Universidad de Cienfuegos

Facultad de Ingeniería

Carrera de Ingeniería Informática

Título:

Análisis de la fortaleza de las contraseñas bajo diferentes políticas de seguridad en la red telemática de la Universidad de Cienfuegos.

Trabajo de diploma para optar por el título de Ingeniería en Informática

Autor:

Dayana Rubio Durán

Tutores:

Dr. C. Eduardo R. Concepción M.
M. Sc. Boris Pérez Cañedo

Cienfuegos, Cuba
Curso 2016 – 2017

Declaración de autoría

Declaro que soy la única autora de este trabajo y autorizo al Departamento de Redes y al Departamento de Informática de la Facultad de Ingeniería en la Universidad de Cienfuegos “Carlos Rafael Rodríguez”, para que hagan el uso que estimen pertinente con el trabajo de diploma.

Para que así conste firmo la presente a los _____ días del mes de_____ del _____.

_____

Firma Autor

Dayana Rubio Durán

Los abajo firmantes certificamos que el presente trabajo ha sido revisado según acuerdo de la dirección de nuestro centro y el mismo cumple los requisitos que debe tener un trabajo de esta envergadura referente a la temática señalada.

_____ _____

Firma Tutor
Dr. C. Eduardo R. Concepción M.

Firma Tutor
M.Sc. Boris Pérez Cañedo

_____ _____

Firma ICT

Firma Vicedecano

Pensamiento

“Investigar es ver lo que todo el mundo ha visto, y pensar lo que nadie más ha pensado”.

Albert Szent–Gyorgyi

Agradecimientos

Le agradezco infinitamente a mi papá, que es sin dudas el principal responsable en mi formación personal y profesional, por creer y confiar siempre en mí, por darme su apoyo y por el amor incondicional que me brinda, por hacer de mí la persona que soy hoy. Porque es lo más importante que tengo y es mi razón de ser, ha sido mi guía y sostén y se ha sacrificado muchos años para contribuir con la realización de mi carrera.

A mi abuela, aunque no se encuentre físicamente conmigo, la llevo siempre presente, por sus consejos, por hacerme una persona de bien, por la luz a lo largo del camino y por su amor y apoyo incondicional estaré siempre agradecida.

A Puti, por apoyarme y ayudarme en todo momento.

A mi novio, por ser esa persona especial que me brinda cada día su amor y confianza, formando parte esencial de mi vida.

A mi madrasta, a mis tías y a mis tíos por contar siempre con su ayuda y dedicación.

A mis queridísimos tutores Eduardo Concepción y Boris Pérez Cañedo, que no tengo palabras como agradecerles todo lo que han hecho por mí; gracias por su guía, esmero, paciencia y su apoyo incondicional durante todo este tiempo.

A mis compañeros de grupo por todo su apoyo y ayuda en estos años compartidos, en especial a Yamila, Beatriz, Noilet y Jorge Carlos.

Al tribunal por sus acertados consejos.

A todos los profesores que me han preparado en conocimientos, enseñanzas y me han dado su ayuda.

A cada persona que de una forma u otra contribuyó con su esfuerzo al éxito de este trabajo.

Dedicatoria

A mi papá Orestes

A mi abuela Miriam

Resumen

La presente investigación titulada “Análisis de la fortaleza de las contraseñas bajo diferentes políticas de seguridad en la red telemática de la Universidad de Cienfuegos” aborda el problema de la fortaleza de las contraseñas y su afectación por la utilización de diferentes mecanismos de adaptación por parte de los usuarios.

Como resultado del estudio realizado, se estableció la prevalencia de los mecanismos de adaptación ante diferentes políticas de contraseñas. Se identificó que la mayor prevalencia ocurre en la política Exhaustiva 8.

Se constató que los mecanismos de adaptación afectan las contraseñas, como una medida de seguridad, en todas las políticas.

Del trabajo emanaron una serie de recomendaciones entre las cuales se encuentra la divulgación entre los usuarios de los resultados del trabajo, en particular, los patrones estructurales más comunes, así como los mecanismos de adaptación más usados, como una vía para contribuir a su educación en los temas de seguridad informática.

Abstract

This study, “Análisis de la fortaleza de las contraseñas bajo diferentes políticas de seguridad en la red telemática de la Universidad de Cienfuegos”, deals with the problem of password strength, and how coping mechanisms used by users affect password strength.

