DecisionStump	94.05%	7936	502
HoeffdingTree	80.97%	6832	1606

Fuente: Elaboración propia

Entonces como se aprecia el mejor desempeño lo realiza el algoritmo de clasificación REPTree, por lo tanto, con este algoritmo se realiza la predicción de los futuros ataques informáticos.

2. ASOCIACION

Para la asociación se empleó el método “Apriori”. En el diseño del modelo se ha considerado un soporte mínimo de regla del 60 y una confianza mínima del 95%. El número máximo de antecedentes fue de 5. A continuación se presentan todos los casos que cumplen los requisitos con los consecuentes indicados:

Tabla 48.

Modelo A priori para asociar ataques informáticos en base a Categoría de amenaza

Consecuente	Antecedente	% de soporte	% de confianza
Categoría de la amenaza = code-execution	Aplicación = web-browsing and País de destino = 10.0.0.0-10.255.255.255 and Acción = reset-both	63.6062078	99.29223319
Categoría de la amenaza = code-execution	Aplicación = web-browsing and País de destino = 10.0.0.0-10.255.255.255 and Zona de origen = WAN and Acción = reset-both	63.6062078	99.29223319
Categoría de la amenaza = code-execution	Aplicación = web-browsing and Zona de origen = WAN and Acción = reset-both	63.64174861	99.23678332
Categoría de la amenaza = code-execution	Aplicación = web-browsing and Acción = reset-both	63.91422817	98.88785913
Categoría de la amenaza = code-execution	Zona de destino = ZONA_DMZ1 and País de destino = 10.0.0.0-10.255.255.255	61.8410141	95.63218391
Categoría de la amenaza = code-execution	Zona de destino = ZONA_DMZ1 and País de destino = 10.0.0.0-10.255.255.255 and Zona de origen = WAN	61.8410141	95.63218391
Categoría de la amenaza = code-execution	Zona de destino = ZONA_DMZ1 and País de destino = 10.0.0.0-10.255.255.255 and Flags = 0x2000	61.8410141	95.63218391
Categoría de la amenaza = code-execution	Zona de destino = ZONA_DMZ1 and País de destino = 10.0.0.0-10.255.255.255 and Zona de origen = WAN and Flags = 0x2000	61.8410141	95.63218391
Categoría de la amenaza = code-execution	Zona de destino = ZONA_DMZ1 and Aplicación = web-browsing and País de destino = 10.0.0.0-10.255.255.255	61.81732022	95.63050977
Categoría de la amenaza = code-execution	Aplicación = web-browsing and País de destino = 10.0.0.0-10.255.255.255 and Flags = 0x2000	61.81732022	95.63050977
Categoría de la amenaza = code-execution	Zona de destino = ZONA_DMZ1 and Aplicación = web-browsing and País de destino = 10.0.0.0-10.255.255.255 and Zona de origen = WAN	61.81732022	95.63050977

Consecuente	Antecedente	% de soporte	% de confianza
Categoría de la amenaza = code-execution	Zona de destino = ZONA_DMZ1 and Aplicación = web-browsing and País de destino = 10.0.0.0-10.255.255.255 and Flags = 0x2000	61.81732022	95.63050977
Categoría de la amenaza = code-execution	Aplicación = web-browsing and País de destino = 10.0.0.0-10.255.255.255 and Zona de origen = WAN and Flags = 0x2000	61.81732022	95.63050977
Categoría de la amenaza = code-execution	Zona de destino = ZONA_DMZ1 and Aplicación = web-browsing and País de destino = 10.0.0.0-10.255.255.255 and Zona de origen = WAN and Flags = 0x2000	61.81732022	95.63050977
Categoría de la amenaza = code-execution	Aplicación = web-browsing and País de destino = 10.0.0.0-10.255.255.255	66.05852387	95.6061693
Categoría de la amenaza = code-execution	Aplicación = web-browsing and País de destino = 10.0.0.0-10.255.255.255 and Zona de origen = WAN	66.05852387	95.6061693
Categoría de la amenaza = code-execution	Zona de destino = ZONA_DMZ1	61.87655491	95.57725445
Categoría de la amenaza = code-execution	Zona de destino = ZONA_DMZ1 and Zona de origen = WAN	61.87655491	95.57725445
Categoría de la amenaza = code-execution	Zona de destino = ZONA_DMZ1 and Flags = 0x2000	61.87655491	95.57725445
Categoría de la amenaza = code-execution	Zona de destino = ZONA_DMZ1 and Zona de origen = WAN and Flags = 0x2000	61.87655491	95.57725445
Categoría de la amenaza = code-execution	Zona de destino = ZONA_DMZ1 and Aplicación = web-browsing	61.85286104	95.57556024
Categoría de la amenaza = code-execution	Aplicación = web-browsing and Flags = 0x2000	61.85286104	95.57556024
Categoría de la amenaza = code-execution	Zona de destino = ZONA_DMZ1 and Aplicación = web-browsing and Zona de origen = WAN	61.85286104	95.57556024
Categoría de la amenaza = code-execution	Zona de destino = ZONA_DMZ1 and Aplicación = web-browsing and Flags = 0x2000	61.85286104	95.57556024
Categoría de la amenaza = code-execution	Aplicación = web-browsing and Zona de origen = WAN and Flags = 0x2000	61.85286104	95.57556024

Consecuente	Antecedente	% de soporte	% de confianza
Categoría de la amenaza = code-execution	Zona de destino = ZONA_DMZ1 and Aplicación = web-browsing and Zona de origen = WAN and Flags = 0x2000	61.85286104	95.57556024
Categoría de la amenaza = code-execution	Aplicación = web-browsing and Zona de origen = WAN	66.09406468	95.55475892
Categoría de la amenaza = code-execution	País de destino = 10.0.0.0-10.255.255.255 and Flags = 0x2000	62.14903447	95.15821578
Categoría de la amenaza = code-execution	País de destino = 10.0.0.0-10.255.255.255 and Zona de origen = WAN and Flags = 0x2000	62.14903447	95.15821578
Categoría de la amenaza = code-execution	Zona de origen = WAN and Flags = 0x2000	62.18457529	95.1038293

Resumen del análisis

Análisis

- Número de reglas: 30
- Número de transacciones válidas: 8.441
- Soporte mínimo: 61,817%
- Soporte máximo: 66,094%
- Confianza mínima: 95,104%
- Confianza máxima: 99,292%
- Elevación mínima: 1,502%
- Elevación máxima: 1,568%
- Capacidad de despliegue mínima: 0,45%
- Capacidad de despliegue máxima: 3,045%
- Soporte de regla mínimo: 59,116%
- Soporte de regla máximo: 63,203%

Campos

- Consecuentes
  - Categoría de la amenaza
- Antecedentes
  - Destination address
  - Regla
  - Aplicación
  - Zona de origen
  - Zona de destino
  - Flags
  - Acción
  - Nombre de la amenaza/contenido
  - Categoría
  - País de origen
  - País de destino

Configuración de creación

- Utilizar los datos en particiones: falso

Número máximo de antecedentes: 5
Soporte mínimo de las reglas: 60,0
Confianza mínima de regla (%): 95,0
Optimizar: Velocidad
Sólo valores verdaderos para las marcas: true
Datos transaccionales: falso

Resumen de entrenamiento

Algoritmo: Apriori
Tipo de modelo: Asociación
Ruta: Ruta2
Usuario: Karina
Fecha de creación: 24/08/19 19:52
Aplicación: IBM® SPSS® Modeler 18
Tiempo transcurrido para la generación del modelo: 0 horas, 0 minutos, 10 segundos

Con un nivel de confianza de 99.29% y soporte mínimo de regla de 63.60%, se predice la categoría de amenaza "code-execution" si la aplicación utilizada es "Web-browsing", el destino de ataque es la red interna 10.0.0.0/24 y la acción de protección es enviar "reset-both"

Con un nivel de confianza de 99.29% y soporte mínimo de regla de 63.60%, se predice la categoría de amenaza "code-execution" si la aplicación utilizada es "Web-browsing", el destino de ataque es la red interna 10.0.0.0/24, la zona de origen del ataque es la WAN y la acción de protección es enviar "reset-both"

Con un nivel de confianza de 99.88% y soporte minimo de regla de 63.91, se predice la categoria de amenaza "code-execution" si la aplicacion utilizada es "Web-browsing"y la accion de proteccion es enviar "reset-both".

Observamos que el modelo construyó 30 conjuntos de reglas que definen las relaciones en base al patrón común encontrado que sería el atributo consecuente: “categoría de la amenaza: code-execution”. Por lo tanto, se concluye que la categoría de amenaza más frecuente en base al modelo de asociación Apriori es “**code-execution**”.

3. SEGMENTACIÓN

Para la segmentación se evaluó el "Modelo de detección de anomalías" de todas las variables del análisis de datos. Se ha considerado como 1% el porcentaje de registros más anómalos de los datos de entrenamiento, que es el cociente del índice de desviación del grupo sobre su media sobre el clúster al que pertenece. Se consideró este valor para que la desviación no sea tan alta sobre la media.

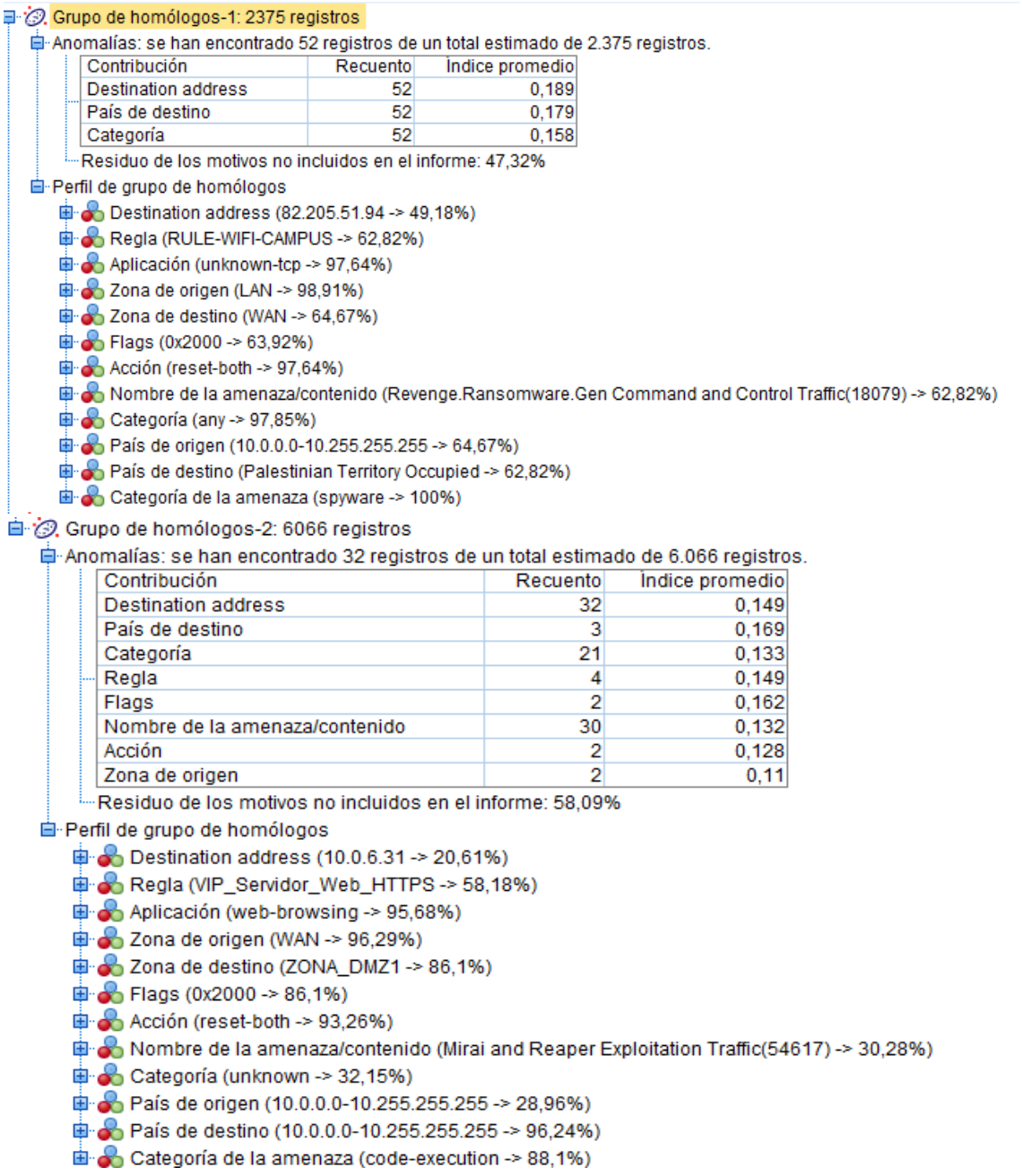

Gráfico 24. *Modelo del análisis de detección de anomalías*

RESUMEN DEL ANÁLISIS

Análisis

Número de grupos de homólogos: 2

Corte de índice de anomalía: 3,67907

Campos

Entradas

Destination address

Regla

Aplicación

Zona de origen

Zona de destino

Flags

Acción

Nombre de la amenaza/contenido

Categoría

País de origen

País de destino

Categoría de la amenaza

Configuración de creación

Utilizar los datos en particiones: falso

Calcular puntuaciones de propensión en bruto: falso

Calcular puntuaciones de propensión ajustada: falso

Determinar valor de corte para la anomalía basado en: Porcentaje de registros más anómalos de los datos de entrenamiento

Porcentaje de registros anómalos: 1,0

Número de campos de anomalía: 3

CONTRASTACIÓN DE TERCERA HIPÓTESIS ESPECÍFICA

Para la prueba de la tercera hipótesis específica, que se pretende probar que "La gestión de seguridad que maneja el personal de TI está relacionada con la aplicación de controles de seguridad basado en estándares de la ISO 27002.", se utilizó los resultados de la encuesta via on-line que se aplicó a los Administradores y Oficiales de seguridad de TI, para cada uno se realizó el análisis estadístico y posteriormente se cruzó información de las preguntas vinculadas a la aplicación de un plan de seguridad basado en estándares como la ISO 27002, con los que quedó demostrado que la aplicación de controles de seguridad dispuesto por una norma internacional de seguridad de la información ISO 27002(2013) no se relaciona con la forma de gestionar la seguridad de las redes informáticas.

Pregunta 1: ¿Utiliza protocolos de cifrado aprobados por la Universidad en las comunicaciones inalámbricas?

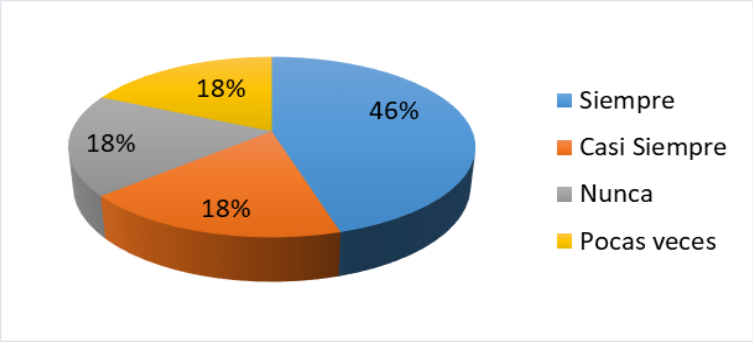

Gráfico 25. *Uso de protocolos de cifrado en la Universidad*

Se aprecia que se tiene configurados protocolos que aseguren las comunicaciones en el campus universitario en un 46%, pero preocupa al ser un tema de alta importancia proteger la data que entra, se procesa y sale de la Universidad, no se tenga un mayor nivel de aplicación.

Pregunta 2: En caso sea afirmativa la respuesta anterior, ¿podría indicar cuáles?

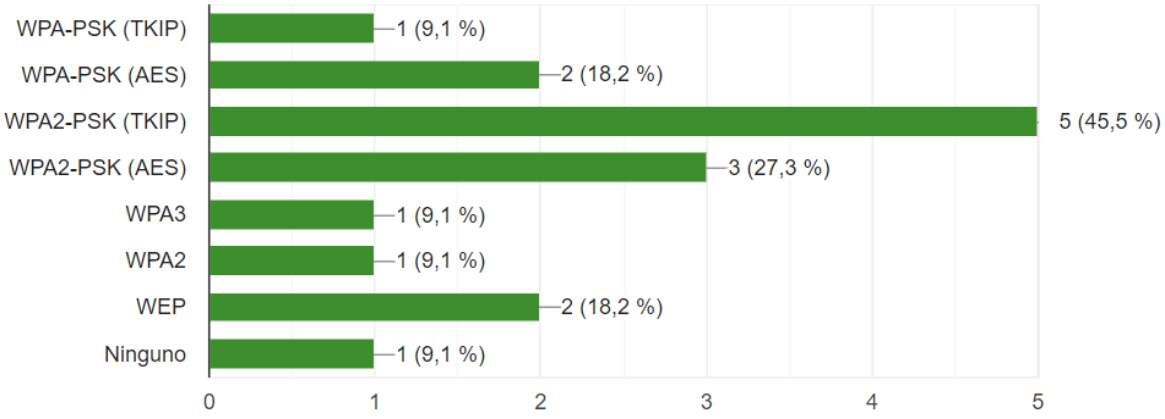

Gráfico 26. *Tipo de protocolos de cifrado utilizado*

Un 45.5% indica que se tiene implementado el protocolo de cifrado WPA2-PSK(TKIP) que es un estándar antiguo de WPA, pero se puede considerar aceptable, le sigue en orden de implementación WPA2-PSK(AES) con un 27.3% que es una solución más reciente para la seguridad de redes Wifi utilizado por WPA2. Aquí lo curioso es que, aún se sigue utilizando el protocolo WEP en 18.2% dado que es un estándar antiguo de encriptación que es vulnerable y ya no se debe utilizar sea el cifrado a 64 o 128 bits.

Pregunta 3: ¿Usa contraseñas o autenticación para usuarios con acceso remoto?

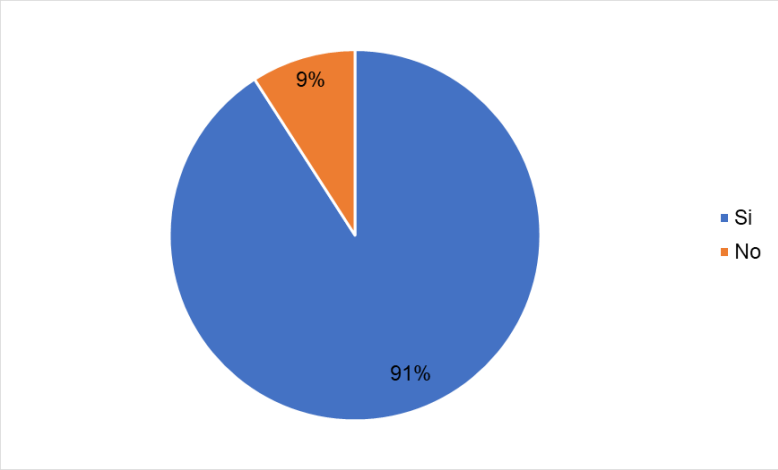

Gráfico 27. *Uso de contraseñas o autenticación*

Se verifica que en 91% en el campus universitario se utilizan contraseñas o algún método de autenticación para conexiones remotas, considerándose una buena práctica.

Pregunta 4: ¿Identifica niveles de acceso de todos los usuarios de la red universitaria para ajustarlos según sea necesario y bloquea intentos de Superusuario o Administrador?

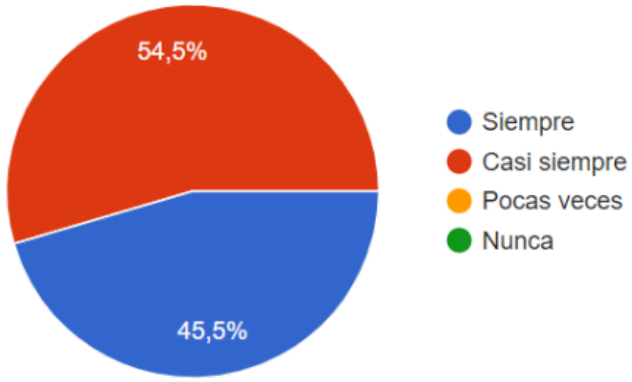

Gráfico 28. *Niveles de acceso y bloqueo de accesos privilegiados*

Se aprecia que la implementación de niveles de acceso por categoría a las instalaciones de la infraestructura de la universidad se aplica siempre en 45.5% y casi siempre en 54.5%, así mismo ésta práctica conlleva al bloqueo de intentos de acceso con privilegios de Superusuario o Administrador, práctica que según la norma ISO 27002 debe siempre aplicarse.

Pregunta 5: ¿Existe alguna política por parte de los encargados de TI para el cambio de su contraseña en su computador cada cierto periodo de tiempo?

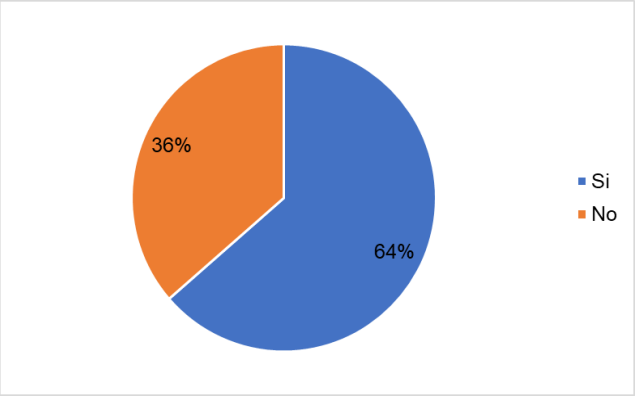

Gráfico 29. *Política de cambio de contraseñas*

En un 64% se indica que, sí existe la política de cambio de contraseñas en el computador, se considera una buena práctica que debería implementarse, porque la protección a estaciones de trabajo también es parte de la gestión de seguridad informática de la Universidad en su conjunto y si el personal no toma conciencia de practicarlas, no se podrá avanzar en mejorar el rendimiento de la red.