As a result, prevalence of coping mechanisms under different password policies was established. Comprehensive 8 was identified as the most susceptible policy.

It was established that coping mechanisms erode password under all policies.

Several recommendations were made: to socialize results among users, specially the most common structural patterns; to socialize the most common coping mechanisms; in order to contribute to support user education.

TABLA DE CONTENIDOS

ÍNDICE DE TABLAS

ÍNDICE DE FIGURAS

Introducción.

La Seguridad Informática es un tema imprescindible, puesto que el mundo está digitalizado a todos los niveles: personal, empresarial y gubernamental. La era de la información, con Internet como principal exponente, ha mejorado el nivel de vida y la convivencia en este planeta. La información es poder, luego hay que protegerla.

Todos los servicios avanzados que se disfrutan están implementados sobre sistemas informáticos que utilizan ordenadores con un hardware y un software concretos, configurados adecuadamente y conectados entre sí y con los clientes mediante redes de comunicaciones. Cualquiera de estos elementos es susceptible de ser atacado por un saboteador o, simplemente, fallar. La seguridad informática intenta evitarlo y, en caso de que ocurra, minimizar los daños para recuperar el servicio lo antes posible. Tenemos que asumir que la seguridad total es imposible. Ante cualquier barrera física o lógica, el atacante buscará una forma de romperla o rodearla, por lo que es necesario tomar todas las medidas posibles para garantizar la integridad de nuestros sistemas y de nuestros datos. En este sentido, la seguridad informática se enfoca en la protección del almacenamiento, procesamiento y transmisión de la información digital.

La seguridad informática es un elemento clave en cualquier organización y, dentro de ella, la autenticación de usuarios constituye un elemento muy importante. La autenticación permite asegurarse con un nivel de confianza razonable de la identidad del usuario. Para analizar la autenticación se clasifica las medidas adoptadas según tres criterios: Algo que sabes. Para acceder al sistema necesitas conocer alguna palabra secreta: la típica contraseña. Algo que tienes. En este caso es imprescindible aportar algún elemento material: generalmente una tarjeta. Algo que eres. El sistema solicita reconocer alguna característica física del individuo (biometría): huella dactilar, escáner de retina, reconocimiento de voz, etc.

La contraseña es el mecanismo más típico. Aplicando la estrategia “algo que sabes”, la pantalla inicial del sistema espera que la persona introduzca el nombre de un

usuario y la contraseña asociada a ese usuario. Mientras lo teclea, el nombre del usuario es visible pero la contraseña no (se suele sustituir por asteriscos, guiones, etc.), para evitar que la vea alguien que se encuentre a nuestra espalda. Si se equivoca, bien porque el usuario no existe, bien porque la contraseña no es la correcta, el sistema le impide la entrada y deja intentarlo de nuevo. Algunos sistemas ofrecen una pista sobre la contraseña (si se puso la última vez que cambiamos la contraseña), y la mayoría tiene un límite de intentos. Alcanzado ese límite, el sistema se puede bloquear durante un tiempo o definitivamente. Con este límite evitamos ataques de fuerza bruta que prueben una a una todas las combinaciones de letras, números y caracteres especiales.

Durante décadas el método de seguridad más usado para autenticarse han sido las contraseñas. La contraseña es una información secreta que se nos solicita para acceder a algún tipo de recurso, y que solo debe conocer el propietario del mismo. El identificador y la contraseña son el par de autenticación más conocido. Simple, robusto, incluso rústico, pero el nivel de seguridad depende directamente de la complejidad de la contraseña. Las contraseñas simples son escasas, y las contraseñas demasiado complejas conducen a los usuarios a aplicar estrategias no siempre correctas para gestionarlas.

Toda organización debe tener una política de seguridad para regular el uso de la información y de los sistemas que la tratan con el fin de mitigar el riesgo de pérdida, deterioro o acceso no autorizado a la misma. Por lo general, poseen una política que dicta normas sobre la especificación de las contraseñas para las cuentas registradas, y cómo estas contraseñas deben ser gestionadas (por ejemplo, con qué frecuencia deben ser cambiadas). Las políticas de contraseñas en las organizaciones pueden ser poco efectivas, y por lo tanto inseguras, si no se considera la experiencia del usuario. Por ejemplo, las políticas pueden determinar que los usuarios cambien regularmente sus contraseñas, cuestión esta que puede complicar los intentos de recordar las mismas.