Pregunta 6: Si existe alguna política o exigencia con qué frecuencia lo realiza

Tabla 49.
Frecuencia de cambio de contraseñas

	Frecuencia	Porcentaje
Cada mes	4	36,4%
Cada seis meses	4	36,4%
Cada tres meses	2	18,2%
Cada año	1	9,1%
Total	11	100,0%

De la pregunta anterior se desprende que, si existe la política de cambio de contraseña, entonces se efectúa cada mes y cada 6 meses en 36.4% y cada 3 meses un 18.2%. Se considera un lapso aceptable de cambio de contraseñas debido a las diferentes vulnerabilidades que se dan en dispositivos de usuario final que son la primera entrada a las redes corporativas.

Pregunta 7: ¿Desactiva las credenciales de los trabajadores que ya no laboran en la universidad en coordinación con RRHH apenas termina su contrato?

Tabla 50.
Frecuencia de desactivación de credenciales a trabajadores ajenos a la institución

	Frecuencia	Porcentaje
Casi siempre	5	45,5%
Siempre	4	36,4$
Pocas veces	2	18,2%
Nunca	0	0,0%
Total	11	100,0%

En un 45.5% el departamento de TI realiza en coordinación con RRHH la desactivación de credenciales a los trabajadores que han dejado de laborar en la institución, siempre se hace esta práctica en 36.4% y haciendo un total de 81.9% que es muy significativo dada la importancia de implementar controles de seguridad; en un 18.2% la práctica se realiza pocas veces, este último valor significa un riesgo a la seguridad.

Pregunta 8: ¿Existe en el departamento de TI, procedimientos a seguir en caso de ocurrir algún problema con los servidores, computadoras o servicios informáticos que necesita para labores cotidianas

Tabla 51.
Existencia de procedimientos en caso de eventualidades

	Frecuencia	Porcentaje
Siempre	5	45,5%
Casi siempre	3	27,3%
Pocas veces	2	18,2%
Nunca	1	9,1%
Total	11	100,0%

En un 45.5% el departamento de TI cuenta con procedimientos a seguir en caso de ocurrir eventualidades vinculadas a la labor con el uso de tecnología de computadoras sea por medios cableados o inalámbricos; casi siempre indican que existe estas actividades en 27.3% y pocas veces en 18,2% seguramente lo aplican. Esta buena práctica no debería quedarse solo en conocer, sino en aplicarse, de aquí se desprenden las siguientes interrogantes.

Pregunta 9: ¿Existe en el departamento de TI, alguna política que indique en qué periodo tiempo se debe respaldar la información de su equipo?

Tabla 52.
Existencia de política de respaldo de la información

	Frecuencia	Porcentaje
Si	6	54,5%
No	5	45,5%
Total	11	100,0

En un 54.5% los encuestados indican que se tiene una política de gestión que indica que se debe realizar el respaldo de la información, pero el porcentaje restante 45,5% no lo sabe, no lo conoce o no lo aplica, que es preocupante.

Pregunta 10. En caso de realizar el respaldo de información indique cada que tiempo lo realiza

Tabla 53.
Intervalo de tiempo de respaldo de información

	Frecuencia	Porcentaje
Cada mes	6	54,5%
No precisa	2	18,2%
Cada 6 meses	2	18,2%%
Cada 3 meses	1	9,1%
Total	11	100,0%

El 54.5% de los encuestados indican que realizan respaldo de su información cada mes, el 18.2% cada 6 meses, período que considero muy extenso dado que ante las continuas amenazas que se presentan a diario, esta actividad debería darse con más frecuencia dada la importancia de la información que se procesa y almacena en el ámbito académico.

Pregunta 11: ¿Los backups y archivos con data sensible están cifrados?

Tabla 54.
Cifrado de información de respaldo

	Frecuencia	Porcentaje
Poca veces	4	36,4%
Casi siempre	3	27,3%
Nunca	2	18,2%
Siempre	2	18,2%
Total	11	100,0%

El 18.2% del total de encuestados indica que su data respaldada siempre se encuentra cifrada, casi siempre el 27.3% y pocas veces respalda el 36.4%. Se puede apreciar que existe la necesidad no solo de sacar backups sino que esta se encuentre protegida

Pregunta 12: Los parches de seguridad solucionan agujeros de seguridad sin modificar la funcionalidad del programa, son frecuentes en aplicaciones que interactúan con Internet. ¿Se aplican estos parches en sus servidores y estaciones de trabajo?

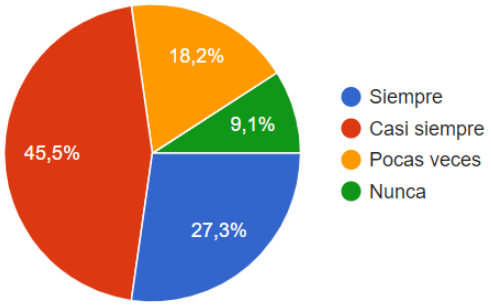

Gráfico 30. *Aplicación de parches en servidores y estaciones de trabajo*

Del total de administradores de seguridad de TI, indica el 27.3% que siempre instala parches de seguridad en su estación de trabajo o servidores; un 45.5% casi siempre lo implementa y 18.2% pocas veces hace esta práctica, de esta forma el envío de la información se realiza con menor riesgo de ser leído por terceros.

Pregunta 13: ¿Qué tipo de parches configura

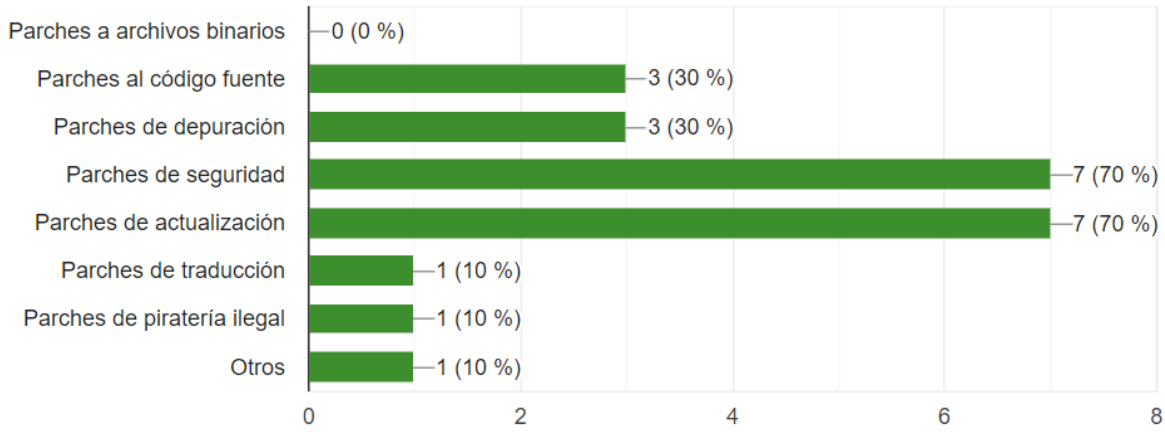

Gráfico 31. *Tipo de parches instalados*

En mayor número (7), los encuestados indican que instalan parches de seguridad y de actualización en sus sistemas, otra menor cantidad (3) indican que instalan parches al condigo fuente y de depuración. Esta es una buena práctica por mantener sus aplicaciones actualizadas.

Pregunta 14: ¿Realiza un registro de los logs de eventos de seguridad, actividades de usuarios, excepciones, fallas?

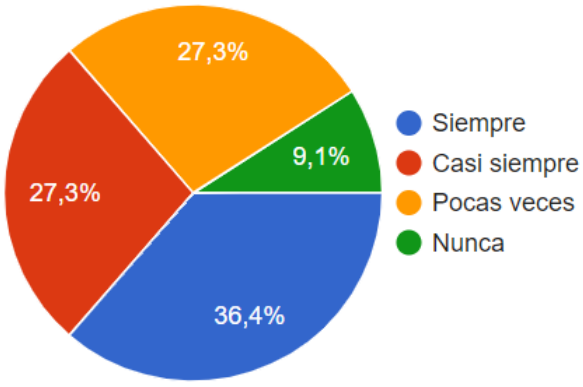

Gráfico 32. *Registro de logs en el sistema*

Un 36.4% siempre registra los logs vinculados a eventos de seguridad u otras acciones que se dan en sus sistemas, casi siempre un 27.3%, considerándose una buena práctica de seguridad en general.

Pregunta 15: ¿Analiza los logs de seguridad de sus sistemas en línea o fuera de línea?

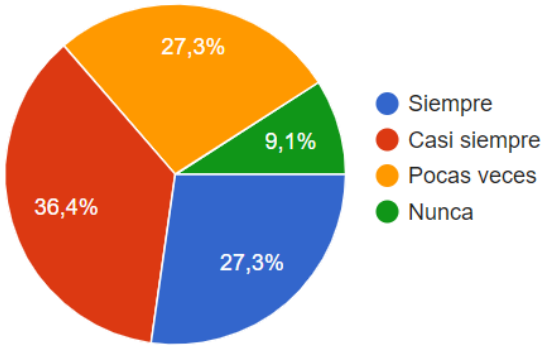

Gráfico 33. *Análisis de logs en línea o fuera de línea*

Los encuestados indican en un 27.3% siempre analizan ya sea en línea o fuera de línea las ocurrencias encontradas en sus sistemas, un 36.4% indica que casi siempre. Una buena práctica sugerida por los organismos de Seguridad Informática, es que el Administrador de TI encargado de velar por la seguridad de los sistemas, debe estar en constante vigilancia y análisis de los logs encontrados, a pesar de ser mucha la información, esta debe tenerse bajo vigilancia, situación que muchas veces no ocurre.

Pregunta 16: ¿Ha realizado actividades de seguimiento de los logs de seguridad?

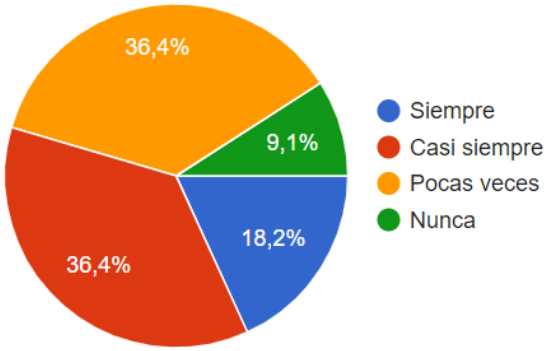

Gráfico 34. *Seguimiento a los logs de seguridad*

En un porcentaje de 54.6% los encuestados realizan seguimiento a los logs de seguridad; pero 18.2% nunca lo hacen, esto es un riesgo para sus sistemas puesto que a través de esta práctica se pueden detectar casos anómalos y situaciones que de pronto sus sistemas de seguridad en la red no los están controlando.

Pregunta 17: ¿Contempla un plan de instalaciones de software en la universidad?

Tabla 55.
Existe plan de instalaciones de software

	Frecuencia	Porcentaje
Si	9	81,8%
No	2	18,2%
Total	11	100,0%

El 81.8% de los encuestados indican que la universidad tiene dentro de sus políticas tener un plan de instalación de software. Es muy importante tener programada esta actividad para que se tenga controlado el software y actualizaciones vinculadas.

Pregunta 18: ¿Se siguen procedimientos o normas establecidos por el departamento de TI en las instalaciones de software por parte de los usuarios?

Tabla 56.
Se siguen procedimientos en las instalaciones de software

	Frecuencia	Porcentaje
Casi siempre	5	45,5
Pocas veces	3	27,3
Siempre	2	18,2
Nunca	1	9,1
Total	11	100,0

Si bien en la pregunta anterior indican que existe un plan de instalación de software, con esta pregunta se comprueba que no siempre se siguen los procedimientos o normas que establece el área de TI, solo el 18.2% lo aplica siempre. Esto indica que el nivel de concientización en aspectos de seguridad no se aplica.

Pregunta 19: De los siguientes servicios ¿cuál/cuáles de ellos utiliza más en sus labores de oficina?

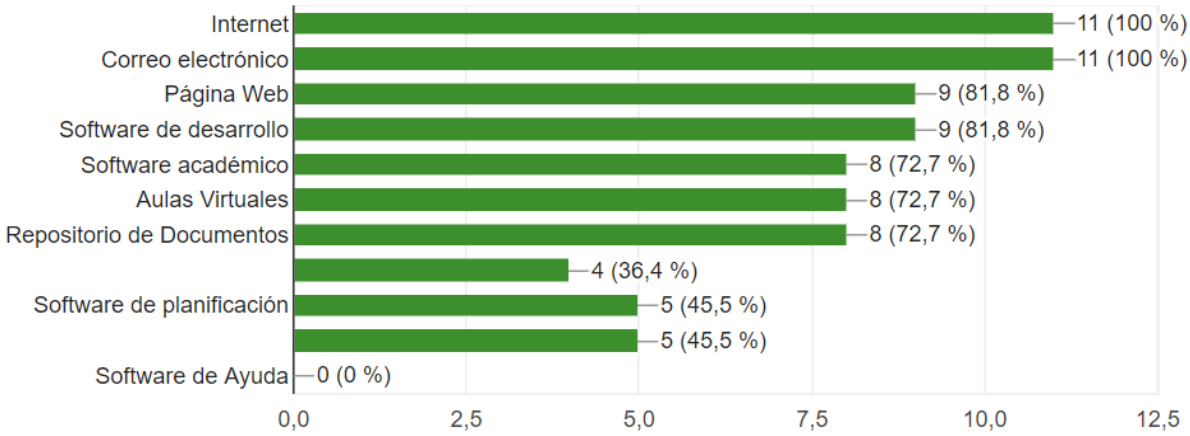

Gráfico 35. *Uso De Servicios En La Universidad*

Los servicios más utilizados en el campus universitario al 100% son Internet y correo electrónico por el total de encuestados, seguido de páginas web y software de desarrollo con un 81.8%, seguidamente se utiliza software académico, aulas virtuales y repositorio de documentos. Esto nos ayuda a contrastar uno de los indicadores de la Variable 1, que indica que los accesos desde fuera de la universidad son los de mayor frecuencia y estos muchas veces son ataques.

Pregunta 20: ¿Ha existido suspensión de algún servicio que usted necesita para realizar su trabajo diario de oficina?

	Frecuencia	Porcentaje
No	7	63,6%
Si	4	36,4%
Total	11	100,0%

Tabla 57. *Existencia de suspensión del servicio utilizado*

Los encuestados manifiestan que ha existido suspensión de servicio en un 36.4% y no existe un 63.6%. lo que puede indicar que se tiene implementado la tolerancia a fallas en la infraestructura de red del campus universitario.

Pregunta 21: ¿De la pregunta anterior señale con qué frecuencia ha sufrido la pérdida de este servicio?

Tabla 58.
Frecuencia de pérdida de suspensión del servicio

	Frecuencia	Porcentaje
Muy poco	4	36,4%
No precisa	3	27,3%
Poco	3	27,3%
Frecuentemente	1	9,1%
Total	11	100,0%

Los administradores de seguridad indican que el servicio que utilizan para sus labores cotidianas se suspende Muy poco (36.4%), poco y frecuentemente un 36.6%, no precisan un 27.3% que deben ser las personas que indicaron en la pregunta anterior que no ha existido suspensión del servicio.

Pregunta 22: ¿Tiene implementados controles de detección de códigos maliciosos en sus dispositivos de defensa para cada uno de estos ataques?

Tabla 59.
Implementación de controles de detección de código malicioso

	Frecuencia	Porcentaje
Casi siempre	5	45,5%
Siempre	3	27,3%
Pocas veces	2	18,2%
Nunca	1	9,1%
Total	11	100,0%

Del total de encuestados, el 27.3% (3) siempre tiene implementado algún mecanismo que **detecta** códigos maliciosos, un 45.5% casi siempre lo implementa y pocas veces el 18.2%.

Pregunta 23: ¿Tiene implementados controles de prevención de códigos maliciosos en sus dispositivos de defensa para cada uno de estos ataques?

Tabla 60.
Implementación de controles de prevención de código malicioso

	Frecuencia	Porcentaje
Pocas veces	5	45,5%
Casi siempre	3	27,3%
Siempre	2	18,2%
Nunca	1	9,1%
Total	11	100,0%

La implementación de controles que **previenen** este tipo de ataques es de solo 18.2%, pocas veces lo implementa un 45.5% y casi siempre un 27.3%. Este resultado nos ayuda a contrastar también las pruebas previas de minería de datos que indican que el ataque más común detectado es "code-execution" un tipo de código malicioso.

Pregunta 24: ¿Tiene implementados controles de protección de códigos maliciosos en sus dispositivos de defensa para cada uno de estos ataques?

Tabla 61.
Implementación de controles de protección de código malicioso

	Frecuencia	Porcentaje
Pocas veces	4	36,4%
Casi siempre	3	27,3%
Siempre	2	18,2%
Nunca	2	18,2%
Total	11	100,0%

Si bien por un lado se tiene siempre implementados controles de **protección** de código malicioso 18.2%, en igual proporción, nunca se implementan, lo que significa un riesgo para la seguridad informática de la universidad. Que pocas veces o casi siempre 63.7% lo apliquen significa que pueden existir los mecanismos, pero no lo utilizan.

Pregunta 25: ¿Hace un análisis detallado de los tipos de tráfico que entran y salen de su perímetro universitario?

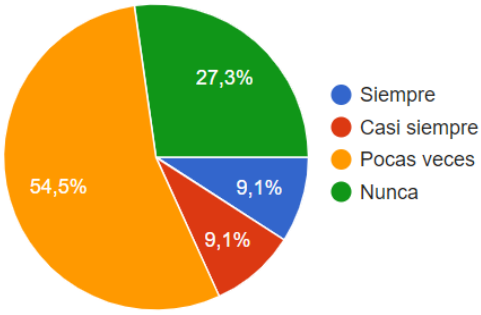

Gráfico 36. *Análisis detallado de los tipos de tráfico*

Muchas veces los esfuerzos por implementar mecanismos de defensa para proteger el perímetro universitario, no tienen efecto positivo porque no se analiza la data que pueda capturarse a traves de snifers o firewalls, esto se verifica en esta pregunta dado que pocas veces 54.5% se hace este análisis detallado del tráfico que atraviesa las redes informáticas. Solo un 9.1% (1 persona) siempre lo hace y nunca lo hace un 27.3%. Este fue el motivo de estudio de la presente investigación.

Pregunta 26: ¿Se realizan auditorias y actividades de verificación de los sistemas para minimizar la interrupción de los procesos?

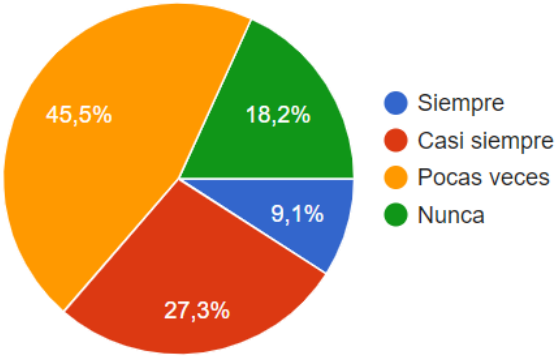

Gráfico 37. *Se realizan auditorias y verificación de sistemas*

Es una buena práctica que se realice periódicamente auditorias de sistemas y verificación de ellos para encontrar falencias y minimizar la interrupción de los procesos informáticos en cualquier entorno. Aquí se observa que pocas veces se realiza en los campus universitarios 45.5% que muy superior a hacerlo siempre y casi siempre con 36.4%.

Pregunta 27. ¿Somete a pruebas de penetración externas a sus sistemas de defensa perimetral?

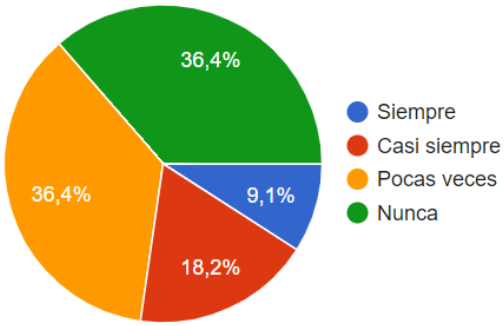

Gráfico 38. *Se realizan pruebas de penetración externas*

Los encuestados indican que "Nunca" realizan pentesting 36.4% y pocas veces 36.4% que hacen un 72.8%. Entonces no se podrá saber si sus políticas de seguridad implementadas son efectivas, no se puede determinar si sus sistemas son vulnerables a los ataques, no se puede determinar si sus defensas son suficientes o ya fueron vulneradas. Es una buena práctica que sugieren los estándares de seguridad informática.

Pregunta 28: ¿Se realiza monitoreo de acceso a la red universitaria?

Tabla 62.
Realización de monitoreo de acceso a la Universidad

	Frecuencia	Porcentaje
Pocas veces	7	63,6
Casi siempre	2	18,2
Siempre	2	18,2
Total	11	100,0

Pocas veces 63.6% los administradores realizan el monitoreo a sus redes informáticas, casi siempre y siempre 18.2% respectivamente.

Pregunta 29: ¿Se realiza monitoreo al firewall principal de la universidad?

Tabla 63.
Realización de monitoreo al firewall de la Universidad

	Frecuencia	Porcentaje
Casi siempre	4	36,4%
Pocas veces	4	36,4%
Siempre	2	18,2%
Nunca	1	9,1%
Total	11	100,0

Casi siempre indican que se monitorea el firewall de la universidad 36.4%, pocas veces lo hacen 36.4%, siempre monitorean 18.2% y nunca solo una persona lo indica así.

Pregunta 30: ¿Se realiza el monitoreo de cuentas privilegiadas a los sistemas y servidores?

Tabla 64.
Realización de monitoreo a cuentas privilegiadas

	Frecuencia	Porcentaje
Casi siempre	4	36,4%
Siempre	4	36,4%
Pocas veces	3	27,3%
Total	11	100,0%

Casi siempre, un 36.4% por lo indica que se hace monitoreo a las cuentas de acceso a los servidores, en igual porcentaje lo hacen siempre, lo que indica una buena práctica para la seguridad informática de la Universidad.