Es muy común olvidar las contraseñas, especialmente cuando se deben recordar muchas a la vez. Para mitigar el impacto del olvido de las contraseñas, se

proporcionan algunos mecanismos de recuperación. Estos por lo general son de dos tipos: (1) auto-servicio; la autenticación se realiza a través de un factor alternativo, como son las preguntas de seguridad, y/o; (2) Servicio de asistencia, donde un técnico ayuda en el proceso de restablecimiento de la contraseña. Para un individuo, el restablecimiento de una contraseña puede requerir cierto esfuerzo adicional. Esta situación puede generar actitudes que afecten de manera negativa la seguridad informática. Por ejemplo, surge la posibilidad que los usuarios utilicen contraseñas débiles y fácilmente predecibles.

Ante la necesidad de cambiar de manera frecuente las contraseñas, los usuarios recurren a diferentes mecanismos de adaptación, que pueden incluir, por ejemplo, utilizar datos personales en calidad de contraseña (nombres de mascotas, fechas significativas, nombres de familiares, ciudades o localidades con vinculaciones personales, aficiones, etc.), una serie de letras dispuestas adyacentemente en el teclado (qwerty) o siguiendo un orden alfabético o numérico (123456, abcde, etc.), la escritura de la misma en un papel, la reutilización de una contraseña anterior, o tomar la anterior y poner un "1" al final. Las contraseñas deben ser algo que se pueda recordar, pero también deben ser seguras.

Aun así, siempre conviene utilizar contraseñas no triviales: palabras que no aparezcan en el diccionario de ninguna lengua, combinar letras mayúsculas con minúsculas, números, signos de puntuación, etc. Y cambiar la contraseña regularmente, porque no sabemos cuánto tiempo llevan intentando atacarla. Los sistemas operativos permiten obligar al usuario a cumplir todas estas normas.

En este sentido, la concientización del usuario para gestionar de modo eficiente su información tiene uno de sus pilares en la correcta gestión y creación de las contraseñas que este ha de utilizar en la mayoría de los procesos y operaciones que requieren de su autenticación. La importancia de la utilización y robustez de las contraseñas es muy elevada.

La situación descrita anteriormente es aplicable también a nuestro país, y particularmente a la Universidad de Cienfuegos. La Resolución No. 127/2007

establece el marco legal que rige la seguridad informática en todas las instituciones del país. En este sentido, se establece la obligatoriedad del diseño, implantación y actualización del Sistema de Seguridad Informática para las tecnologías de la información, así como la elaboración del Plan de Seguridad Informática a partir del Sistema diseñado.

La Universidad de Cienfuegos posee un Plan de Seguridad Informática que establece los principios organizativos y funcionales de la actividad de seguridad informática en el centro.

Como se planteó con anterioridad, e independientemente de las políticas establecidas, el comportamiento de los usuarios ante los cambios de contraseñas puede representar un riesgo potencial para la seguridad.

La Universidad de Cienfuegos, ubicada en la provincia de Cienfuegos, cuenta con un Departamento de Redes, el cual es el encargado de velar por la Seguridad Informática del cetro, por lo que le es de gran interés analizar la fortaleza de las contraseñas que emplean los usuarios. La autenticación es la primera línea de defensa, aunque existen diferentes alternativas de autenticación. En la Universidad se utiliza la combinación nombre de usuario/contraseña. Los usuarios para crear o cambiar sus contraseñas tienen que regirse por una política de seguridad, sin embargo, cuando esta es demasiado compleja el usuario acude a mecanismos de adaptación para evitar olvidar su contraseña, lo que trae consigo un efecto negativo sobre la fortaleza de la misma, así como en la Seguridad Informática.

Teniendo en cuenta la situación anterior se plantea como **problema a resolver**:

¿Qué efecto ejerce en la fortaleza de las contraseñas la utilización de mecanismos de adaptación por parte de los usuarios de la red telemática de la Universidad de Cienfuegos, bajo diferentes políticas de seguridad?

Se define como **objeto de estudio** el siguiente**:**

El proceso de autenticación basado en contraseñas que utilizan los usuarios de la red telemática de la Universidad de Cienfuegos.

El **campo de acción** se enmarca en:

La fortaleza de las contraseñas generadas por los usuarios de la red telemática de la Universidad de Cienfuegos.

El **objetivo general** de esta investigación consiste en:

Establecer la relación entre los mecanismos de adaptación que utilizan los usuarios de la red telemática de la Universidad de Cienfuegos y la fortaleza de las contraseñas.

Para dar cumplimiento al objetivo general se plantean las siguientes **preguntas de investigación**:

1. ¿Cuáles mecanismos de adaptación utilizan los usuarios antes los cambios de contraseñas?
    - Información conocida.
    - Reutilización parcial o completa de información.
    - Patrones fácilmente identificables.
2. ¿Cómo se comporta la utilización de los mecanismos de adaptación en el transcurso de los cambios según la política de contraseña?
3. ¿Disminuye la fortaleza de las contraseñas con la utilización de mecanismos de adaptación a un nivel que pueda afectar las políticas de seguridad?

Las **tareas** a realizar son:

1- Revisión y análisis de la bibliografía concerniente a la Seguridad Informática.
2- Entrevista con los profesores del departamento de redes de la Universidad de Cienfuegos.

3- Elaboración de un software que posibilita recopilar cambios de contraseñas por parte de los usuarios rigiéndose por una política de seguridad.
4- Recopilación y procesamiento de la información asociada a las contraseñas de los usuarios.
5- Análisis detallado de los resultados obtenidos.
6- Elaboración del documento final de la investigación.

Significación y aporte:

Ganar en comprensión acerca de cómo los cambios de contraseñas de los usuarios de la red telemática de la Universidad de Cienfuegos pudieran afectar las políticas de seguridad.

Los resultados del trabajo permitirán implementar un conjunto de acciones para contribuir a la educación de los usuarios, y elaborar recomendaciones para la actualización de las políticas de seguridad.

La tesis está estructurada en introducción, 3 capítulos, conclusiones y recomendaciones.

Capítulo I: Fundamentación teórica. Se exponen los principales conceptos y terminología: contraseñas, tipos de posibles ataques, entropía, mecanismos de adaptación. Se incluye una revisión bibliográfica: panorámica sobre los trabajos de investigación realizados en el campo del análisis de la fortaleza de las contraseñas.

Capítulo II: Diseño experimental. Se presenta la metodología de trabajo utilizada. Se incluye la descripción del proceso de recolección de datos, las variables estudiadas, el proceso de análisis propuesto, así como las herramientas utilizados.

Capítulo III: Presentación y análisis de los resultados. En este capítulo se muestran y analizan los resultados obtenidos durante el estudio y se presentan las posibles acciones derivadas de estos análisis.

Capítulo I: Fundamentación teórica.

En el presente capítulo exponen los principales conceptos y terminología: contraseñas, tipos de posibles ataques, entropía, mecanismos de adaptación. Se incluye una revisión bibliográfica: panorámica sobre los trabajos de investigación realizados en el campo del análisis de la fortaleza de las contraseñas.

1.1 La seguridad informática.

La seguridad informática tiene como propósito fundamental la protección del almacenamiento, el procesamiento y la transmisión de la información.

En el artículo 2 de la Resolución No. 6/96 se plantea que “la seguridad informática, es el conjunto de medidas administrativas, organizativas, físicas, técnicas, legales y educativas dirigidas a prevenir, detectar y responder a acciones que pongan en riesgo la confidencialidad, integridad y disponibilidad de la información que se procese, intercambie, reproduzca y conserve a través de las tecnologías de información” [1, p. 3].

Precisamente, la seguridad de las Tecnologías de la Información y las Comunicaciones descansa en esos tres pilares:

- Confidencialidad**:** condición que asegura que la información no pueda estar disponible o ser descubierta por o para personas, entidades o procesos no autorizados [2, p. 19].

  Los mecanismos de control de acceso contribuyen a la confidencialidad. Un buen ejemplo es la criptografía que se utiliza para proteger mensajes secretos (aunque, a su vez, la clave utilizada se convierte en objeto de protección).

- Integridad**:** Condición que garantiza que la información sólo puede ser modificada, incluyendo su creación y borrado, por el personal autorizado. El concepto de integridad significa que el sistema no debe modificar o corromper la información que almacene, o permitir que alguien no autorizado lo haga [2, p. 20].
- Disponibilidad**:** Propiedad que garantiza que los usuarios autorizados tengan acceso a la información y activos asociados cuando se requiera. Significa que el sistema, tanto hardware como software, se mantienen funcionando y que está en capacidad de recuperarse rápidamente en caso de fallo [2, p. 20].