Pregunta 31: ¿Se realiza el monitoreo del volumen de tráfico para identificar el uso indebido de los recursos de la universidad?

Tabla 65.
Realización de monitoreo al tráfico de la red

	Frecuencia	Porcentaje
Pocas veces	4	36,4
Casi siempre	3	27,3
Siempre	3	27,3
Nunca	1	9,1
Total	11	100,0

El monitoreo del volumen del tráfico de la universidad que identifique su uso indebido por parte de usuarios internos se realiza pocas veces 36.4% y casi siempre 27.3%.

Pregunta 32: ¿Supervisa los puertos comunes para los protocolos que permiten sesiones remotas?

Tabla 66.
Supervisión de puertos comunes para sesiones remotas

	Frecuencia	Porcentaje
Casi siempre	5	45,5
Pocas veces	4	36,4
Siempre	2	18,2
Total	11	100,0

Es una buena actividad de seguridad, la supervisión de puertos conocidos para detectar protocolos que permiten sesiones remotas de usuarios autorizados o no a la Universidad, la tabla nos indica que casi siempre se realiza esta actividad con 45.5% y pocas veces 36.4%.

Pregunta 33: ¿Ha recibido en el último mes, algún ataque informático a sus redes de datos en general?

Tabla 67.
Recepción de ataque en el último mes

	Frecuencia	Porcentaje
No	7	63,6%
Si	4	36,4%
Total	11	100,0%

Indican que en el último mes si han recibido ataques informáticos un 63.3% y no lo han percibido o recibido 36.4%

Pregunta 34: ¿Ha indagado que tipo de ataques son más frecuentes en las redes WLAN de su universidad?

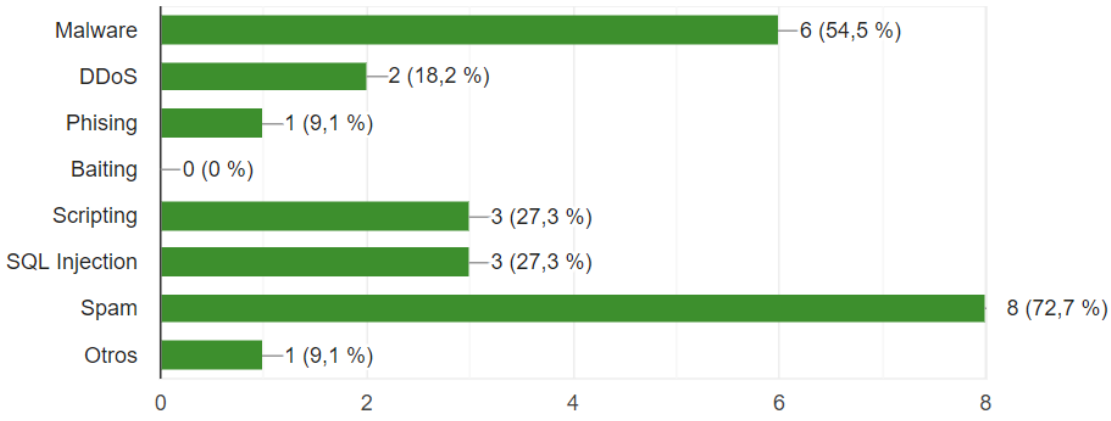

Gráfico 39. *Tipo de ataque frecuente en WLAN*

El ataque más frecuente que se ha detectado en las redes inalámbricas del campus universitario son de la categoría spam con 72.7%, seguido del malware con 54.5%, lo que puede sugerir al departamento de TI que sus controles no están siendo eficientes.

Pregunta 35: ¿Aproximadamente cuantos ataques ha detectado?

Tabla 68.
Cantidad de ataques detectados

	Frecuencia	Porcentaje
Entre 1 y 30	10	90,9
Entre 31 y 60	1	9,1
Total	11	100,0

Los encuestados han detectado entre 1 y 30 ataques, que hace un 90.9% del total percibido.

Pregunta 36: ¿Cómo clasificaría a los ataques potenciales detectados?

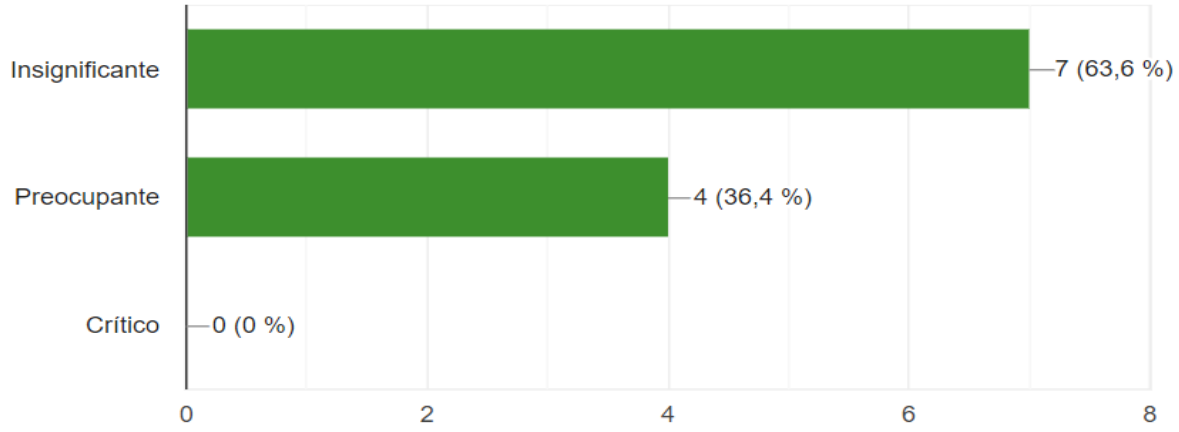

Gráfico 40. *Clasificación del ataque detectado*

El encuestado clasifica el ataque que detectó como insignificante, que es 63.6% y preocupante 36.4%, se tendría que hacer un análisis más detallado para saber cuál específicamente fue y el nivel de daño que ocasionó el mismo para catalogarlo en este nivel.

Pregunta 37: ¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?

Tabla 69.
Aplicación de Plan de seguridad ISO 27002

	Frecuencia	Porcentaje
Si	6	54,5
No	5	45,5
Total	11	100,0

Los encuestados en 54.5%, indicaron que se aplica en su entidad un plan de seguridad informática basada en el estándar ISO 27002, pero casi un número similar indicó que no 45.5%. Sería aconsejable que el total ya pueda indicar que si se aplica.

ANALISIS E INTERPRETACION

Para el análisis y validación de la tercera hipótesis específica, se procedió a evaluar los resultados del análisis estadístico, se utilizó la prueba de Chi Cuadrado, considerando al Plan de seguridad de la información en el departamento de TI basado en estándares ISO 27002 como variable dependiente y las demás preguntas pertinentes de la Ficha de Recolección de datos aplicado a los Administradores y Oficiales de seguridad de TI, como independientes.

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Uso de protocolos de cifrado en las comunicaciones inalámbricas

Tabla 70.
Plan de seguridad de la Información según ISO 27002 por Protocolos de cifrado en Wifi

	Aplicación de Plan de seguridad de la información en base a ISO 27002		
	No	Si	Total
Casi Siempre	0(0,0%)	2(33,3%)	2(18,2%)
Nunca	1(20,0%)	1(16,7%)	2(18,2%)
Pocas veces	2(40,0%)	0(0,0%)	2(18,2%)
Siempre	2(40,0%)	3(50,0%)	5(45,5%)
Total	5(100,0%)	6(100,0%)	11(100,0%)

Interpretación: En las personas que participan de este estudio se ha encontrado que de los que aplican el plan de seguridad de la información en base a la ISO 27002, 33.3% casi siempre aplica el plan de seguridad, 40% siempre aplica. De los que han declarado que no aplican el Plan de seguridad, 40% aplican siempre y 40% pocas veces. Las ligeras diferencias entre los que aplican y no aplican el plan, hacen que las diferencias entre ambos nos sean estadísticamente significativas entre si (χ^2: 0.246; p=246).

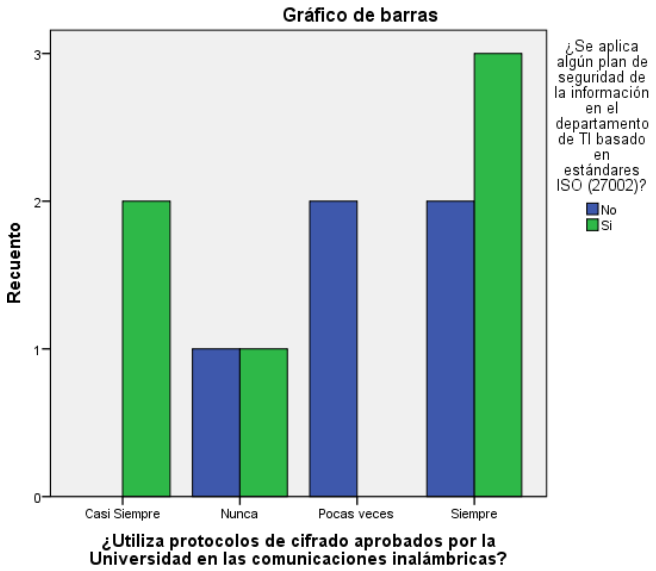

Gráfico 41. *Plan de seguridad de la Información según ISO 27002 por Protocolos de cifrado en Wifi*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Tipo de protocolos usados en comunicaciones inalámbricas

Tabla 71.
Plan de seguridad de la Información según ISO 27002
Por Tipo de protocolos usados en comunicaciones inalámbricas

	Aplicación de Plan de seguridad de la información en base a ISO 27002		Total
	No	Si	
Ninguno	0(0,0%)	1(16,7%)	1(9,1%)
WEP	1(20,0%)	0(0,0%)	1
WPA-PSK (AES)	1(20,0%)	0(0,0%)	1(9,1%)
WPA-PSK (TKIP), WPA-PSK (AES), WPA2-PSK (TKIP)	1(20,0%)	0(0,0%)	1(9,1%)
WPA2-PSK (AES)	0(0,0%)	1(16,7%)	1(9,1%)
WPA2-PSK (AES), WPA2, WEP	0(0,0%)	1(16,7%)	1(9,1%)
WPA2-PSK (TKIP)	1(20,0%)	2(33,3%)	3(27,3%)
WPA2-PSK (TKIP), WPA2-PSK (AES)	0(0,0%)	1(16,7%)	1(9,1%)
WPA3	1(20,0%)	0(0,0%)	1(9,1%)
TOTAL	5(100,0%)	6(100,0%)	11(100,0%)

Interpretación: Las diferencias entre los que aplican y no aplican el Plan de seguridad en base al ISO 27002 varía. Se tiene que los protocolos que usan los que no aplican el plan son WEP, WPA-PSK (TKIP), WPA-PSK (AES), WPA2-PSK (TKIP) homogéneamente. Los que aplican el plan en base al ISO 27002 usan el protocolo WPA2-PSK (TKIP) en 33.3%, uno de este grupo reportó que no usa protocolos. (χ^2: 8.311; p=0.4037). Se acepta Ho, no hay significancia.

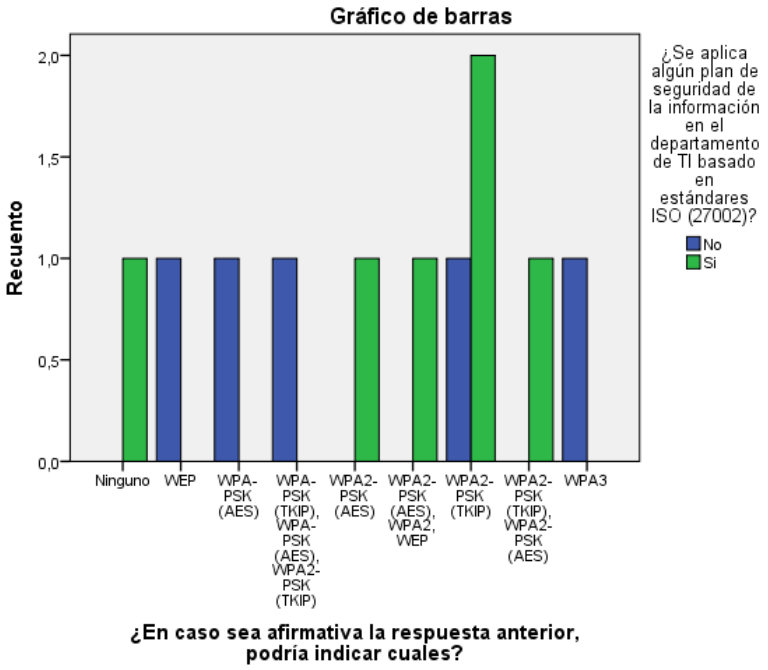

Gráfico 42. *Plan de seguridad de la Información según ISO 27002 por Tipo de protocolos usados en comunicaciones inalámbricas*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Uso de contraseñas o autenticación para usuarios con acceso remoto

Tabla 72.
Plan de seguridad de la Información según ISO 27002
Por Uso de contraseñas o autenticación para usuarios con acceso remoto

Aplicación de Plan de seguridad de la información en base a ISO 27002

	No	Si	Total
No	1(20,0%)	0(0,0%)	1(9,1%)
Si	4(80,0%)	6(100,0%)	10(90,9%)
Total	5(100,0%)	6(100,0%)	11(100,0%)

Interpretación: El 100% de los que aplican el plan de seguridad utiliza contraseñas para la autenticación, de los que no aplican el plan de seguridad el 80% también aplica contraseñas. Esta ligera diferencia hace que no exista diferencias estadísticamente significativas entre ambos grupos (χ^2: 0.009; p=0.924).

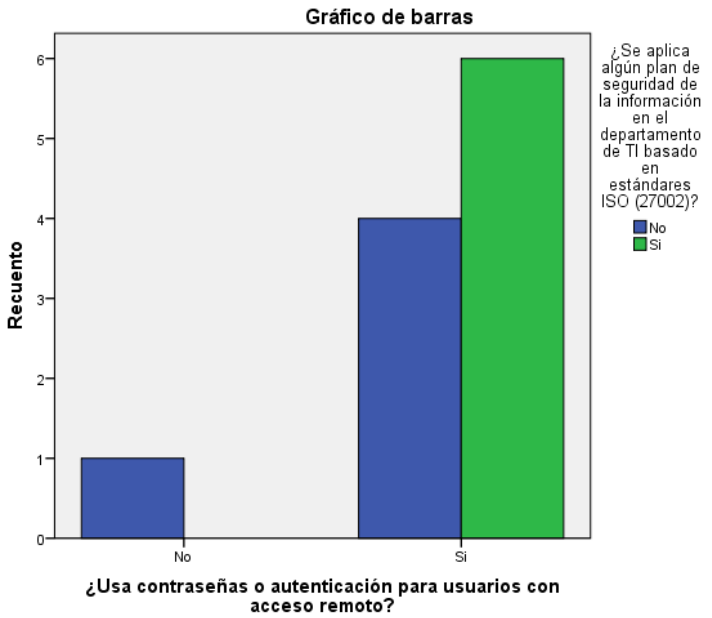

Gráfico 43. *Plan de seguridad de la Información según ISO 27002 por Uso de contraseñas o autenticación para usuarios con acceso remoto*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Identificación de niveles de acceso de usuarios a la red universitaria y bloqueo de intentos de Superusuario o Administrador

Tabla 73.
Plan de seguridad de la Información según ISO 27002 por Identificación de niveles de acceso de usuarios a la red universitaria y bloqueo de intentos de Superusuario o Administrador

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		
	No	Si	Total
Casi siempre	3(60,0%)	3(50,0%)	6(54,5%)
Siempre	2(40,0%)	3(50,0%)	5(45,5%)
TOTAL	5(100,0%)	6(100,0%)	11(100,0%)

Interpretación: Se ha encontrado que todos los encuestados cumplen casi siempre o siempre en la identificación de los niveles de acceso de usuarios a la red y bloqueos de intentos de superusuario, esto no tiene ninguna relación estadísticamente significativa en comparación de aplicación del plan de seguridad basado en estándares ISO 27002 (χ^2: 0.08; p=0.773).

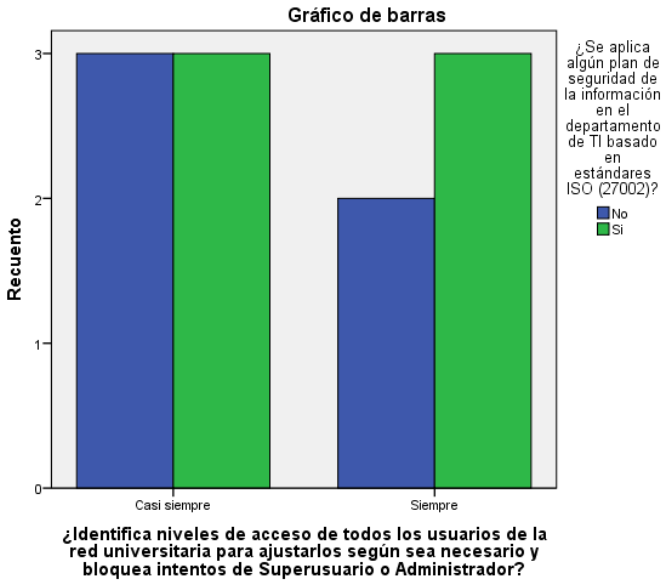

Gráfico 44. *Plan de seguridad de la Información según ISO 27002 por Identificación de niveles de acceso de usuarios a la red universitaria y bloqueo de intentos de Superusuario o Administrador*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Política de cambio de contraseña cada cierto periodo de tiempo

Tabla 74.
Plan de seguridad de la Información según ISO 27002
Por Política de cambio de contraseña cada cierto periodo de tiempo

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		
	No	Si	Total
No	2(40,0%)	2(33,3%)	4(36,4%)
Si	3(60,0%)	4(66,7%)	7(63,6%)
TOTAL	5(100,0%)	6(100,0%)	11(100,0%)

Interpretación: La política de cambio de contraseña cada periodo de tiempo es similar de acuerdo con la aplicación del Plan de Seguridad, los que sí tienen la política en el cambio de contraseño en los que no aplican el Plan de seguridad fue del 60% y en los que si aplican fue del 66.7% (χ^2: 0.16; p=0.68). No existiendo diferencia significativa entre ellos.

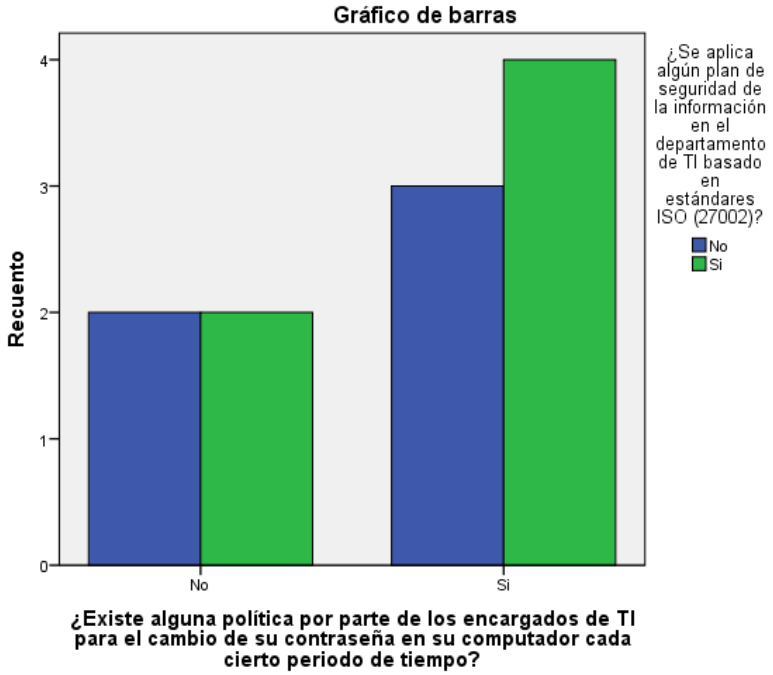

Gráfico 45. *Plan de seguridad de la Información según ISO 27002 por Política de cambio de contraseña cada cierto periodo de tiempo*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Frecuencia de cambio de contraseña.

Tabla 75.
Plan de seguridad de la Información según ISO 27002
Por Frecuencia de cambio de contraseña

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		
	No	Si	Total
No especifica	2(40,0%)	2(33,3%)	4(36,4%)
Cada año	1(20,0%)	0(0,0%)	1(9,1%)
Cada seis meses	2(40,0%)	2(33,3%)	4(36,4%)
Cada tres meses	0(0,0%)	2(33,3%)	2(18,2%)
TOTAL	5(100,0%)	6(100,0%)	11(100,0%)

Interpretación: La frecuencia del cambio de contraseña no han sido influenciada o asociada a la aplicación del Plan de Seguridad (χ^2: 2.93; p=0.4025). Aunque se ha logrado observar ligeras diferencias, como en la que el 33.3% de los que si aplica el Plan hace el cambio cada tres meses, a diferencia de los que no aplican en donde ninguno hace el cambio cada tres meses.

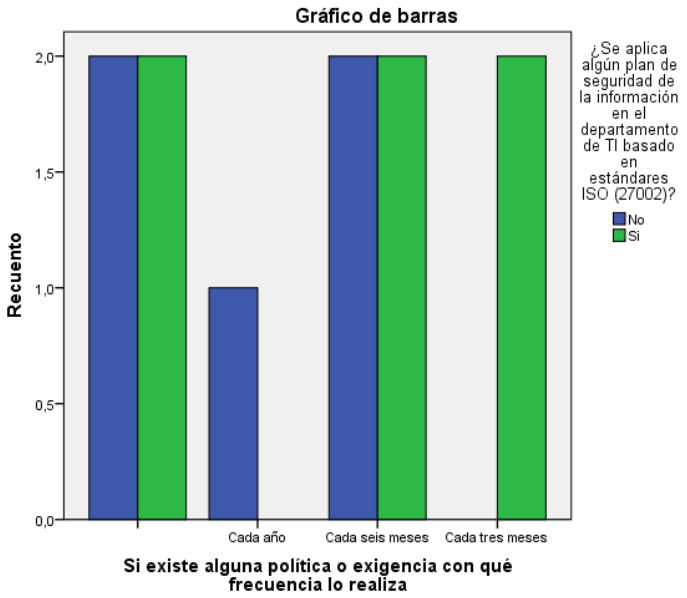

Gráfico 46. *Plan de seguridad de la Información según ISO 27002 por Frecuencia de cambio de contraseña*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Desactivación de credenciales de trabajadores en coordinación con RRHH.