Por su parte, el artículo 69 de la Resolución No.127 /2007 expone que "en las redes donde se establezcan servicios de intercambio de datos o mensajes con otras redes o usuarios externos se implementarán mecanismos de seguridad que garanticen la confidencialidad, la integridad, el control de accesos, la autenticación y el no repudio, según corresponda" [2, p. 13].

1.2 Políticas de Seguridad Informática.

En general, las políticas de seguridad informática establecen qué se puede hacer y qué no se puede hacer; mientras que los mecanismos pretenden hacer cumplir las políticas [3, p. 7].

Los mecanismos pueden ser:

- Técnicos, en los cuales los controles en la computadora hacen cumplir la política. Por ejemplo, el requerimiento que el usuario suministre una contraseña para autenticarse antes de utilizar la computadora.
- Procedimentales, en los cuales los controles fuera del sistema hacen cumplir la política. Por ejemplo, deshabilitar la cuenta de un usuario por haber cometido una infracción prevista.

Según el Plan de Seguridad Informática de la Universidad de Cienfuegos, las políticas de seguridad informática se definen “como el conjunto de normas y procedimientos establecidos por una organización para regular el uso de la información y de los sistemas que la tratan con el fin de mitigar el riesgo de pérdida, deterioro o acceso no autorizado a la misma” [4, p. 7].

En el mismo documento se establece que:

> Las políticas que se definen están en correspondencia con el Decreto Ley 199/1999 del Consejo de Estado, la Instrucción 1/2007 del MININT y las Resoluciones: 1/2000 y 2/2002 del MININT, la 65/2003, 127/2007, la 85/2007 y la 33/2008; todas del MIC. La Resolución conjunta 340/2004 del MFP y el MIC. Estas están encaminadas a minimizar los riesgos, las vulnerabilidades y hacer funcionar de forma segura las aplicaciones, software y sistemas. Comprenden además las acciones a acometer por todos los usuarios de los servicios de red de transmisión de datos dentro de las áreas de la UCF, las Filiales Pedagógicas y de aquellos que tengan vínculos de trabajo con esta. [4, p. 7].

Un elemento muy importante dentro de las políticas es el papel que juegan los usuarios; aspecto que también es contemplado en las legislaciones vigentes.

Así, por ejemplo, la Resolución 127/2007 plantea lo siguiente:

> ARTÍCULO 65**:** Los usuarios que han recibido la autorización para el empleo de los servicios que brindan las redes son responsables por su propia conducta. Los usuarios deben conocer las políticas de seguridad para las computadoras y redes a que ellos acceden y están en la obligación de cumplir estas políticas [2, p. 12].

1.2.1 Políticas de contraseñas.

A pesar de que se han desarrollado otros métodos de autenticación como la basada en imágenes, las contraseñas siguen siendo la alternativa viable para la mayoría de los sistemas de información [5].

Para evitar que los usuarios usen contraseñas débiles, las organizaciones, como parte de las políticas de seguridad informática, crean políticas de contraseñas que establecen, entre otras, la cantidad de caracteres, la complejidad, los límites de caducidad y/o la cantidad de veces que un usuario puede reutilizar la misma contraseña. Sin embargo, cuando los requisitos de contraseñas son demasiado complejos, se puede crear una situación en la que el usuario olvida su contraseña.

En el caso de la Universidad de Cienfuegos, se establecen las siguientes exigencias [4, p. 20]:

- Forzar el historial de contraseñas a 10 contraseñas recordadas.
- Longitud mínima de la contraseña: 10 caracteres.
- Vigencia máxima de la contraseña: 45 días.
- Vigencia mínima de la contraseña: 6 días.
- Almacenar las contraseñas utilizando una encriptación no reversible.
- Duración del bloqueo de cuenta: 30 minutos.
- Restablecer la cuenta de bloqueos después de 30 minutos.
- Umbral de bloqueos de cuenta: 3 intentos incorrectos de inicio de sesión.
- Activar protector de pantalla con contraseña.

Si bien las políticas de contraseñas pudieran mejorar la seguridad de la información, existe un desafío ya que los usuarios pueden tener dificultades para recordar las contraseñas. El uso de contraseñas fuertes y su constante cambio puede tener efectos contraproducentes, ya que pone demasiada carga cognitiva en los usuarios [5].