Tabla 76.
Plan de seguridad de la Información según ISO 27002 por Desactivación de credenciales de trabajadores en coordinación con RRHH. Frecuencia de cambio de contraseña

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Casi siempre	1(20,0%)	4(66,7%)	5(45,5%)
Pocas veces	2(40,0%)	0(0,0%)	2(18,2%)
Siempre	2(40,0%)	(33,3%)	4(36,4%)
Total	5(100,0%)	6(100,0%)	11(100,0%)

Interpretación: La desactivación de credenciales se da en un 40% en los que no aplican el Plan de Seguridad, esto contrasta con el 0% de los que si aplican el Plan. Complementado con el 100% de los que aplican el Plan que casi siempre o siempre desactivan las credenciales. Sin embargo, estos valores igual nos indican que no existen diferencias estadísticamente significativas (χ^2: 3.74; p=0.154).

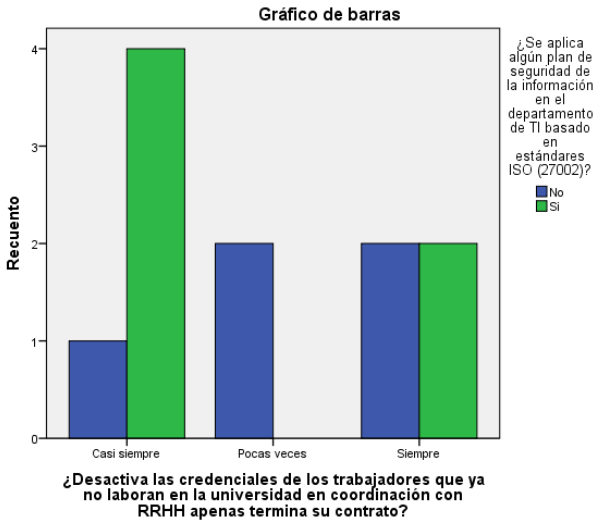

Gráfico 47. *Plan de seguridad de la Información según ISO 27002 por Desactivación de credenciales de trabajadores en coordinación con RRHH. Frecuencia de cambio de contraseña*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Existencia de procedimientos en caso de ocurrir problemas con los servidores, computadoras o servicios informáticos.

Tabla 77.
Plan de seguridad de la Información según ISO 27002 por Existencia de procedimientos en caso de ocurrir problemas con los servidores, computadoras o servicios informáticos

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		
	No	Si	Total
Casi siempre	1(20,0%)	2(33,3%)	3(27,3%)
Nunca	1(20,0%)	0(0,0%)	1(9,1%)
Pocas veces	1(20,0%)	1(16,7%)	2(18,2%)
Siempre	2(40,0%)	3(50,0%)	5(45,5%)
Total	5(100,0%)	6(100,0%)	11(100,0%)

Interpretación: 60% de los que no aplican el Plan de Seguridad basado en la ISO 27002 saben (siempre/casi siempre) de la existencia de procedimientos en caso de ocurrir problemas informáticos, ligeramente por debajo del 83.3% de los que si aplican el Plan. Sin embargo, no existen diferencias estadísticamente significativas entre ambos grupos (χ^2: 1.45; p=0.6939).

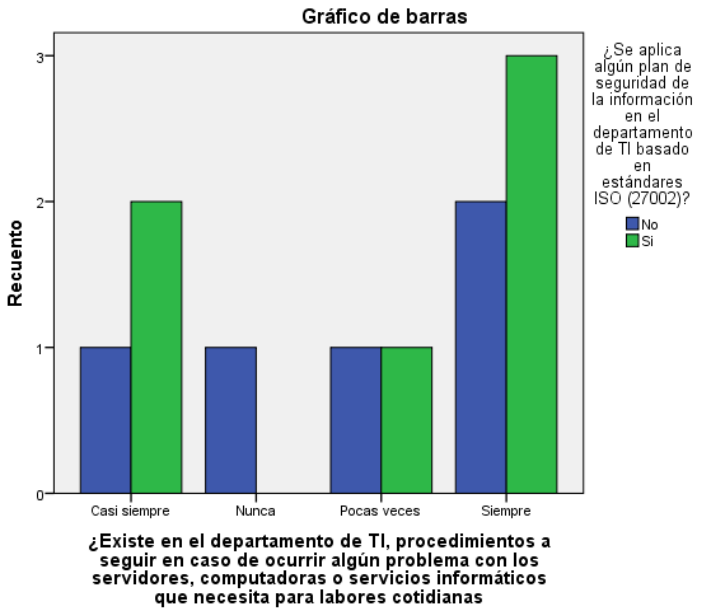

Gráfico 48. *Plan de seguridad de la Información según ISO 27002 por Existencia de procedimientos en caso de ocurrir problemas con los servidores, computadoras o servicios informáticos*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Existencia de política que indique en qué periodo tiempo se debe respaldar la información de su equipo.

Tabla 78.
Plan de seguridad de la Información según ISO 27002 por Existencia de política que indique en qué periodo tiempo se debe respaldar la información de su equipo

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		
	No	Si	Total
No	3(60,0%)	2(33,3%)	5(45,5%)
Si	2(40,0%)	4(66,7%)	6(54,5%)
Total	5(100,0%)	6(100,0%)	11(100,0%)

Interpretación: Al consultar sobre la presencia de políticas que nos indiquen el periodo de tiempo para respaldar la información y asociarlo con la aplicación del Plan de Seguridad basado en la ISO 27002 no se ha encontrado relación estadísticamente significativa (χ^2: 0.076; p=0.782). Las frecuencias en ambos casos son muy similares entre sí.

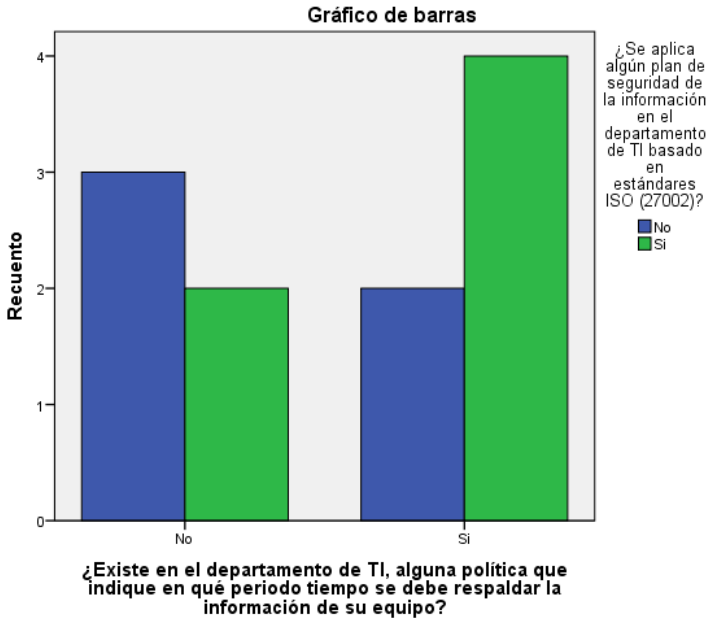

Gráfico 49. *Plan de seguridad de la Información según ISO 27002 por Existencia de política que indique en qué periodo tiempo se debe respaldar la información de su equipo*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por periodicidad de respaldo de la información.

Tabla 79.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Periodicidad de respaldo de la información

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		
	No	Si	Total
No especifica	1(20,0%)	1(16,7%)	2(18,2%)
Cada 3 meses	1(20,0%)	0(0.0%)	1
Cada 6 meses	1(20,0%)	1(16,7%)	2(18,2%9
Cada mes	2(40,0%)	4(66,7%)	6(54,5%)
Total	5(100,0%)	6(100,0%)	11(100,0%)

Interpretación: La periodicidad del respaldo de la información cada mes es ligeramente superior en los casos en los que se aplica el Plan de Seguridad basado en estándares ISO 27002 (66.7% Por 40%) (χ^2: 1.589; p=0.662). Se debe considerar que en ambas situaciones comparativas han tenido un caso cada una sin especificar.

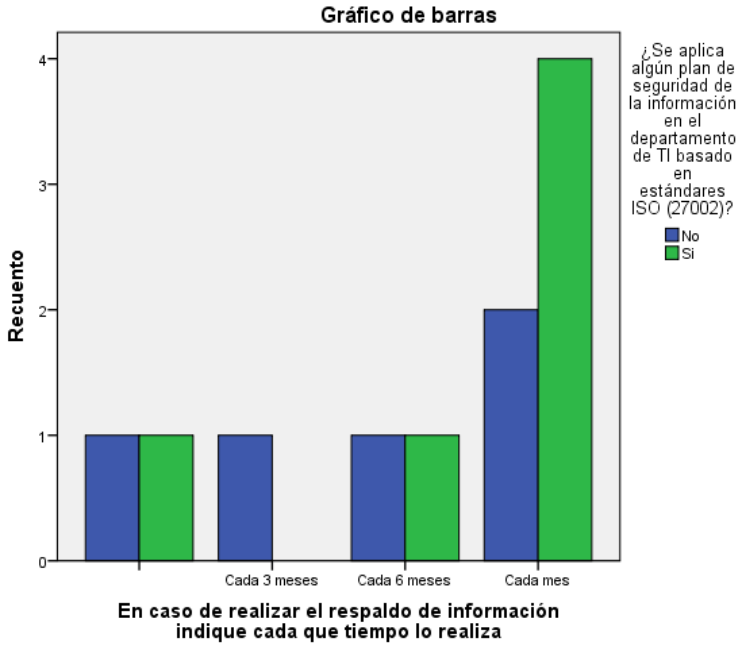

Gráfico 50. *Plan de seguridad de la Información según ISO 27002 por Periodicidad de respaldo de la información*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por los backups y archivos con data sensible están cifrados.

Tabla 80. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Los backups y archivos con data sensible están cifrados*

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		
	No	Si	Total
Casi siempre	2(40,0%)	1(16,7%)	3(27,3%)
Nunca	1(20,0%)	1(16,7%)	2(18,2%)
Poca veces	2(40,0%)	2(33,3%)	4)36,4%)
Siempre	0(0,0%)	2(33,3%)	2(18,2%9
Total	5(100,0%)	6(100,0%)	11(100,0%)

Interpretación: La frecuencia del cifrado de los backups y archivos con data sensible siempre se ha dado en 33.3% los casos que aplican el Plan de Seguridad basado en la ISO 27002, a diferencia del grupo en los que no se aplica el plan en donde mayoritariamente se reportó casi siempre (40%) o pocas veces (40%) el cifrado. Cada uno tuvo un caso en los que nunca se han cifrado la información en mención (χ^2: 2.261; p=0.52).

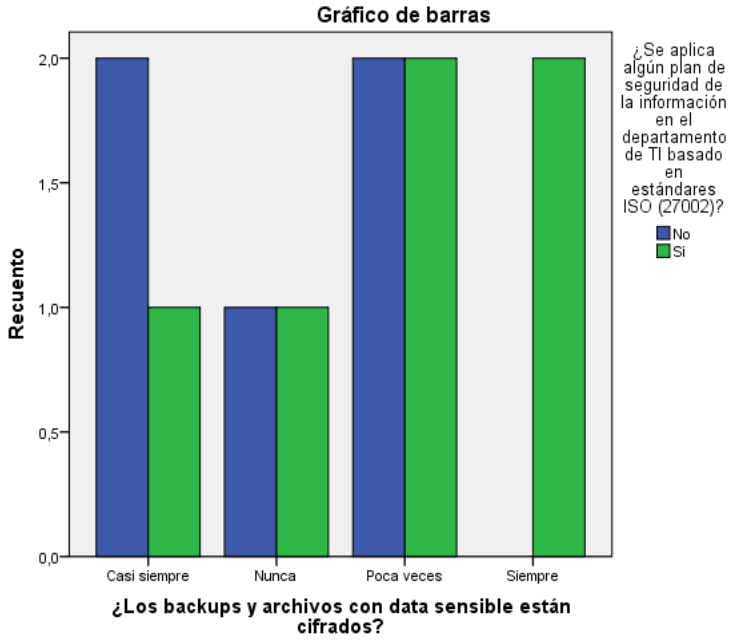

Gráfico 51. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Los backups y archivos con data sensible están cifrados*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por aplicación de parches en servidores y estaciones de trabajo.

Tabla 81.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 Por Aplicación de parches en servidores y estaciones de trabajo

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		
	No	Si	Total
Casi siempre	1(20,0%)	(66,7%)4	5(45,5%)
Nunca	0(0,0%)	1(16,7%)	1(9,1%)
Pocas veces	2(40,0%)	0(0,0%)	2(18,2%)
Siempre	2(40,0%)	1(16,7%)	3(27,3%)
TOTAL	5(100,0%)	6(100,0%)	11(100,0%)

Interpretación: La aplicación de parches en servidores y estaciones de trabajo en los que aplican el Plan de seguridad basado en la ISO 27002 se dio siempre/casi siempre en el 83.4%, en los que no aplican el Plan se ha dado en el 60% (χ^2: 5.08; p=0.166). Hubo un caso en el que nunca se aplicó el parche, y correspondió al grupo que aplica el Plan de Seguridad.

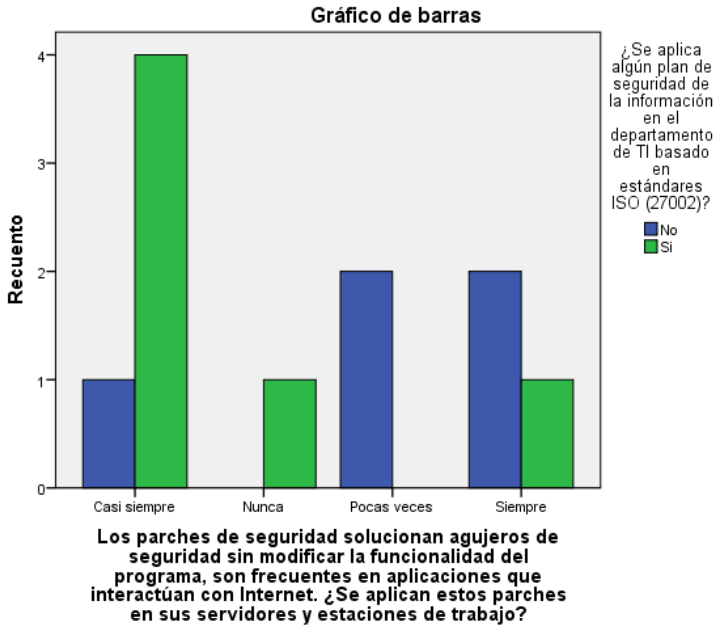

Gráfico 52. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Aplicación de parches en servidores y estaciones de trabajo*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 Por tipo de parches configurados

Tabla 82.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 Por Tipo de parches configurados

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		
	No	Si	Total
	0(0,0%)	1(16,7%)	1(9,1%)
Parches al código fuente, Parches de depuración, Parches de seguridad	0(0,0%)	1(16,7%)	1(9,1%)
Parches al código fuente, Parches de depuración, Parches de seguridad, Parches de actualización	0(0,0%)	1(16,7%)	1(9,1%)
Parches al código fuente, Parches de seguridad, Parches de actualización	1(20,0%)	0(0,0%)	1(9,1%)
Parches de actualización	1(20,0%)	0(0,0%)	1(9,1%)
Parches de actualización, Parches de piratería ilegal	1(20,0%)	0(0,0%)	1(9,1%)
Parches de depuración, Parches de actualización, Parches de traducción	0(0,0%)	1(16,7%)	1(9,1%)
Parches de seguridad	0(0,0%)	1(16,7%)	1(9,1%)
Parches de seguridad, Otros	1(20,0%)	0(0,0%)	1(9,1%)
Parches de seguridad, Parches de actualización	1(20,0%)	1(16,7%)	2(18,2%)
Total	5(100,0%)	6(100,0%)	11(100,0%)

Interpretación: Los diferentes tipos de parches tuvo una gran dispersión entre los que aplicaban y los que no aplicaban el Plan de Seguridad basado en la ISO 27002. En ninguno de los casos hubo la presencia del uso del parche en más de dos individuos evaluados (χ^2: 8.983; p=0.439).

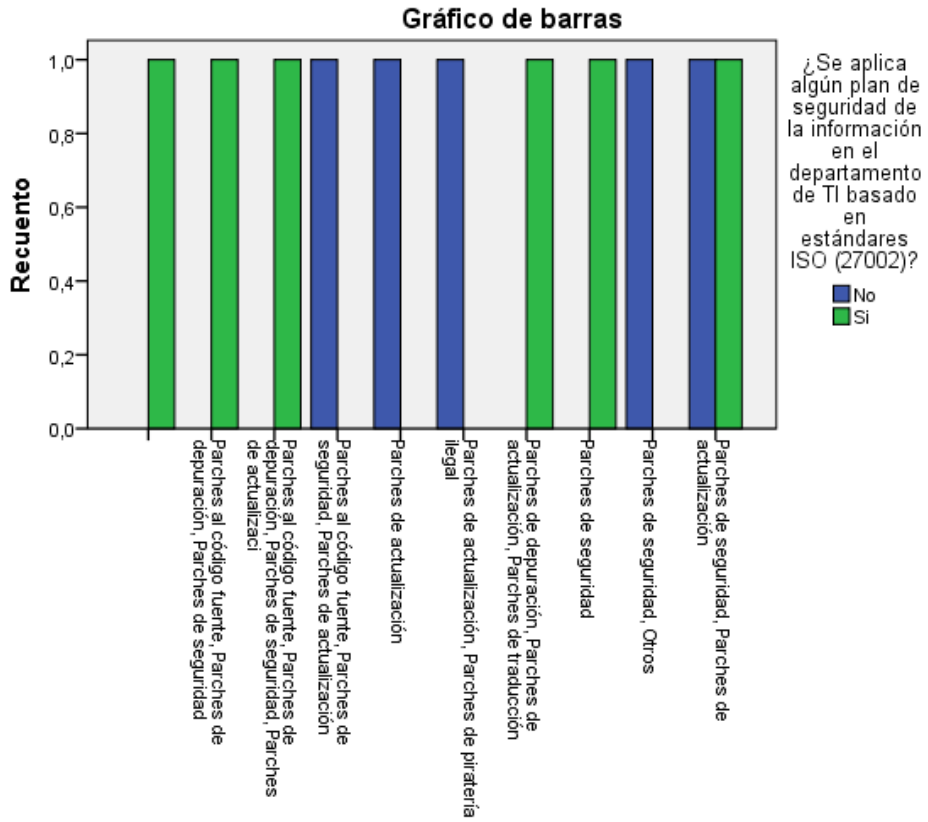

Gráfico 53. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Tipo de parches configurados*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por registro de los logs de eventos de seguridad, actividades de usuarios, excepciones, fallas

Tabla 83.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002
Por Registro de los logs de eventos de seguridad, actividades de usuarios, excepciones, fallas

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		
	No	Si	Total
Siempre	1 (20%)	3 (50%)	4 63.4%)
Casi siempre	2 (40%)	1 (16.1%)	3 (27.3%)
Pocas veces	2 (40%)	1 (16.7%)	3 (27.3%)
Nunca	0 (0%)	1 (16.7%)	1 (9.1%)
TOTAL	5 (100%)	6 (100%)	11 (100%)

Interpretación: En lo que respecta el registro de los logs de eventos de seguridad la distribución de su uso en cuanto a los que aplican y no aplican el Plan de Seguridad basado en la ISO 270002 no guardan diferencias estadísticas entre sí (χ^2: 8.983; p=0.439). se ha reportado un caso que nunca ha hecho el registro de los logs, este pertenece a los que si aplican el Plan de seguridad.

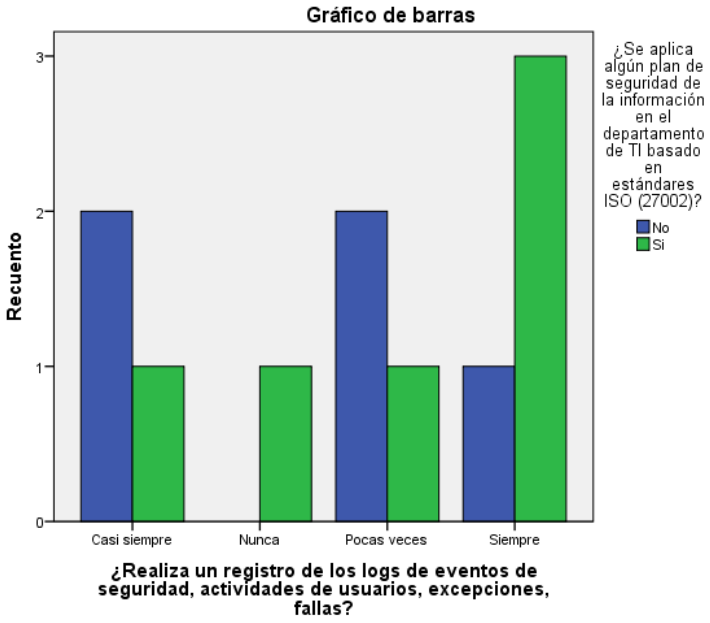

Gráfico 54. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Registro de los logs de eventos de seguridad, actividades de usuarios, excepciones, fallas*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Análisis de logs de seguridad en línea o fuera de línea

Tabla 84.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 Por Análisis de logs de seguridad en línea o fuera de línea

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Siempre	1 (20%)	2 (33.3%)	3 (27.3%)
Casi siempre	1 (20%)	3 (50%)	4 (36.4%)
Pocas veces	3 (60%)	0 (0%)	3 (27.3%)
Nunca	0 (0%)	1 (16.7%)	1 (9.1%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: La aplicación de los logs de seguridad se aplica siempre o casi siempre en 40% de los que no siguen el Plan de Seguridad basado en ISO 27002, mientras que en los que siguen el Plan se conforma el 83.3%. Sin embargo, existe un caso en los que nunca ha aplicado el log de seguridad y se encuentra dentro de los que refirieron que si aplican el Plan (χ^2: 5.286; p=0.152).