Otras políticas de contraseñas comunes en las organizaciones, y que serán objeto de análisis en el presente trabajo son las siguientes [6]:

- **Básica16**: El usuario debe generar una palabra o frase de 16 caracteres como mínimo. No se exige el uso de mayúsculas, minúsculas, números o caracteres especiales, pero no se impide su uso.
- **Exhaustiva8**: Requiere una contraseña de ocho caracteres de longitud y debe incluir al menos un: carácter especial, número, letra mayúscula y letra minúscula. El usuario es entonces libre de incluir cualquier carácter adicional para satisfacer el requisito de longitud de ocho caracteres.
- **Blacklist Hard:** Requiere una contraseña con una longitud de ocho caracteres y que no contenga una palabra del diccionario. El usuario es libre de incluir o excluir números, caracteres especiales, letras mayúsculas y minúsculas en su generación de contraseñas bajo la política *Blacklist Hard*.

1.3 La autenticación de usuarios.

La confidencialidad intenta que la información solo sea utilizada por las personas o máquinas debidamente autorizadas. Para garantizar la confidencialidad necesitamos disponer de tres tipos de mecanismos [7, p. 15]:

- Autenticación**.** La autenticación intenta confirmar que una persona o máquina es quien dice ser, que no estamos hablando con un impostor.

- Autorización**.** Una vez autenticado, los distintos usuarios de la información tendrán distintos privilegios sobre ella. Básicamente dos: solo lectura, o lectura y modificación.

- Cifrado**.** La información estará cifrada para que sea inútil para cualquiera que no supere la autenticación.

Una de las funciones esenciales de la seguridad es el control de acceso. El objetivo principal del control de acceso es la protección de los recursos contra el acceso no autorizado, garantizando al mismo tiempo el acceso autorizado. Uno de los requisitos previos del control de acceso dentro de la seguridad de la información es la autenticación, que es la que garantiza el establecimiento de la identidad de la persona que intenta acceder a un sistema o red [5].

La autenticación es un procedimiento básico en la gestión de la seguridad de cualquier sistema de información. La Resolución No.127/2007 plantea que la autenticación es el método para comprobar la identificación de un usuario o proceso. Una vez identificado al usuario, es necesario que este demuestre de algún modo la veracidad de su identidad [2, p. 19].

En el artículo 45 de la misma resolución se estipula que “en los sistemas en que es posible el acceso por múltiples usuarios se dispondrá para cada uno de ellos de un identificador de usuario personal y único” [2, p. 10].

Para poder realizar un procedimiento de autenticación es necesario poner en práctica algunos de estos métodos: autenticación por conocimiento (algo que sabes); autenticación por pertenencia (algo que tienes); autenticación por características (fundamentada en algún rasgo físico característico de la persona, por ejemplo, verificación de voz, huellas dactilares, patrones de iris, etc.) [7, p. 16].

De estas tres categorías, el método de autenticación de usuario más utilizado es autenticación basada en el conocimiento. La autenticación de usuarios basada en el conocimiento puede dividirse en categorías diferentes, que incluyen (a) basado en caracteres, (b) basado en imágenes, y (c) basado en preguntas/respuestas.

Las contraseñas se incluyen en la categoría de preguntas/respuestas y son el método más utilizado de autenticación en los sistemas de información [5, p. 13].

1.3.1 La autenticación basada en contraseñas.

La autenticación basada en contraseñas se basa, en esencia, en un secreto compartido (las credenciales, por ejemplo, la contraseña) que sirve para identificar un usuario (que reclama una identidad) ante un sistema (que verifica la contraseña). La autenticación se reduce entonces al proceso de determinar la validez de las credenciales que se utilizan para reclamar una identidad [8].

En general, el protocolo de autenticación consiste en una serie de mensajes que se intercambian entre el reclamante de una identidad y un verificador, y que permite demostrar que el reclamante posee y tiene el control sobre una o varias credenciales válidas para establecer su identidad. Opcionalmente, esta serie de mensajes permite demostrar que el reclamante se está comunicando con el verificador deseado [8].

Obviamente, el proceso de autenticación no demuestra si la persona que suministra las credenciales es la correcta o si es un impostor que ha logrado conseguir esas credenciales.