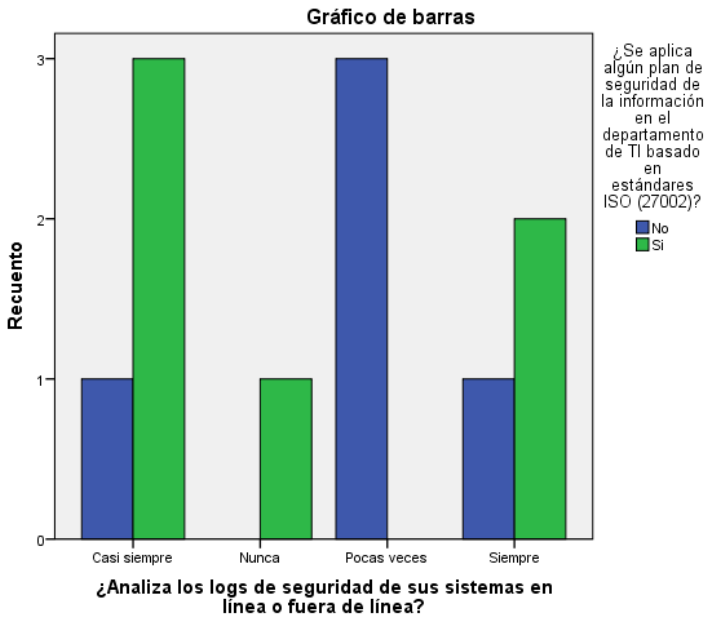

Gráfico 55. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Análisis de logs de seguridad en línea o fuera de línea*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Actividades de seguimiento de los logs de seguridad?

Tabla 85.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 Por Seguimiento de logs de seguridad en línea o fuera de línea

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Siempre	1 (20%)	1 (16.7%)	2 (18.2%)
Casi siempre	2 (40%)	2 (33.3%)	4 (36.4%)
Pocas veces	2 (40%)	2 (33.3%)	4 (36.4%)
Nunca	0 (0%)	1 (16.7%)	1 (9.1%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: No se han encontrado diferencias estadísticamente significativas entre los que hacen actividades para el seguimiento de los logs de seguridad y los que aplican el Plan de Seguridad (ISO 270002) (χ^2: 0.917; p=0.821). Se ha encontrado un caso de los que aplican el Plan que nunca ha realizado actividades de seguimiento de los logs.

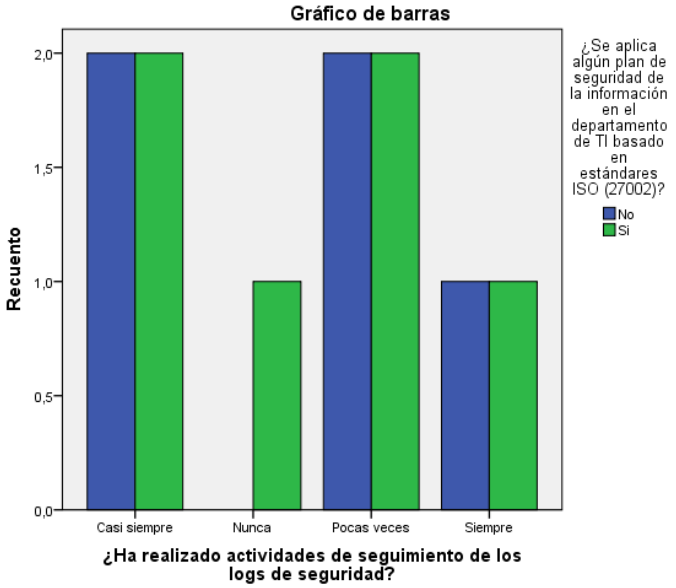

Gráfico 56. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Seguimiento de logs de seguridad en línea o fuera de línea*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por plan de instalaciones de software

Tabla 86.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Plan de instalaciones de software

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
No	2 (40%)	0 (0%)	2 (18.2%)
Si	3 (60%)	6 (100%)	9 (81.8%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: El 100% de los que aplican un plan de Seguridad (ISO 27002) ha indicado que contemplan un plan de instalaciones de software en la universidad, mientras que solo el 60% de los que no aplica el Plan ha contemplado el plan de instalaciones de software, esto nos puede dar indicios claros del uso, aplicabilidad e importancia del Plan de Seguridad (χ^2: 0.861; p=0.3549

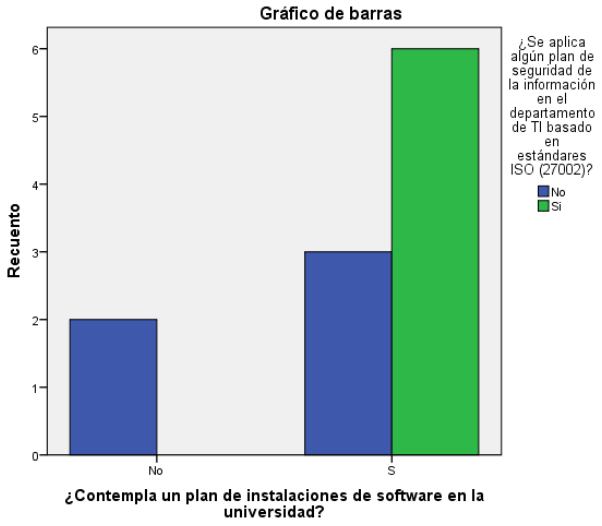

Gráfico 57. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Plan de instalaciones de software*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por procedimientos de instalación de software

Tabla 87.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Procedimientos de instalación de software

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Siempre	0 (0%)	2 (33.3%)	2 (18.2%)
Casi siempre	2 (40%)	3 (50%)	5 (45.5%)
Pocas veces	2 (40%)	1 (16.7%)	3 (27.3%)
Nunca	1 (20%)	0 (0%)	1 (9.1%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: 33% de los que aplican el Plan de Seguridad (ISO 27002) siempre siguen los procedimientos y normas establecidos por el departamento de TI de la Universidad. 20% de los que no aplican el Plan nunca han seguido los procedimientos o normas establecidas por el departamento de TI (χ^2: 3.471; p=0.325).

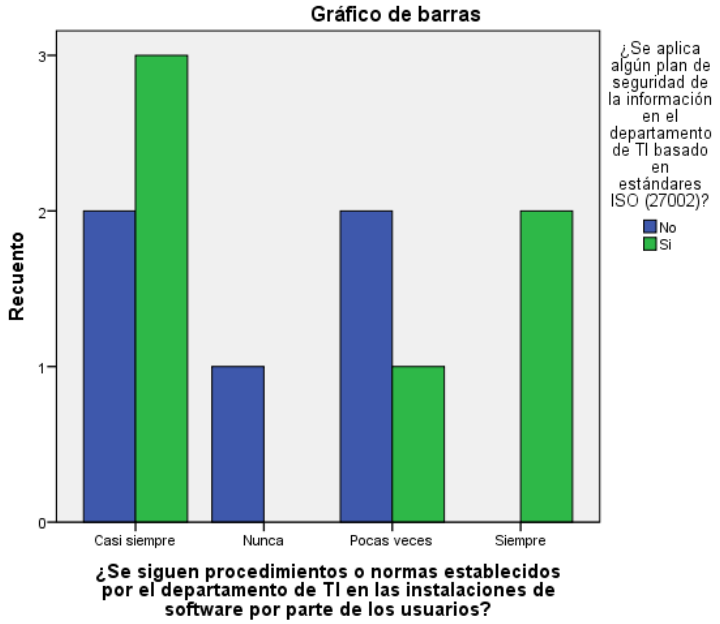

Gráfico 58. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Procedimientos de instalación de software*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Servicios utilizados en labores de oficina

Tabla 88.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 Por Servicios utilizados en labores de oficina

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Internet, Correo electrónico, Página Web, Aulas Virtuales	1 (20%)	0 (0%)	1 (9.1%)
Internet, Correo electrónico, Página Web, Software de desarrollo, Aulas Virtuales, Repositorio de Documentos, Software de Gestión de Bibliotecas	1 (20%)	0 (0%)	1 (9.1%)
Internet, Correo electrónico, Página Web, Software de desarrollo, Software académico, Aulas Virtuales	0 (0%)	1 (16.7%)	1 (9.1%)
Internet, Correo electrónico, Página Web, Software de desarrollo, Software académico, Aulas Virtuales, Repositorio de Documentos, Software de configuración de dispositivos	0 (0%)	1 (16.7%)	1 (9.1%)
Internet, Correo electrónico, Página Web, Software de desarrollo, Software académico, Aulas Virtuales, Repositorio de Documentos, Software de Gestión de Bibliotecas, Software de planificación	0 (0%)	2 (33.3%)	2 (18.2%)
Internet, Correo electrónico, Página Web, Software de desarrollo, Software académico, Aulas Virtuales, Repositorio de Documentos, Software de planificación, Software de configuración de dispositivos	1 (20%)	0 (0%)	1 (9.1%)
Internet, Correo electrónico, Página Web, Software de desarrollo, Software académico, Repositorio de Documentos, Software de Gestión de Bibliotecas, Software de configuración de dispositivos	1 (20%)	0 (0%)	1 (9.1%)
Internet, Correo electrónico, Página Web, Software de desarrollo, Software académico, Repositorio de Documentos, Software de planificación, Software de configuración de dispositivos	0 (0%)	1 (16.7%)	1 (9.1%)
Internet, Correo electrónico, Software académico, Repositorio de Documentos, Software de planificación, Software de configuración de dispositivos	0 (0%)	1 (16.7%)	1 (9.1%)
Internet, Correo electrónico, Software de desarrollo, Aulas Virtuales	1 (20%)	0 (0%)	1 (9.1%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: La distribución de los servicios que se utiliza en la oficina son variados. Se observa una diferencia interesante en los que utilizan el Plan de Seguridad (ISO 27002), ya que el 33.3% utiliza software de gestión de bibliotecas y de planificación, y ninguno de los que no aplicaban el plan utilizaron estos servicios.

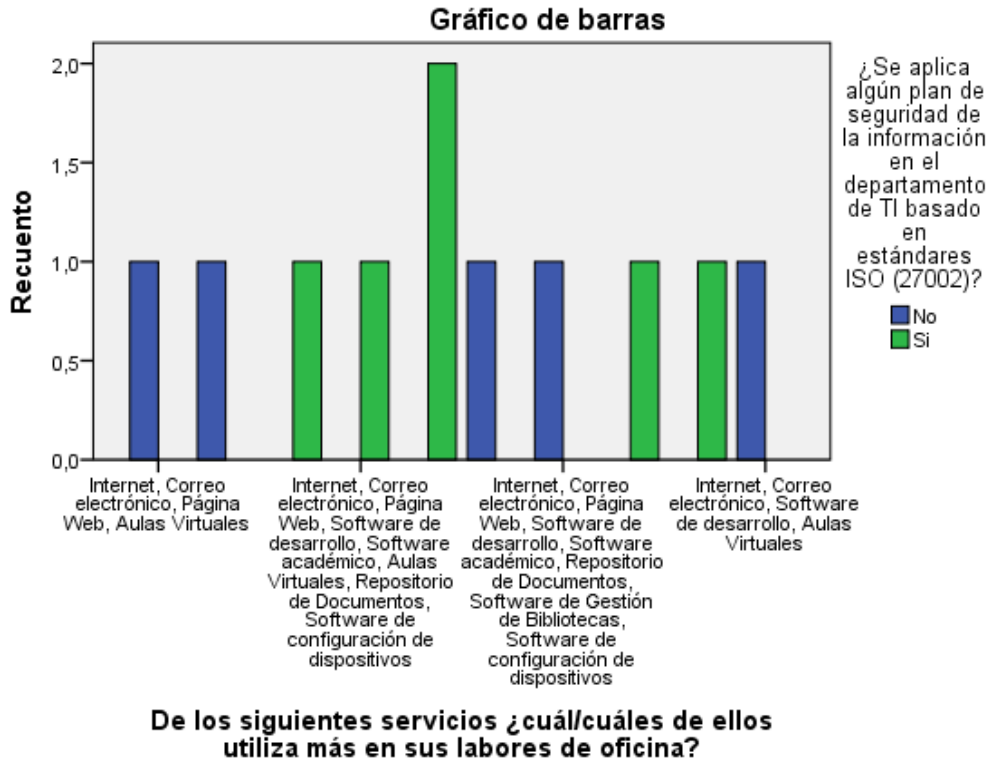

Gráfico 59. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Servicios utilizados en labores de oficina*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Suspensión de servicios

Tabla 89.

Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 POR Suspensión de servicios

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
No	3 (60%)	4 (66.7%)	7 (63.6%)
Si	2 (40%)	2 (33.3%)	4 (36.4%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: Se han presentado proporciones similares, sin diferencias estadísticamente significativas (χ^2: 0.052; p=0.819) en relación con la suspensión de algún servicio que se necesita para el trabajo diario de oficina entre los que aplican y los que no aplican el Plan de Seguridad de la Información basado en el ISO 27002.

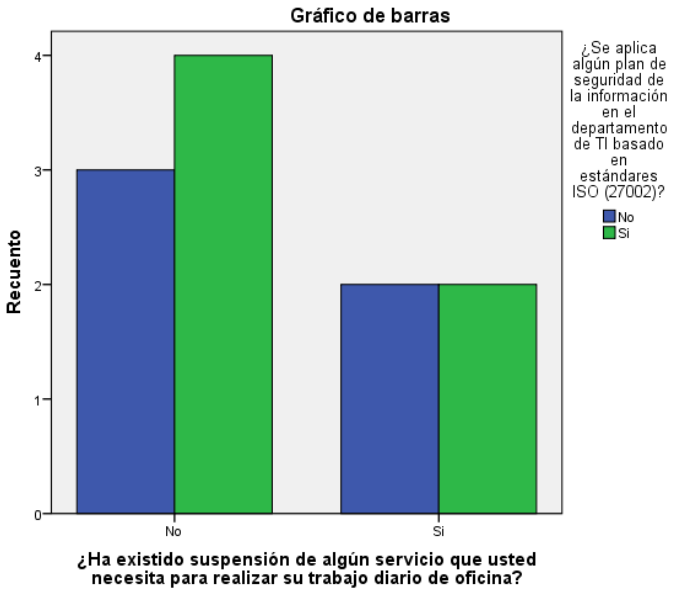

Gráfico 60. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Suspensión de servicios*

¿Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Frecuencia de pérdida del servicio?

Tabla 90.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 Por Frecuencia de pérdida del servicio?

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Frecuentemente	0 (0%)	1 (16.7%)	1 (9.1%)
Poco	3 (60%)	0 (0%)	3 (27.3%)
Muy poco	1 (20%)	3 (50%)	4 (36.4%)
Nada	1 (20%)	2 (33.3%)	3 (27.3%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: No existen diferencias estadísticamente significativas en las frecuencias de los que han perdido los servicios que utilizan durante su trabajo entre los grupos que aplican el Plan de Seguridad y los que no lo aplican (χ^2: 5.286; p=0.152).

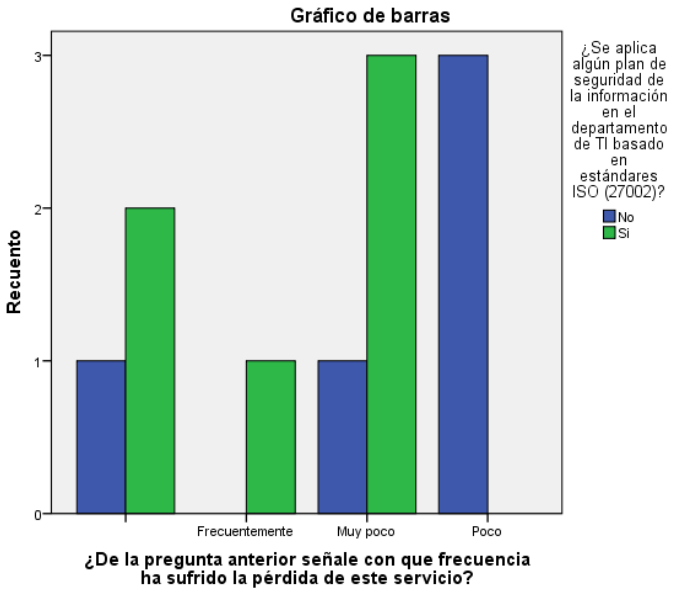

Gráfico 61. *¿Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Frecuencia de pérdida del servicio?*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Controles de detección de códigos maliciosos en dispositivos de defensa

Tabla 91.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002
Controles de detección de códigos maliciosos en dispositivos de defensa

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Siempre	1 (20%)	2 (33.3%)	3 (27.3%)
Casi siempre	2 (40%)	3 (50%)	5 (45.5%)
Pocas veces	1 (20%)	1 (16.7%)	2 (18.2%)
Nunca	1 (20%)	0 (0%)	1 (9.1%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: Se ha encontrado que 83.3% de los que aplican el Plan de Seguridad (ISO 27002) siempre/casi siempre tiene implementados controles de detección de códigos malicioso en sus dispositivos de defensa, cercano a ellos, con una frecuencia del 60% se encuentran los que no aplican el Plan, pero tienen la misma implementación (χ^2: 1.454; p=0.693). Existe un caso en los que nunca se ha implementado los controles de detección, perteneciente a los que no aplican el Plan.

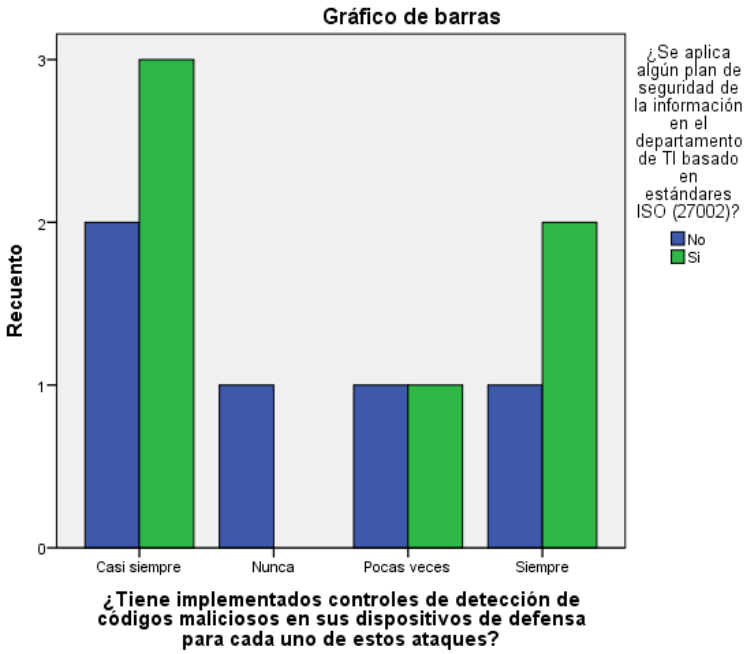

Gráfico 62. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 Controles de detección de códigos maliciosos en dispositivos de defensa*

Prueba de Aplicación de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por controles de prevención de códigos maliciosos en dispositivos de defensa

Tabla 92.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 Por Controles de prevención de códigos maliciosos en dispositivos de defensa

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Siempre	1 (20%)	1 (16.7%)	2 (18.2%)
Casi siempre	1 (20%)	2 (33.3%)	3 (27.3%)
Pocas veces	2 (40%)	3 (50%)	5 (45.5%)
Nunca	1 (20%)	0 (0%)	1 (9.1%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: La implementación de controles de prevención de códigos malicioso en los dispositivos de defensa se ha dado de manera similar entre los que aplican y no aplican plan de seguridad (ISO 27002) (χ^2: 1.451; p=0.693). Ligeramente llama la atención que exista un caso (20%) dentro de los que no aplican el Plan que nunca haya implementado los controles de prevención.

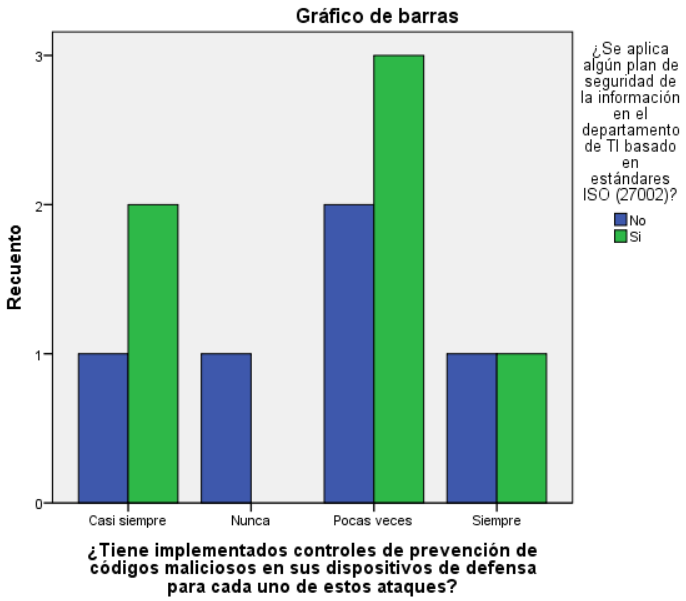

Gráfico 63. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Controles de prevención de códigos maliciosos en dispositivos de defensa*

Prueba de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Implementación de controles de protección de códigos maliciosos en dispositivos de defensa

Tabla 93.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Implementación de controles de protección de códigos maliciosos en dispositivos de defensa

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Siempre	1 (20%)	1 (16.7%)	2 (18.2%)
Casi siempre	1 (20%)	2 (33.3%)	3 (27.3%)
Pocas veces	1 (20%)	3 (50%)	4 (36.4%)
Nunca	2 (40%)	0 (0%)	2 (18.2%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: 40% de los que no aplican controles de seguridad (ISO 27002) nunca ha implementado controles de protección de códigos maliciosos, difiere de los que si han aplicado el Plan de seguridad en donde 16.7% siempre, 33.3% casi siempre y 50% pocas veces tienen implementados los controles de protección de códigos maliciosos (χ^2: 3.269; p=0.352).