A pesar de todas las críticas, este método de autenticación sigue siendo el más popular. En un reciente artículo, los autores [9] realizan un interesante análisis sobre las alternativas para reemplazar las contraseñas. Allí se propone un conjunto de criterios relacionados con la usabilidad, el despliegue y la seguridad. Los resultados que se presentan, aunque limitados a la autenticación en la web, brindan una perspectiva interesante del por qué la popularidad del uso de las contraseñas como método de autenticación.

1.3.2 Implementación de sistemas basados en contraseñas.

La variante más simple de implementación de un sistema basado en contraseñas consiste en mantener una tabla que asocia a cada nombre de usuario con su

contraseña. El usuario suministra nombre y su contraseña, se realiza una búsqueda en dicha tabla y si se encuentra, el usuario es autenticado.

Esta variante tiene el problema que si un atacante compromete al sistema y logra obtener la tabla de nombres de usuario y contraseñas obtendría acceso pleno usando cualquiera de las credenciales almacenadas en texto plano.

Para mejorar la seguridad de este esquema es posible almacenar las contraseñas criptográficamente protegidas. En este sentido se utilizan las denominadas funciones *hash* criptográficas que permiten codificar la contraseña y tienen la propiedad de ser muy difícil de invertir. Entre ellas se encuentran las conocidas, comúnmente utilizadas, y no recomendadas, MD5 y SHA1, y las más recomendadas *bcrypt*, *scrypt* y PBKDF2 [10, p. 47].

Además del almacenamiento de las contraseñas, es necesario tener en cuenta su transmisión por la red cuando sea el caso. Aquí la recomendación es transmitir las contraseñas solamente sobre TLS [10, p. 15].

Otras variantes para la implementación de sistemas basados en contraseñas son los denominados métodos multifactoriales de autenticación que incluyen, entre otros, al protocolo de reto – respuesta, los rasgos biométricos, la localización del usuario, así como sus combinaciones [3, pp. 186-195].

1.3.3 Ataques a los sistemas basados en contraseñas.

Independientemente de las diversas formas de implementación de los sistemas basados en contraseñas, los posibles ataques se concentran en los siguientes:

- Enumeración de usuarios.
- Predicción de contraseñas.

1.3.3.1 Enumeración de usuarios.

Este paso está dedicado a mejorar la eficiencia del proceso de predicción de contraseñas. El principal objetivo es desechar nombres de usuario no válidos en una aplicación dada. Los atacantes utilizan toda la información disponible. El estudio de la información de los perfiles de los usuarios, la manera en que algunas aplicaciones sugieren los nombres de usuario y otras, son elementos que pueden ser utilizados para tratar de adivinar los nombres de los usuarios. Supongamos que nuestra aplicación emite mensajes específicos a la hora de la autenticación como los siguientes:

- El nombre de usuario es incorrecto.
- La contraseña es incorrecta.
- La combinación “usuario/contraseña” es incorrecta.

En este caso, si no se logra ingresar correctamente, se recibe una pista adicional acerca de qué es lo que no está correcto. Esto es algo que se debe evitar, es decir, se debe emitir un mensaje genérico que no brinde información acerca del posible error.

El proceso de registro de nuevos usuarios en algunas aplicaciones brinda posibilidades para la determinación de nombres de usuarios. Si se ingresa un nombre ya existente, el sistema reporta sobre la coincidencia. De esta manera ya se tiene un nombre de usuario válido.

Cuando se brinda al usuario la posibilidad de especificar su nombre de usuario, por lo general se siguen patrones predecibles. Por ejemplo, un usuario de nombre Juan Pérez podría sugerir variantes tales como juan, juanp, jperez, etc. Esto puede facilitarle el trabajo a un atacante. Según sea el caso, podríamos pensar en utilizar nombres de usuario menos evidentes. La utilización de la tecnología CAPTCHA

(*Completely Automated Public Turing test to tell Computers and Humans Apart*) puede ayudar a mitigar estos riesgos. El bloqueo de cuentas es otra alternativa para minimizar estos riesgos. Después de varios intentos (de 3 a 5, por ejemplo) la cuenta se bloquea. Los mecanismos de desbloqueo pueden variar también. En algunos casos se define un período de tiempo para el desbloqueo (30 minutos, 1 hora, 24 horas, etc.). En otros casos se requiere la acción directa del administrador del sitio. Existen muchas otras variantes, o vulnerabilidades de nuestras aplicaciones, que pueden ser utilizadas por un atacante para predecir los nombres de usuarios. Finalmente, debemos señalar que la posibilidad de determinar los nombres de los usuarios es un riesgo que muchas aplicaciones asumen, aun conociendo que es una cuestión que puede facilitar un ingreso ilegal.