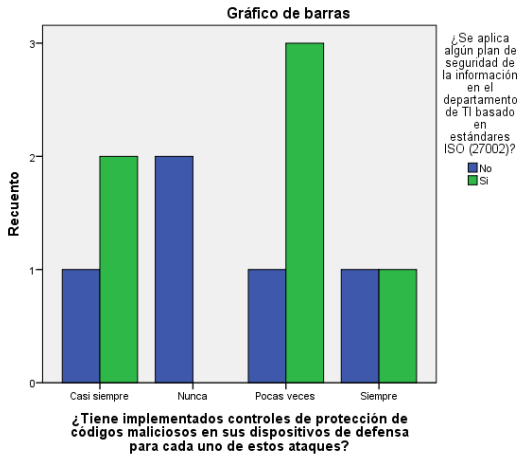

Gráfico 64. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Implementación de controles de protección de códigos maliciosos en dispositivos de defensa*

Prueba de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Análisis detallado de tipos de tráfico que entran y salen de su perímetro universitario

Tabla 94.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 Por análisis detallado de tipos de tráfico que entran y salen de su perímetro universitario

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Siempre	0 (0%)	1 (16.7%)	1 (9.1%)
Casi siempre	0 (0%)	1 (16.7%)	1 (9.1%)
Pocas veces	3 (60%)	3 (50%)	6 (54.5%)
Nunca	2 (40%)	1 (16.6%)	3 (27.3%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: En relación con el análisis detallado de los tipos de tráficos que entran y salen del perímetro universitario existen diferencias entre los que ha reportado que realizan casi siempre y siempre este análisis, puesto que 33.4% de los que aplican el Plan de Seguridad (ISO 27002) están en ese grupo y no está ninguno de los que no aplican el Plan de seguridad. 40% de los que no aplica el Plan nunca ha hecho un análisis detallado a diferencia de los que aplican el Plan en donde 16.6% nunca lo realizó (χ^2: 2.261; p=0.520).

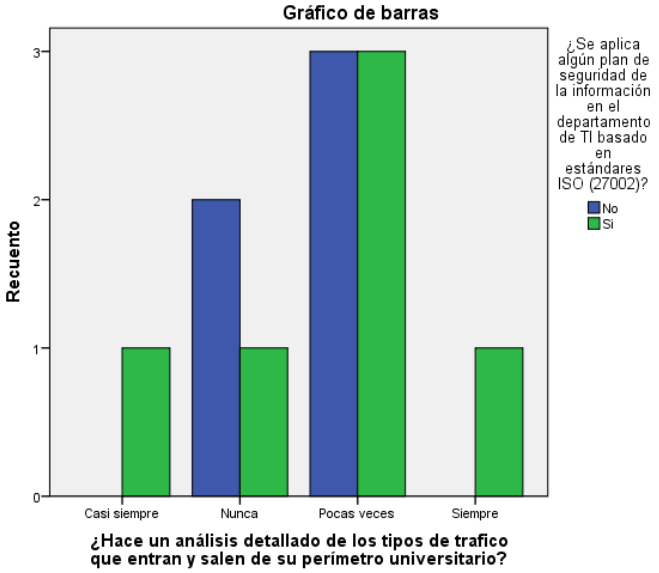

Gráfico 65. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por análisis detallado de tipos de tráfico que entran y salen de su perímetro universitario*

Prueba de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por realización de auditorías y actividades de verificación de los sistemas para minimizar la interrupción de los procesos

Tabla 95.

Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Realización de auditorías y actividades de verificación de los sistemas para minimizar la interrupción de los procesos

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Siempre	0 (0%)	1 (16.7%)	1 (9.1%)
Casi siempre	0 (0%)	3 (50%)	3 (27.3%)
Pocas veces	4 (80%)	1 (16.7%)	5 (45.5%)
Nunca	1 (20%)	1 (16.7%)	2 (18.2%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: Se observan que las auditorias y actividades de verificación son siempre y casi siempre en el 66.7% de los que aplican el Plan de Seguridad (ISO 27002), el 80% de los que no aplican el Plan pocas veces han realizado la auditoria, y 20% del grupo en mención nunca han realizado auditorias y actividades de verificación; sin embargo, estas no llegan a ser estadísticamente significativas (χ^2: 5.757; p=0.124)

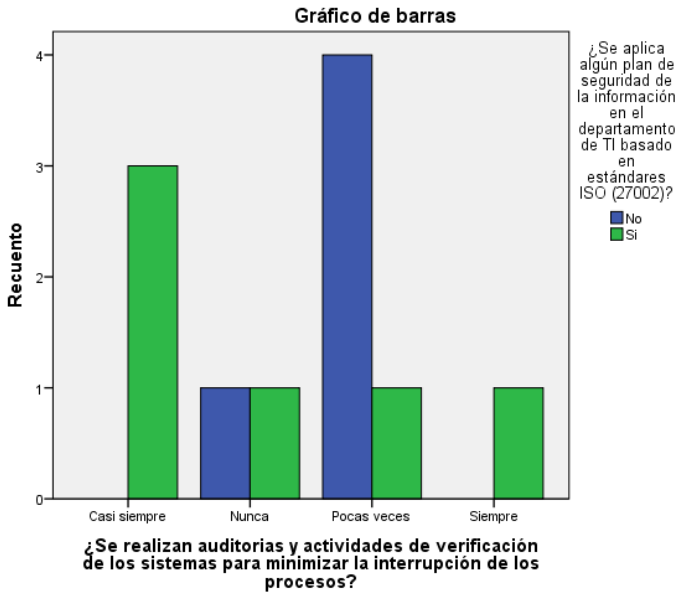

Gráfico 66. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Realización de auditorías y actividades de verificación de los sistemas para minimizar la interrupción de los procesos*

Prueba de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por realización de pruebas de penetración externas a sistemas de defensa perimetral

Tabla 96.

Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Realización de pruebas de penetración externas a sistemas de defensa perimetral

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Siempre	0 (0%)	1 (16.7%)	1 (9.1%)
Casi siempre	1 (20%)	1 (16.7%)	2 (18.2%)
Pocas veces	2 (40%)	2 (33.3%)	4 (36.4%)
Nunca	2 (40%)	2 (33.3%)	4 (36.4%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: La sumisión a pruebas de penetración externas a sus sistemas tiene ligeras variaciones entre los que aplican y los que no aplican el plan de seguridad (ISO 27002). En los primeros se ha encontrado un caso (16.7%) en los que siempre somete a pruebas, ese caso vendría a ser la única diferencia entre ambos grupos, se observa que no existen diferencias estadísticamente significativas (χ^2: 0.917; p=0.821).

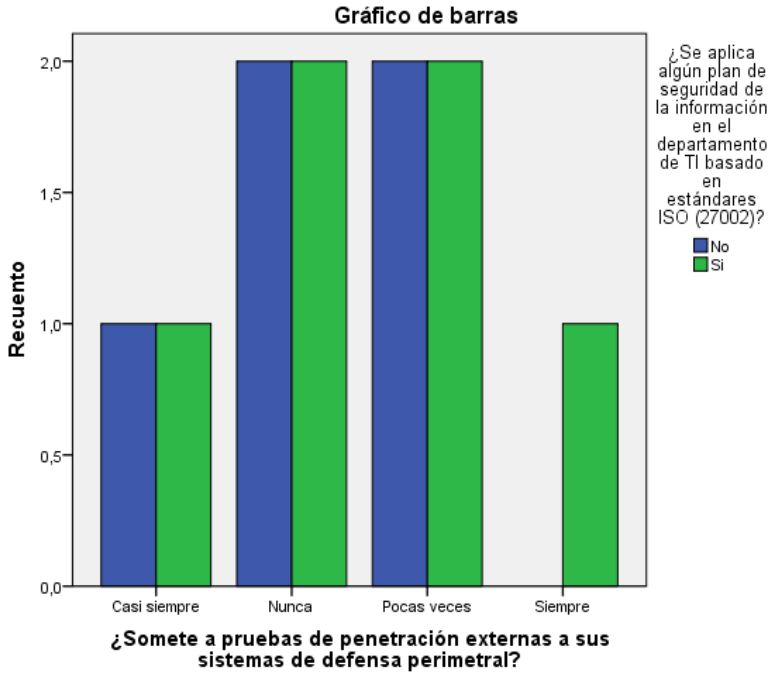

Gráfico 67. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Realización de pruebas de penetración externas a sistemas de defensa perimetral*

Prueba de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Aplicación de monitoreo de acceso a la red universitaria

Tabla 97.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 Por Aplicación de monitoreo de acceso a la red universitaria

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Siempre	1 (20%)	1 (16.7%)	2 (18.2%)
Casi siempre	1 (20%)	1 (16.7%)	2 (18.2%)
Pocas veces	3 (60%)	4 (66.7%)	7 (63.6%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: En cuanto a la realización del monitoreo de acceso a la red universitaria, la frecuencia es similar en los grupos que no aplican y los que aplican el plan de seguridad de la información (ISO 27002) (χ^2: 0.52; p=0.974).

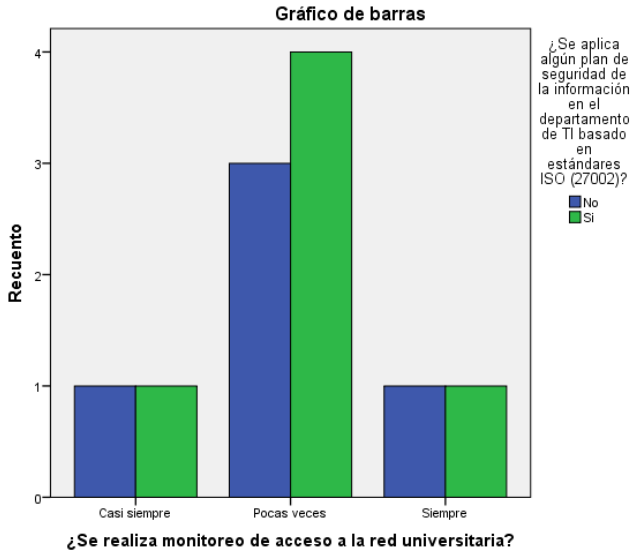

Gráfico 68. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Aplicación de monitoreo de acceso a la red universitaria.*

¿Prueba de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Realización de monitoreo al firewall principal de la universidad?

Tabla 98.

Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Realización de monitoreo al firewall principal de la universidad

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Siempre	1 (20%)	1 (16.7%)	2 (18.2%)
Casi siempre	1 (20%)	3 (50%)	4 (36.4%)
Pocas veces	2 (40%)	2 (33.3%)	4 (36.4%)
Nunca	1 (20%)	0 (0%)	1 (9.1%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: En cuanto al monitoreo del firewall principal de la universidad se ha encontrado un caso que no aplica el Plan de Seguridad (ISO 270002) que nunca ha monitoreado, 40% de los que no aplican el Plan siempre y casi siempre realizan el monitoreo, comparativamente 66.7% de los que si aplican el Plan han realizado el monitoreo (χ^2: 1.925; p=0.588).

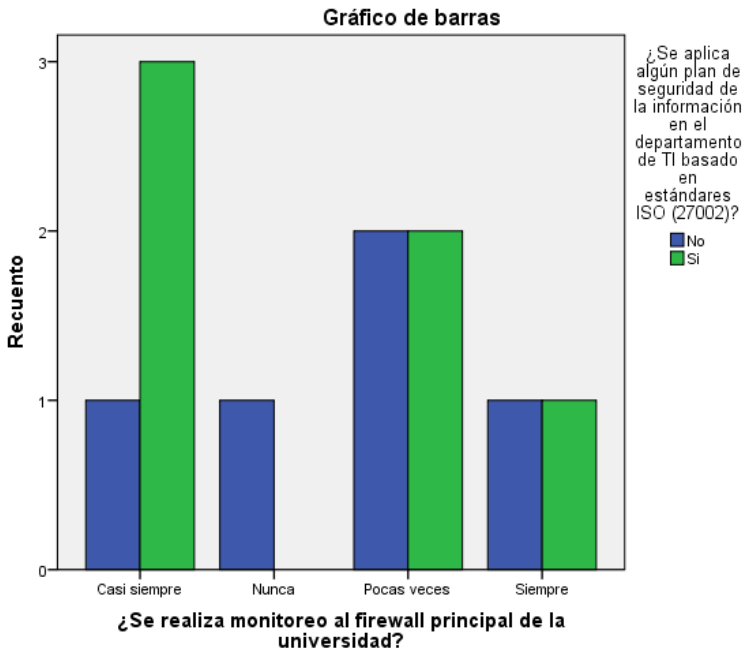

Gráfico 69. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Realización de monitoreo al firewall principal de la universidad*

Prueba de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Realización de monitoreo de cuentas privilegiadas a los sistemas y servidores.

Tabla 99.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Realización de monitoreo de cuentas privilegiadas a los sistemas y servidores

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Siempre	2 (40%)	2 (33.3%)	4 (36.4%)
Casi siempre	2 (40%)	2 (33.3%)	4 (36.4%)
Pocas veces	1 (20%)	2 (33.3%)	3 (27.3%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: La distribución de las frecuencias del monitoreo de cuentas privilegiadas a los sistemas y servidores es similar y proporcional en los grupos que aplican y no aplican los planes de seguridad de la información basados en la ISO 27002 (χ^2: 0.244; p=0.885).

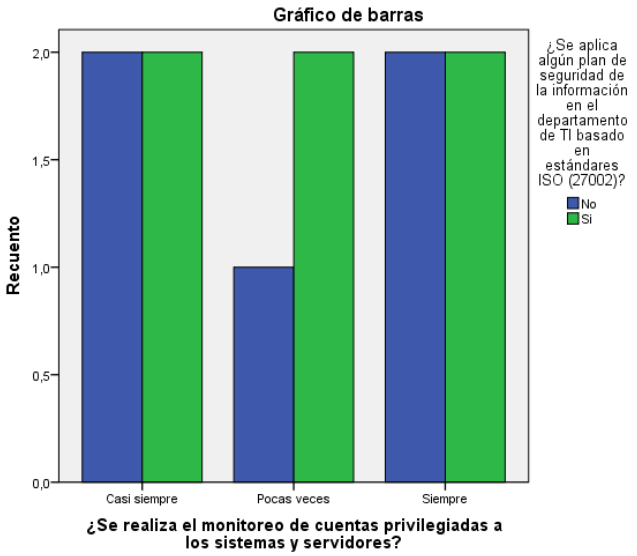

Gráfico 70. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Realización de monitoreo de cuentas privilegiadas a los sistemas y servidores*

Prueba de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Monitoreo al volumen de tráfico para identificar el uso indebido de los recursos de la universidad

Tabla 100.

Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Monitoreo al volumen de tráfico para identificar el uso indebido de los recursos de la universidad

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Siempre	1 (20%)	2 (33.3%)	3 (27.3%)
Casi siempre	1 (20%)	2 (33.3%)	3 (27.3%)
Pocas veces	2 (40%)	2 (33.3%)	4 (36.4%)
Nunca	1 (20%)	0 (0%)	1 (9.1%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: El monitoreo del volumen de tráfico (para identificación de uso indebido) tiene ligeras diferencias entre ambos grupos (los que aplican y los que no aplican el Plan de Seguridad basado en la ISO 27002). Se ha encontrado un caso (20%) en los que nunca ha aplicado el monitoreo correspondiente a los que no aplican el plan. 40% de los que no aplican el plan casi siempre o siempre monitorean, contrasta con el 66.6% de los que aplican el Plan (χ^2: 1.589; p=0.662). Aunque las proporciones se mantienen en ambos grupos.

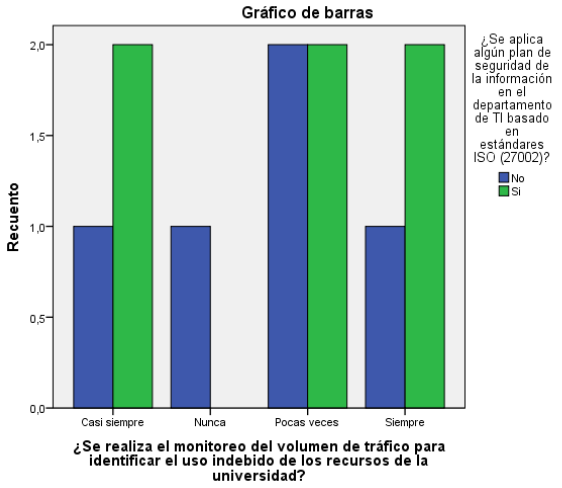

Gráfico 71. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Monitoreo al volumen de tráfico para identificar el uso indebido de los recursos de la universidad*

Prueba de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Supervisión de puertos comunes para los protocolos que permiten sesiones remotas

Tabla 101.

Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Supervisión de puertos comunes para los protocolos que permiten sesiones remotas

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Siempre	1 (20%)	1 (16.7%)	2 (18.2%)
Casi siempre	2 (40%)	3 (50%)	5 (45.5%)
Pocas veces	2 (40%)	2 (33.3%)	4 (36.4%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: La supervisión de los puertos comunes para los protocolos que permiten sesiones remotas se ha dado siempre y casi siempre en 60% en los que no aplican el Plan de Seguridad, y en 56.7% en los que aplican el Plan, guardando similitudes entre ambos grupos (χ^2: 0.11; p=0.946).

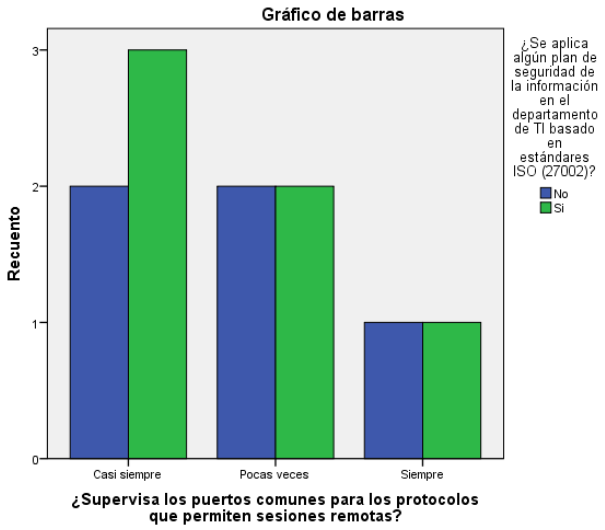

Gráfico 72. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Supervisión de puertos comunes para los protocolos que permiten sesiones remotas*

Prueba de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Recepción de ataque informático a sus redes de datos en general en el último mes

Tabla 102.
Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Recepción de ataque informático a sus redes de datos en general en el último mes

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
No	2 (40%)	5 (83.3%)	7 (63.6%)
Si	3 (60%)	1 (16.7%)	4 (36.4%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: En relación con la recepción de algún ataque informático durante el último el último mes, si existen diferencias cuantitativas en los que aplican y no aplican el Plan de Seguridad basado en la ISO 27002. 60% de los que no aplican han recibo ataques a sus redes. 16.7% de los que aplica el Plan han recibido; aunque dichos valores tienen el suficiente poder de poder generar diferencias estadísticamente significativas (χ^2: 0.737; p=0.391).

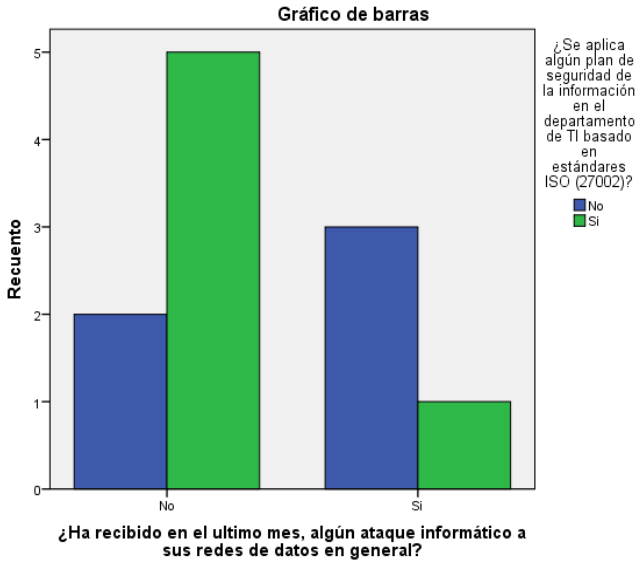

Gráfico 73. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Recepción de ataque informático a sus redes de datos en general en el último mes*

Plan de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Ataques más frecuentes en las redes WLAN de la universidad

Tabla 103.

Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Ataques más frecuentes en las redes WLAN de la universidad

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
DDoS, Scripting, SQL Injection	0 (0%)	1 (16.7%)	1 (9.1%)
Malware	0 (0%)	1 (16.7%)	1 (9.1%)
Malware, DDoS, SQL Injection, Spam	1 (20%)	0 (0%)	1 (9.1%)
Malware, Scripting, SQL Injection, Spam	1 (20%)	0 (0%)	1 (9.1%)
Malware, Spam	1 (20%)	1 (16.7%)	2 (18.2%)
Malware, Spam, Otros	0 (0%)	1 (16.7%)	1 (9.1%)
Phising, Spam	1 (20%)	0 (0%)	1 (9.1%)
Scripting	0 (0%)	1 (16.7%)	1 (9.1%)
Spam	1 (20%)	1 (16.7%)	2 (18.2%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: En lo que respecta la indagación de ataques informáticos a las redes WLAN la descripción tiene una amplia variación entre los que aplica y los que no aplican el Plan de Seguridad basado en la ISO 27002 (χ^2: 6.967; p=0.54). Sin embargo, dentro de todos ellos se tiene mayor coincidencia la selección de spam, por lo que podría considerarse uno de los potenciales problemas que presenta la Universidad.

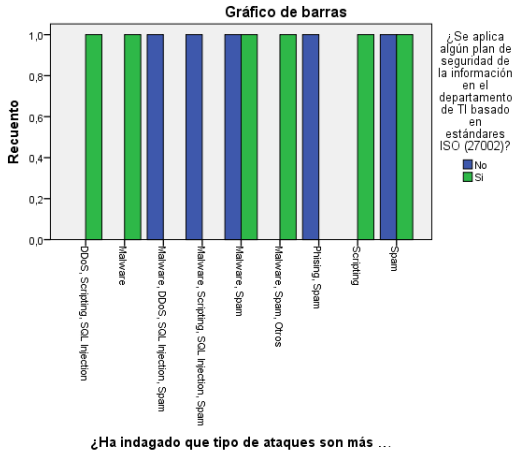

Gráfico 74. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Ataques más frecuentes en las redes WLAN de la universidad*

Prueba de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Cantidad de ataques detectados

Tabla 104.

Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Cantidad de ataques detectados

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Entre 1 y 30	5 (100%)	5 (83.3%)	10 (90.9%)
Entre 31 y 60	0 (0%)	1 (16.7%)	1 (9.1%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: Mayoritariamente se han detectado entre 1 a 30 ataques, no existe diferencia entre los que aplican el Plan de Seguridad basado en la ISO 27002 (χ^2: 0.01; p=0.92), sin embargo, llama la atención que uno (16.7%) de los que aplican el Plan ha detectado entre 31 y 60 ataques.

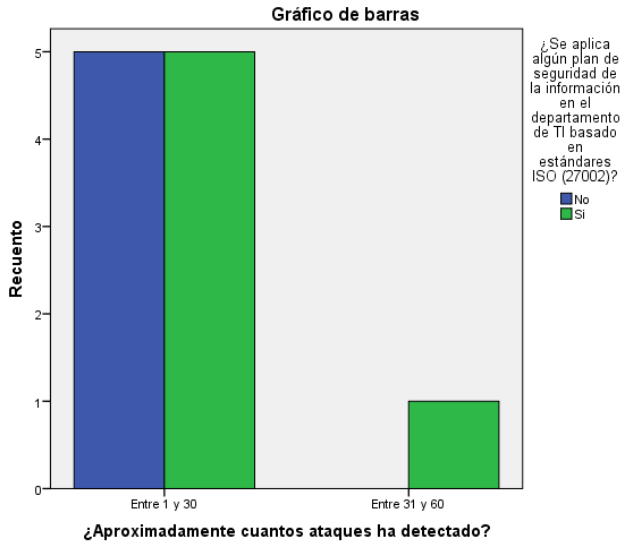

Gráfico 75. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Cantidad de ataques detectados*

Plan de Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Clasificación de ataques potenciales detectados

Tabla 105.

Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Clasificación de ataques potenciales detectados

	¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?		Total
	No	Si	
Insignificante	3 (60%)	4 (66.7%)	7 (63.6%)
Preocupante	2 (40%)	2 (33.3%)	4 (36.4%)
Total	5 (100%)	6 (100%)	11 (100%)

Interpretación: En lo que respecta a la clasificación a los ataques potencialmente detectados en ambos casos (los que aplican el Plan de Seguridad basado en ISO 27002) lo consideran mayoritariamente insignificante (60% los que no aplican el Plan y 66.7% los que si aplican) (χ^2: 0.16; p=0.6892).

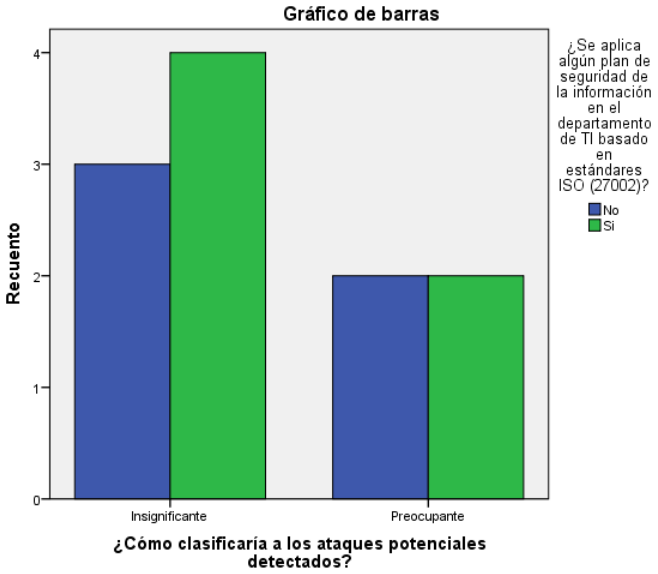

Gráfico 76. *Plan de seguridad informática en el departamento de TI basado en estándares ISO 27002 por Clasificación de ataques potenciales detectados*

V. DISCUSIÓN DE RESULTADOS

- El presente trabajo de investigación tuvo como objetivo realizar el análisis exploratorio de ataques informativos utilizando modelos de minería de datos para la predicción de los ataques futuros y frecuentes que se dan en el campus universitario a partir de los datos que resultaron de la captura de sniffers y firewall; para ello se utilizó la metodología CRISP-DM con una secuencia de pasos ordenados, se pudo determinar el algoritmo más eficiente para la predicción de ataques.
- Klenzi y Lopez (2017), en su propuesta "Detección de Ataques DoS con Herramientas de Minería de Datos", trabajó en modelar y mitigar ataques a un servidor de red por denegación de Servicios (Denial of Services) mediante el análisis offline de un flujo de datos simulados y la utilización de algoritmos y herramientas correspondientes a Minería de datos en flujos de datos continuos. En nuestro estudio se realizó también el análisis off-line pero de todo tipo de ataques, incluyendo Denegación de Servicios, se utilizó herramientas y algoritmos de asociación, segmentación y clasificación que nos permitió predecir futuros ataques.
- En el trabajo de Galán (2015) "Aplicación de la metodología CRISP-DM a un proyecto de minería de datos en el entorno universitario", indica que es importante hacer minería de datos partir de una base de datos o data warehouse (almacén de datos) que contenga la información que se quiere analizar y que ésta información esté correctamente estructurada, aplica estrictamente cada una de las distintas etapas de la metodología CRISP-DM sobre los datos académicos almacenados por la universidad en sus sistemas informáticos. De esta forma se pretende sacar conclusiones que ayuden a mejorar los servicios que ofrece la universidad a sus alumnos. En la presente investigación, se partió de una base de datos de ataques capturados durante más de un año, la que se encontraba correctamente estructurada en campos numéricos, alfanuméricos y de tipo fecha; igualmente se aplicó la metodología CRISP-DM en sus diferentes etapas para que nos ayude a elegir los algoritmos y modelos adecuados al objetivo para clasificar la información y compararlos para hacer predicciones.
- Mejía (2015) propone desarrollar un plan de seguridad informática basándose en la serie de normas ISO 27000 y de esta forma contemplando el uso de software libre y licenciado, facilitó conocer las fortalezas y debilidades de la institución pública, pero que sirve como referencia a otras instituciones educativas privadas para proveer directrices y comenzar con el análisis de riesgo de sus activos y así

fijar una escala del riesgo y proceder a realizar sus planes de seguridad informática. Esta investigación utiliza 3 controles de gestión de seguridad de la información definidos en la ISO 27002(2013): Cifrado, Seguridad en la operativa y Seguridad en las telecomunicaciones, que permitió relacionar los controles aplicados en la Universidad contrastado con los sugeridos por la Normativa internacional.

- Vallejo y Tenelanda (2012), indican que es necesario detectar las anomalías que se presenten en los registros de acceso de las redes de datos, indican que la minería de datos basada en una metodología adecuada, puede ser muy útil en el proceso de exploración de datos y que mediante tecnologías analíticas y procesos estadísticos permite generar reglas a partir de datos históricos de capturas, para generar reglas y patrones que permiten predecir intrusiones. El presente estudio utiliza una metodología adecuada para procesar los datos con tecnologías de software licenciado y software libre, procesos estadísticos y analíticos, no solo para predecir intrusiones, sino que se realizó la asociación de datos y la segmentación de los mismos para determinar casos atípicos dentro de los ataques informáticos que muchas veces no son detectados por los analizadores, snifers o el mismo firewall.
- Según la prueba de categoría de amenaza frente a la hora, se observa que los ataques de spyware se producen de 9 a 10am, y provienen casi en su totalidad del medio externo (98.6%); asi mismo los ataques de la categoría code-execution proviene del interior de la Universidad en 99.8%. esto permite afirmar que, a esa hora, la mayoría de estudiantes accede a las redes inalámbricas de la universidad, ocupan los laboratorios para las sesiones prácticas e inicia la labor administrativa en la universidad.
- El análisis de la primer hipótesis específica, indica que las categorías de ataques que tienen más incidencias son spyware y code-execution, el primero tiene como objetivo la zona de destino WAN y el segundo tiene por objetivo atacar a la Zona DMZ1, en la cual se encuentran los servidores públicos que comúnmente se configuran en las zonas desmilitarizadas de cualquier red empresarial, esto supone que los accesos a los servidores Web, servidor DNS y servidor de correo son los blancos perfectos para este tipo de ataque. Esta afirmación se comprueba también con la prueba de categoría de amenaza con el puerto de destino, que indica los ataques a los puertos 80 y 8000 utilizados por los servidores web, uno de ellos sería el Servidor_Web_HTTPS quien se verificó también recibe la mayor cantidad de ataques desde IP internas a la Universidad por code-execution.

- Los algoritmos con el mejor nivel de confiabilidad para la predicción de ataques informáticos futuros fueron el algoritmo de REPTree con un 99.94%, seguido del CHAID con un 98% tomando como predictor a la categoría de amenaza.
- Con el modelo de clasificación de minería de datos, se encontraron 2 grupos homólogos, el primero de 2375 registros y el segundo de 6066 registros, que no se ajustan a patrones de datos normales, lo que permitió identificar valores atípicos en el primer grupo en los campos: destination address, país de destino y categoría. En el segundo grupo las anomalías fueron de 32 registros encontrados de los campos: destination address, país de destino, categoría, regla, flags, nombre de amenaza, acción y zona de origen, lo que nos permitió detectar los campos comunes: destination address, país de destino y categoría como los casos extraños, es decir que se encuentran algunos registros que podrían considerarse falsos positivos.
- A través de la encuesta aplicada a los administradores de seguridad del departamento de TI, se concluye que las universidades pocas veces realizan auditorias y actividades de verificación de los sistemas para minimizar la interrupción de los procesos, la pregunta 26 de la ficha de recolección de datos lo verifica con 45.5%, siendo inclusive muy superior si sumamos los aspectos de hacerlo “siempre” y “casi siempre” que alcanzan solo al 36.4%. esto es preocupante porque a través de estas acciones se pueden identificar si se realiza una eficaz gestión de los recursos informáticos, entre ellos de la seguridad y si se aplican normativas que ayudan a minimizar los riesgos informáticos.
- Después de realizar la contratación de las hipótesis y el analizar los resultados estadísticos y exploratorios con las herramientas tecnológicas utilizadas, podemos afirmar que el trabajo de investigación logró el objetivo de determinar cómo el análisis exploratorio de ataques informáticos provisto por las herramientas de captura de datos, llevado a las herramientas de minería de datos, está estrechamente relacionado con la gestión de seguridad de las redes inalámbricas en universidades locales, permitiendo determinar que controles de seguridad informática se aplican correctamente o cuales tienen que modificarse para alcanzar un nivel de servicio adecuado.

VI. CONCLUSIONES

- Se comprobó que mediante el análisis exploratorio de ataques informáticos utilizando herramientas de minería de datos, existe una fuerte relación con la gestión de seguridad en redes inalámbricas en las universidades de la ciudad de Arequipa, que fue verificado a través de las pruebas de las hipótesis específicas.
- Se realizó la caracterización de los ataques informáticos a través de la metodología CRISP-DM que permitió gestionar de mejor forma la seguridad de las redes inalámbricas
- Con la minería de datos se pudo clasificar de manera óptima los datos capturados a través de los algoritmos de clasificación, los cuales nos indicaron que las amenazas más frecuentes fueron code-execution y spyware, aceptándose para la predicción de futuros ataques. Con un nivel de confianza predictiva del modelo REPTree de 99.94%, se determinó que las amenazas más frecuentes en el campus universitario fueron de tipo spyware hacia las redes inalámbricas. El modelo CHAID utiliza estadísticos de chi-cuadrado para identificar las divisiones óptimas de la data, encontrándose que la categoría con más alta probabilidad de predictor es el nombre de la amenaza/contenido, seguido del país de origen. Para el modelo C5.0, la importancia del predictor “Aplicación” es Zona de origen, nombre de la amenaza/contenido que son una posible clasificación de la predicción. La categoría de amenaza más frecuente en base al modelo de asociación Apriori es “code-execution” si la aplicación utilizada es "Web-browsing" y la acción de proteccion es enviar "reset-both". Los modelos de segmentación permitieron observar los valores atípicos que dispersaron la construcción de los clústeres en campos donde el valor era único o muy bajo, que no se ajustan a patrones de datos normales lo que nos permitió detectar los campos comunes.
- Con las pruebas estadísticas de la tercera hipótesis específica, se determina que la implementación de controles que previenen el tipo de ataque “código malicioso”, este hallazgo ayudó a contrastar también las realizadas con las pruebas de minería de datos que indicaron que el ataque más común detectado es “code-execution”, un tipo de código malicioso y con las pruebas estadísticas de la primera hipótesis específica, también se comprueba lo mismo. Con lo que quedó demostrado que la aplicación de controles de seguridad dispuesto por una

norma internacional de seguridad de la información ISO 27002(2013) no se relaciona con la forma de gestionar la seguridad de las redes informáticas.

VII. RECOMENDACIONES

- Gracias a los resultados devueltos por el análisis con minería de datos, se recomienda poner en el plan de trabajo de las oficinas de Tecnologías de la Información de las universidades, el uso de la metodología CRISP-DM utilizado en este trabajo de investigación, para tener una guía de referencia de las acciones que se deben ejecutar ante los miles de ataques que cada día se reciben en los servicios informáticos.
- Se sugiere desarrollar un plan de análisis de datos off-line de todo el tráfico que ingresa, atraviesa y sale del campus universitario, basado en herramientas inteligentes para estar mejor preparados ante los múltiples ataques de a cada minuto se perpetran por parte de los ciberdelincuentes.
- Se recomienda realizar un análisis más exhaustivo en los resultados del modelo de clusterización de detección de anomalías ya que la sospecha de anomalía, puede que resulte en anomalías reales y perjudique el servicio de seguridad y sea un falso positivo de ataque informático.
- Se recomienda establecer un plan de capacitación y concientización a todo el personal que utiliza los servicios informáticos de la universidad en aspectos de seguridad informática, dado que los pocos controles que tienen, no se aplican. Este fue el diagnóstico que este estudio pudo obtener.

VIII. REFERENCIAS

AGUIRRE, Julio (2017). Metodología ISO 27000 para optimizar rendimiento de redes corporativas medianas móviles manteniendo estándares de disponibilidad y seguridad. Universidad Tecnológica Israel.

ANDERSON, James, P. 1972. Computer Security Technology Planning Study. ESD-TR-73-51, v II. Electronic Systems Division, Air Force Systems Command, Hanscom Filed, Bedford, MA.

AUDITOOL (2012). Los medidores del desempeño en la seguridad informática. Revista Contaduría Pública. Instituto Mexicano de Contadores Públicos. Consultado el 24.02.19. Disponible en: https://www.auditool.org/blog/auditoria-de-ti/1264-los-medidores-del-desempeño-en-la-seguridad-informatica

BALUJA, W. y Fernandez, Y. Detección de Intrusiones en Redes Inalámbricas. 2009.

BRISA, Sergio (2015). Squert, SecurityArtWork. Consultado el 29.07.19. Disponible en https://www.securityartwork.es/2015/06/22/squert/.

BRITOS, J. (2010). "Detección de Intrusiones en redes de datos con captura distribuida y procesamiento estadístico".

CANDIA, Dennis (2019). Predicción del rendimiento académico de los estudiantes de la UNSAAC a partir de sus datos de ingreso utilizando algoritmos de aprendizaje automático". UNSAAC. Cusco

CASTILLO Mora Anabel F. 2006. "Consultorías para la determinación de brechas de seguridad de una red inalámbrica". Escuela Superior Politecnico del Litoral: Guayaquil, Ecuador. p. 247.

CRIPS-DM(2016). "La metodología para poner orden en los proyectos de Data Science". Cross Industry Standard Process for Data Mining. Consultado el 13.08.19. Disponible en https://data.sngular.com/es/art/25/crisp-dm-la-metodologia-para-poner-orden-en-los-proyectos-de-data-science

CRUZ, F.L.C.d.l., 2004. "Fuerzas Armadas de la region andina en el contexto de la seguridad cibernetica, en concordancia con la resolucion de la O.E.A. AG/RES. 2004 (XXXIV-O/04)". Del Salvador, Buenos Aires, Argentina

CHAPARRO B. Fabián, Sáchica A. María, Vargas J. Andres, "Análisis de la gestión y seguridad de las redes WLAN en el rango de frecuencias cercano a la banda de 2.4 Ghz". 2001

ECHEVERRÍA José Antonio "Detección de intrusos en la capa de enlace del protocolo 802.11", La Habana : Instituto Superior Politécnico (CUJAE), 2012.

EBEL Frank, Guillaume FORTUNATO, Jéróme HENNECART, Laurent SCHALKWIJK, Marion AGÉ, Nicolas CROCFER, Robert CROCFER y Sébastien LASSON "Seguridad informática - Hacking Ético: Conocer el ataque para una mejor defensa" 4ª edición), Edic. ENI, Barcelona España, 2015.

GALAN Cortina Victor, "Aplicación de la metodología CRISP-DM a un proyecto de minería de datos en el entorno universitario", Universidad Carlos III de Madrid Escuela Politécnica Superior Ingeniería en Informática, Madrid España, 2015.

GARCÍA Jose Antonio, Acevedo Angela. Análisis Para Predicción De Ventas Utilizando Minería De Datos En Almacenes De Ventas De Grandes Superficies. Universidad Tecnológica de Pereira. Colombia, 2010.

GESTION (2015) Alerta: los 4 ataques informáticos más frecuentes y cómo evitarlos. Consultado el 15.03.19. Disponible en: https://destinonegocio.com/pe/gestion-pe/alerta-los-4-ataques-informaticos-mas-frecuentes-y-como-evitarlos/

IMÉNEZ M.I, Gómez J, Padilla N. Utilización de sistemas de detección de intrusos como elemento de seguridad perimetral". Universidad de Almería. España, 2008.

GUTIERREZ Juan Jesus, "Propuesta de una metodología de extracción de conocimientos a partir de datos de las prestaciones del seguro integral de salud en la región Piura en el año 2016", pag.60. Universidad Católica Los Angeles de Chimbote, Perú. 2017

GTDI (2017) Proyecto de Norma Técnica peruana (PNTP) ISO/IEC 27002:2017 en consulta pública. Consultado el 12.08.19. Disponible en https://www.gtdi.pe/PNTP-ISO-IEC_27002_2017_en_consulta_publica

HARALD Sundmaeker, Patrick Guillemin, Peter Friess, Sylvie Woelfflé, "Vision and Challenges for realising the Internet of Things", Marzo 2010.

IANA.ORG (2019). Service Name and Transport Protocol Port Number Registry. Consultado el 23.08.19. Disponible en https://www.iana.org/assignments/service-names-port-numbers/service-names-port-numbers.txt

INFORMÁTICA DIGITAL (2018). Los 7 Mejores Analizadores de red y Sniffers para windows y Linux. Consultado el 22.07.19. Disponible en https://www.locura informaticadigital.com/2018/03/02/7-mejores-analizadores-de-red-sniffers-windows-y-linux/

ISO 27002.ES (2012). Controles de seguridad. Consultado el 24.03.19. Disponible en: http://www.iso27000.es/iso27002.html

KASPERSKY (2016) Kaspersky Internet Security. Consultado el 24.03.19. Disponible en: https://support.kaspersky.com/sp/kis2016.

KEITA Soy, Fanta. “Detección de Intrusos en la Capa de Enlace del Protocolo 802.11”. Instituto Superior Politecnico Jose Antonio Echeverria. CUJAE. 2012.

KHALIL El-Khatib, M.G., Aboubakr Lbekkouri (2008). Selecting the Best Set of Features for Efficient Intrusion Detection in 802.11 Networks. Volume, 4.

KLENZI, Raúl; Lopez, Marcelo. “Detección de ataques con Herraminetas de Minería de Datos”. Universidad Nacional de San Juan. Argentina, 2017.

MEJIA Viteri José Teodoro, “Plan de seguridad informática del departamento de tencologias de la información y comunicación de la Universidad Tecnica de Babahoyo para mejorar la gestión en la confidencialidad e integridad de la información y disponibilidad de los servicios”, Universidad Autónoma de los Andes. 2015

PASCUAL, A.E., Seguridad en Redes Inalámbricas. 2007.

PALO ALTO Networks-OS (2019). “Guía del administrador de PAN-OS®” Consultado el 10.08.19. Disponible en https://docs.paloaltonetworks.com/pan-os/7-1/pan-os-admin/monitoring/syslog-field-descriptions.html#67983.

SGSI. Blog especializado en Sistemas de Gestión de Seguridad de la Información. Consultado el 24.03.15. Disponible en: https://www.pmg-ssi.com/2015/12/iso-27001-2013-seguridad-informacion-nube/

TÉLLEZ Castaño, Carlos Felipe. Detección de intrusos y seguridad en redes móviles ad-hoc. Seminario de Investigación. Colombia. 2010

TSB, U.-T.O.c.d.N.d.l.T., 2006. La seguridad de las telecomunicaciones y las tecnologías de la información, UIT, Editor. UIT: Ginebra

VALLEJO P. Diego, Tenelanda V Germán. «Minería de datos aplicada en detección de intrusos». Ingenierías USBMed 3, n.º 1 (30 de junio de 2012). https://doi.org/10.21500/20275846.264.

VARELA Carlos, Domínguez Luis. 2002, “Redes Inalámbricas”. Escuela Técnica Superior de Ingeniería Informática, Universidad de Valladolid. Volumen 18. Pag.18.

24 Estadísticas de Seguridad Informática que Importan en el 2019. Consultado el 05.06.19. Disponible en https://preyproject.com/blog/es/24-estadisticas-seguridad-informatica-2019/

IX. ANEXOS

ANEXO 1

MATRIZ DE CONSISTENCIA

MATRIZ DE CONSISTENCIA

TEMA: ANÁLISIS EXPLORATORIO DE ATAQUES INFORMÁTICOS APLICANDO HERRAMIENTAS DE MINERIA DE DATOS, PARA LA GESTIÓN DE LA SEGURIDAD DE REDES INALÁMBRICAS EN UNIVERSIDADES DE AREQUIPA

PROBLEMA	OBJETIVOS	HIPOTESIS	VARIABLES	METODOLOGIA
Problema general: ¿Cómo el análisis exploratorio de ataques informáticos aplicando herramientas de minería de datos, está relacionado con la gestión de seguridad de las redes inalámbricas en universidades de Arequipa? **Problemas específicos:** a) ¿Cómo realizar la caracterización de ataques informáticos para la gestión de seguridad de las redes inalámbricas? b) ¿Cómo las características de los ataques informáticos determinan los patrones de predicción de acuerdo a la categoría de la amenaza utilizando modelos de clasificación, asociación y segmentación? c) ¿Cómo es la relación de la gestión de seguridad y los controles aplicados según estándares ISO 27002?	**Objetivo general:** Determinar cómo el análisis exploratorio de ataques informáticos aplicando herramientas de minería de datos, está relacionado con la gestión de seguridad de las redes inalámbricas en universidades de Arequipa. **Objetivos específicos:** a) Realizar la caracterización de ataques informáticos para gestionar la seguridad de las redes inalámbricas b) Caracterizar los ataques informáticos para determinar los patrones de predicción de acuerdo a la categoría de la amenaza utilizando modelos de clasificación, asociación y segmentación. c) Evaluar la relación de la gestión de seguridad con los controles de seguridad aplicados según estándares ISO 27002.	**Hipótesis general:** El análisis exploratorio de ataques informáticos aplicando herramientas de minería de datos, está relacionado con la gestión de seguridad de las redes inalámbricas en universidades de Arequipa **Hipótesis específicas:** a) La caracterización de ataques informáticos está asociado a la gestión de seguridad de las redes inalámbricas. b) Las características de los ataques informáticos permiten determinar los patrones de predicción de acuerdo a la categoría de la amenaza utilizando modelos de clasificación, asociación y segmentación. c) La gestión de seguridad está relacionada con los controles de seguridad aplicados según estándares ISO 27002.	**Variable 1:** **ANÁLISIS EXPLORATORIO DE ATAQUES INFORMÁTICOS APLICANDO HERRAMIENTAS DE MINERIA DE DATOS** **Indicadores**: - Número de ataques por categoría - Temporalidad del ataque - Número de ataques por dispositivo - Frecuencia de ataques por país de procedencia - Frecuencia de ataques por país de destino - Número de incidencias por mes - Incidencia de zona de ataque - Severidad de la amenaza - Numero de repeticiones de acceso - Ataques por zona de origen - Ataques por IP origen - Ataques por puerto origen - Aplicaciones atacadas - Tipo de acción de defensa ejecutada **Variable 2:** **GESTIÓN DE SEGURIDAD DE LAS REDES INALAMBRICAS**	**Tipo de Investigación:** Aplicada **Nivel de Investigación** Descriptivo, correlacional y predictivo **Diseño:** No experimental **Técnicas de recolección de datos:** • Minería de datos • Encuesta **Instrumentos** • Ficha de registro de reporte obtenido del firewall • Ficha de recolección de datos (gestión de seguridad) **Población:** • Tráfico de ataques. • Personal del departamento de Seguridad de TI **Muestra:** - No se realizó un muestreo porque se trabajó con la totalidad de la población: ○ 8441 registros de ataques ○ 11 especialistas

			Indicadores: - Uso de protocolos de cifrado autorizados. - Uso de contraseñas para acceso remoto. - Identificación de acceso de usuarios de la red. - Política de cambio de contraseñas. - Desactivación de credenciales de trabajadores retirados. - Existencia de procedimientos frente a problemas informáticos. - Periodo de tiempo y cifrado de respaldo de la información. - Aplicación de parches de seguridad. - Registro, análisis y seguimiento de logs. - Plan de instalaciones de software. - Uso de servicios informáticos - Control de detección y prevención de códigos maliciosos. - Análisis de tráfico, auditoría y monitoreo de sistemas. - Plan de seguridad informática utilizando ISO 27002.	

ANEXO 2

FICHA DE RECOLECCION DE DATOS

Dirigido al: Administradores y Oficiales de seguridad de TI

La siguiente ficha de recolección de datos, tiene por objetivo recopilar información sobre la aplicación de políticas, normas y procedimientos de seguridad informática utilizados en el campus universitario, organizado en base a 3 controles de Seguridad de la Norma ISO 27002(2013)

Control 10: CIFRADO

1. Utiliza protocolos de cifrado aprobados por la universidad en las comunicaciones inalámbricas?
   Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐
2. En caso sea afirmativo, podría indicar cuales?
   WPA-PSK (TKIP)
   WPA-PSK (AES)
   WPA2-PSK (TKIP)
   WPA2-PSK (AES)
   WEP
   WPA3
   WPA2
   Ninguno
3. Usa contraseñas o autenticación para usuarios con acceso remoto?
   Si ☐ No ☐
4. ¿Identifica niveles de acceso de todos los usuarios de la red universitaria para ajustarlos según sea necesario y bloquea intentos de Superusuario o Administrador??
   Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐
5. ¿Existe alguna política por parte de los encargados de TI para el cambio de su contraseña en su computador cada cierto periodo de tiempo?
   Si ☐ No ☐
6. Si existe alguna política o exigencia con qué frecuencia lo realiza.
   Cada mes ☐
   Cada Tres meses ☐
   Cada seis meses ☐
   Cada año o más ☐

CONTROL 12: SEGURIDAD EN LA OPERATIVA

7. ¿Desactiva las credenciales de los trabajadores que ya no laboran en la universidad en coordinación con RRHH apenas termina su contrato?
8. Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐

9. ¿Existe en el departamento de TI, procedimientos a seguir en caso de ocurrir algún problema con los servidores, computador o servicios informáticos que necesita para labores cotidianas
   Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐
10. Existe en el departamento de TI, alguna política que indique en qué periodo tiempo se debe respaldar la información de su equipo?
    Si ☐ No ☐
11. En caso de realizar el respaldo de información indique cada que tiempo lo realiza
    Cada mes ☐
    Cada Tres meses ☐
    Cada seis meses ☐
    Cada año o más ☐
12. Los backups y archivos con data sensible están cifrados?
    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐
13. Los parches de seguridad solucionan agujeros de seguridad sin modificar la funcionalidad del programa, son frecuentes en aplicaciones que interactúan con Internet. ¿Se aplican estos parches en sus servidores y estaciones de trabajo?
    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐
14. Qué tipo de parches configura?
    Parches a archivos binarios ☐
    Parches al código fuente ☐
    Parches de depuración ☐
    Parches de seguridad ☐
    Parches de actualización ☐
    Parches de traducción ☐
    Parches de piratería ilegal ☐
    Otros ☐
15. Realiza un registro de los logs de eventos de seguridad, actividades de usuarios, excepciones, fallas?
    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐
16. Analiza los logs de seguridad de sus sistemas en línea o fuera de línea
    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐
17. Ha realizado actividades de seguimiento de los logs de seguridad
    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐
18. Contempla un plan de instalaciones de software en la universidad?
    Si ☐ No ☐
19. Se sigue procedimientos o normas establecidos por el departamento de TI en las instalaciones de software por parte de los usuarios?
    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐
20. De los siguientes servicios ¿cuál de ellos utiliza más en sus labores en la oficina?
    Internet ☐
    Correo electrónico ☐
    ☐

Página Web ☐
Software de desarrollo ☐
Software académico ☐
Aulas Virtuales ☐
Repositorio de Documentos. ☐
Software de Gestión de Bibliotecas ☐
Software de planificación ☐
Software de configuración de dispositivos ☐
Software de Ayuda. ☐

21. ¿Ha existido suspensión de algún servicio que usted necesita para realizar su trabajo diario en la oficina?
    Si ☐ No ☐
22. ¿De la pregunta anterior señale con una X con qué frecuencia ha sufrido la pérdida de este servicio?
    Muy Poco ☐
    Poco ☐
    Frecuentemente ☐
    Muy frecuentemente ☐
    Siempre ☐
23. ¿Tiene implementados controles de detección de códigos maliciosos en sus dispositivos de defensa para cada uno de estos ataques?
    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐
24. ¿Tiene implementados controles de prevención de códigos maliciosos en sus dispositivos de defensa para cada uno de estos ataques?
    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐
25. Tiene implementados controles de protección de códigos maliciosos en sus dispositivos de defensa para cada uno de estos ataques?
    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐
26. Hace un análisis detallado de los tipos de trafico que entran y salen de su perímetro empresarial?
    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐
27. Se realizan auditorias y actividades de verificación de los sistemas para minimizar la interrupción de los procesos?
    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐
28. ¿Somete a pruebas de penetración externas a sus sistemas de defensa perimetral?
    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐

Control 13: Gestión de seguridad en las redes

29. Se realiza monitoreo de acceso a la red universitaria
    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐

30. Se realiza monitoreo al firewall principal de la universidad?

    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐

31. Se realiza algún monitoreo de cuentas privilegiadas a los sistemas y servidores?

    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐

32. ¿Se realiza el monitoreo del volumen de tráfico para identificar el uso indebido de los recursos de la universidad?

    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐

33. ¿Supervisa los puertos comunes para los protocolos que permiten sesiones remotas??

    Siempre ☐ Casi Siempre ☐ Pocas veces ☐ Nunca ☐

34. Ha indagado que tipo de ataques son más frecuentes en las redes WLAN de su universidad?

    Malware ☐
    DDoS ☐
    Phising ☐
    Baiting ☐
    Scripting ☐
    SQL Injection ☐
    Spam ☐
    Otros ☐

35. Aproximadamente cuantos ataques ha detectado?

    Entre 1 y 30 ☐
    Entre 31 y 60 ☐
    Entre 61 y 90 ☐
    Mas de 100 ☐

36. Cómo clasificaría a los ataques potenciales detectados?

    Insignificante ☐ Preocupante ☐ Crítico ☐

37. ¿Se aplica algún plan de seguridad de la información en el departamento de TI basado en estándares ISO (27002)?

    Si ☐ No ☐

ANEXO Nº 03
FICHA DE VALORACION DEL INSTRUMENTO

TABLA DE VALIDACIÓN DE INSTRUMENTO POR EXPERTOS

Apellidos y Nombres del Informante	Institución donde labora	Nombre del Instrumento	Autor del Instrumento
Dr. Máximo Rondón Rondón	UCSM	CUESTIONARIO	Mg. Karina Rosas Paredes

Título de la Investigación:
ANÁLISIS EXPLORATORIO DE ATAQUES INFORMÁTICOS APLICANDO HERRAMIENTAS DE MINERÍA DE DATOS, PARA LA GESTIÓN DE LA SEGURIDAD DE LAS REDES INALÁMBRICAS EN UNIVERSIDADES DE AREQUIPA

		DEFICIENTE				REGULAR				BUENA				MUY BUENA				EXCELENTE			
		5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
1. CLARIDAD	Esta formulado con lenguaje apropiado.																	X			
2. OBJETIVIDAD	Esta expresado en conductas observables.																	X			
3. ACTUALIZACIÓN	Está adecuado al avance de la ciencia y la tecnología.																	X			
4. ORGANIZACIÓN	Esta organizado en forma lógica.																	X			
5. SUFICIENCIA	Comprende aspectos cuantitativos																	X			
6. INTENCIONALIDAD	Es adecuado para valorar el aprendizaje de estadística																	X			
7. CONSISTENCIA	Está basado en aspectos teóricos científicos.																	X			
8. COHERENCIA	Entre las variables, dimensiones, Indicadores e ítems.																	X			
9. METODOLOGÍA.	La estrategia responde al propósito de la investigación.																	X			
10. PERTINENCIA	La escala es aplicable.																	X			

6. ASPECTOS DE EVALUACIÓN

OPINIÓN DE APLICACIÓN

Bien para ser aplicado. Buen instrumento

PROMEDIO DE VALORACIÓN: 85

LUGAR Y FECHA	DNI	FIRMA DEL EXPERTO INFORMANTE	N° DE TELEFONO
AQP 01-08-2019	29202992		987 965 584

TABLA DE VALIDACIÓN DE INSTRUMENTO POR EXPERTOS

Apellidos y Nombres del Informante	Institución donde labora	Nombre del Instrumento	Autor del Instrumento
Dr. Álvaro Fernández Del Carpio	UCSM	Cuestionario	Mg. Karina Rosas Paredes

Título de la Investigación:

Análisis exploratorio de ataques informáticos aplicando herramientas de minería de datos, para la gestión de la seguridad de redes inalámbricas en universidades de Arequipa.

		DEFICIENTE				REGULAR				BUENA				MUY BUENA				EXCELENTE			
		5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
1. CLARIDAD	Esta formulado con lenguaje apropiado.																	X			
2. OBJETIVIDAD	Esta expresado en conductas observables.																	X			
3. ACTUALIZACIÓN	Está adecuado al avance de la ciencia y la tecnología.																	X			
4. ORGANIZACIÓN	Esta organizado en forma lógica.																	X			
5. SUFICIENCIA	Comprende aspectos cuantitativos																	X			
6. INTENCIONALIDAD	Es adecuado para valorar el aprendizaje de estadística																	X			
7. CONSISTENCIA	Está basado en aspectos teóricos científicos.																	X			
8. COHERENCIA	Entre las variables, dimensiones, indicadores e ítems.																	X			
9. METODOLOGÍA.	La estrategia responde al propósito de la investigación.																	X			
10. PERTINENCIA	La escala es aplicable.																	X			

1. ASPECTOS DE EVALUACIÓN

I. OPINIÓN DE APLICACIÓN

Se sugiere 2 objetivos

II. PROMEDIO DE VALORACIÓN: 85

LUGAR Y FECHA	DNI	FIRMA DEL EXPERTO INFORMANTE	N° DE TELEFONO
AREQUIPA 24/07/19	29470715		947968915

TABLA DE VALIDACIÓN DE INSTRUMENTO POR EXPERTOS

Apellidos y Nombres del Informante	Institución donde labora	Nombre del Instrumento	Autor del Instrumento
Mg. José Esquicha Tejada	UCSM	CUESTIONARIO	Mg. Karina Rosas Paredes

Título de la Investigación:

ANÁLISIS EXPLORATORIO DE ATAQUES INFORMÁTICOS APLICANDO HERRAMIENTAS DE MINERÍA DE DATOS, PARA LA GESTIÓN DE LA SEGURIDAD DE LAS REDES INALÁMBRICAS EN UNIVERSIDADES DE AREQUIPA

		DEFICIENTE				REGULAR				BUENA				MUY BUENA				EXCELENTE			
		5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
1. CLARIDAD	Esta formulado con lenguaje apropiado.																	X			
2. OBJETIVIDAD	Esta expresado en conductas observables.																	X			
3. ACTUALIZACIÓN	Está adecuado al avance de la ciencia y la tecnología.																	X			
4. ORGANIZACIÓN	Esta organizado en forma lógica.																	X			
5. SUFICIENCIA	Comprende aspectos cuantitativos																	X			
6. INTENCIONALIDAD	Es adecuado para valorar el aprendizaje de estadística																	X			
7. CONSISTENCIA	Está basado en aspectos teóricos científicos.																	X			
8. COHERENCIA	Entre las variables, dimensiones, Indicadores e ítems.																	X			
9. METODOLOGÍA.	La estrategia responde al propósito de la investigación.																	X			
10. PERTINENCIA	La escala es aplicable.																	X			

6. ASPECTOS DE EVALUACIÓN

OPINIÓN DE APLICACIÓN

Es aceptable el nivel, muy buenas preguntas

PROMEDIO DE VALORACIÓN: 85

LUGAR Y FECHA	DNI	FIRMA DEL EXPERTO INFORMANTE	N° DE TELEFONO
Arequipa 02/08/19	43635330		959143336

TABLA DE VALIDACIÓN DE INSTRUMENTO POR EXPERTOS

Apellidos y Nombres del Informante	Institución donde labora	Nombre del Instrumento	Autor del Instrumento
Dr. José Sulla Torres	UCSM	Cuestionario	Mg. Karina Rosas Paredes

Título de la Investigación:

Análisis exploratorio de ataques informáticos aplicando herramientas de minería de datos, para la gestión de la seguridad de redes inalámbricas en universidades de Arequipa.

		DEFICIENTE				REGULAR				BUENA				MUY BUENA				EXCELENTE			
		5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
1. CLARIDAD	Esta formulado con lenguaje apropiado.																X				
2. OBJETIVIDAD	Esta expresado en conductas observables.																X				
3. ACTUALIZACIÓN	Está adecuado al avance de la ciencia y la tecnología.																X				
4. ORGANIZACIÓN	Esta organizado en forma lógica.																X				
5. SUFICIENCIA	Comprende aspectos cuantitativos																X				
6. INTENCIONALIDAD	Es adecuado para valorar el aprendizaje de estadística																X				
7. CONSISTENCIA	Está basado en aspectos teóricos científicos.																X				
8. COHERENCIA	Entre las variables, dimensiones, indicadores e ítems.																X				
9. METODOLOGÍA.	La estrategia responde al propósito de la investigación.																X				
10. PERTINENCIA	La escala es aplicable.																X				

1. **ASPECTOS DE EVALUACIÓN**

I. **OPINIÓN DE APLICACIÓN**

Aplicable

II. **PROMEDIO DE VALORACIÓN:** 80

LUGAR Y FECHA	DNI	FIRMA DEL EXPERTO INFORMANTE	N° DE TELEFONO
12/08/2019	29612305		959950467

TABLA DE VALIDACIÓN DE INSTRUMENTO POR EXPERTOS

Apellidos y Nombres del Informante	Institución donde labora	Nombre del Instrumento	Autor del Instrumento
Dr. Héctor Velarde Bedregal	UCSM	Cuestionario	Mg. Karina Rosas Paredes

Título de la Investigación:

Análisis exploratorio de ataques informáticos aplicando herramientas de minería de datos, para la gestión de la seguridad de redes inalámbricas en universidades de Arequipa.

		DEFICIENTE				REGULAR				BUENA				MUY BUENA				EXCELENTE			
		5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
1. CLARIDAD	Esta formulado con lenguaje apropiado.																X				
2. OBJETIVIDAD	Esta expresado en conductas observables.																X				
3. ACTUALIZACIÓN	Está adecuado al avance de la ciencia y la tecnología.																X				
4. ORGANIZACIÓN	Esta organizado en forma lógica.																X				
5. SUFICIENCIA	Comprende aspectos cuantitativos																X				
6. INTENCIONALIDAD	Es adecuado para valorar el aprendizaje de estadística																X				
7. CONSISTENCIA	Está basado en aspectos teóricos científicos.																X				
8. COHERENCIA	Entre las variables, dimensiones, Indicadores e ítems.																X				
9. METODOLOGÍA.	La estrategia responde al propósito de la investigación.																X				
10. PERTINENCIA	La escala es aplicable.																X				

1. **ASPECTOS DE EVALUACIÓN**

I. **OPINIÓN DE APLICACIÓN**

Muy buen instrumento

II. **PROMEDIO DE VALORACIÓN:** 80.0

LUGAR Y FECHA	DNI	FIRMA DEL EXPERTO INFORMANTE	N° DE TELEFONO
AQP, 12-08-2019	29204610		931854660

TABLA DE VALIDACIÓN DE INSTRUMENTO POR EXPERTOS

Apellidos y Nombres del Informante	Institución donde labora	Nombre del Instrumento	Autor del Instrumento
Dr. Guillermo Calderón Ruiz	UCSM	Cuestionario	Mg. Karina Rosas Paredes

Título de la Investigación:

Análisis exploratorio de ataques informáticos aplicando herramientas de minería de datos, para la gestión de la seguridad de redes inalámbricas en universidades de Arequipa.

		DEFICIENTE				REGULAR				BUENA				MUY BUENA				EXCELENTE			
		5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
1. CLARIDAD	Esta formulado con lenguaje apropiado.																	X			
2. OBJETIVIDAD	Esta expresado en conductas observables.																	X			
3. ACTUALIZACIÓN	Está adecuado al avance de la ciencia y la tecnología.																	X			
4. ORGANIZACIÓN	Esta organizado en forma lógica.																	X			
5. SUFICIENCIA	Comprende aspectos cuantitativos																	X			
6. INTENCIONALIDAD	Es adecuado para valorar el aprendizaje de estadística																	X			
7. CONSISTENCIA	Está basado en aspectos teóricos científicos.																	X			
8. COHERENCIA	Entre las variables, dimensiones, indicadores e ítems.																	X			
9. METODOLOGÍA.	La estrategia responde al propósito de la investigación.																	X			
10. PERTINENCIA	La escala es aplicable.																	X			

1. **ASPECTOS DE EVALUACIÓN**

I. **OPINIÓN DE APLICACIÓN**

Buen instrumento, recaba información importante

II. **PROMEDIO DE VALORACIÓN:** 85

LUGAR Y FECHA	DNI	FIRMA DEL EXPERTO INFORMANTE	N° DE TELEFONO
Arequipa 30/07/2019	29591972		954182554