1.3.3.2 Predicción de contraseñas.

Esto constituye la pesadilla de los esquemas de autenticación basados en la combinación nombre de usuario/contraseña. Estos ataques pueden ser realizados de manera manual o de manera automatizada. Para ellos existen diferentes enfoques y herramientas.

Un ataque de diccionario es un evento malicioso en el que un atacante construye una base de datos con varias combinaciones de posibles contraseñas, denominadas "el diccionario". El atacante entonces intenta iniciar sesión en el sistema usando las contraseñas de esa base de datos; si una contraseña falla, el atacante pasa a la siguiente hasta que se hayan agotado todas las opciones de la base de datos o se bloquee el sistema. Dicho proceso puede ser automatizado utilizando código para acelerar los senderos de ataque, incluyendo el tiempo común de retraso para superar los bloqueos del sistema. Los ataques de diccionario pueden ser ataques de diccionario fuera de línea, si son no interactivos o ataques de diccionario en línea si están en línea e interactivos. El ataque de fuerza bruta como un ataque que ocurre cuando todas las combinaciones posibles de letras, números y símbolos son utilizadas en un esfuerzo para adivinar una contraseña.

En cuanto a las estrategias, se puede utilizar un enfoque primero en profundidad o uno primero a lo ancho. En el primer caso se prueban todas las combinaciones de contraseñas para un usuario dado. Esto puede conducir a bloqueos de cuentas por sobrepasar la cantidad de intentos. En el segundo caso se chequean diferentes nombres de usuarios para una misma contraseña. Este último enfoque es menos propenso a bloqueos de cuentas. Este tipo de ataque automatizado puede conducir a una denegación de servicios debido a la gran cantidad de peticiones que recibiría el servidor.

Los sistemas automatizados utilizan diccionarios de contraseñas para sus intentos. Se tienen en cuenta las contraseñas más comunes utilizadas por los usuarios: 1234567, admin, <nombre-de usuario-alguna-fecha>, <nombre-de-la-empresa-año-actual>. También se utilizan palabras comunes del idioma y sus combinaciones. No es sorprendente que estos ataques puedan tener éxito, si se tiene en cuenta que muchos usuarios no le prestan suficiente atención a la selección de sus contraseñas.

Otro elemento que facilita estos ataques es el hecho de no renovar de manera frecuente las contraseñas, así como utilizar la misma contraseña para ingresar en diferentes sitios. La mejor medida contra estos ataques es una política de contraseñas fuertes, unido a una fuerte política de bloqueo de cuentas. También se utilizan alternativas más fuertes, como el hecho de enviar un mensaje de texto con un número de confirmación que debe ser ingresado en la aplicación para una autenticación exitosa.

A pesar de esto, las contraseñas tienen muchas limitaciones, sufren de problemas de seguridad y usabilidad. Debido a que los usuarios tienen limitaciones en la memoria a largo plazo, tienden a utilizar diferentes mecanismos de adaptación, ejemplo ellos son: contraseñas cortas, fácilmente predecibles pues le posibilita al usuario recordarlas con facilidad, además los usuarios incluyen dígitos y símbolos al final de estas y letras mayúsculas al inicio, toman como base palabras y frases predecibles, incluyendo nombres, fechas, letras de canciones, relacionadas

semánticamente, utilizan patrones de teclado así como: QAZXSW, QWERTY, 12345678, realizan sustituciones de caracteres por ejemplo, remplazando la (a) por @, reutilizan contraseñas de una cuenta a otra, aunque en ocasiones no en su totalidad, solo con pequeños cambios por lo general predecibles, utilizan palabras de diccionario, escriben las contraseñas en algún lugar para recordarlas y suelen elegir una contraseña que sea similar a una pregunta de seguridad. El uso de contraseñas fáciles de recordar presenta un riesgo de seguridad para la organización debido a ataques, como el ataque de fuerza bruta ya antes mencionado [5, p. 4].

En el caso de que las contraseñas se encuentren criptográficamente protegidas mediante funciones *hash*, una alternativa a los diccionarios la constituye las denominadas tablas *rainbow*. Una tabla *rainbow* no es más que una tabla que pone en correspondencia contraseñas y *hashes*. Esta variante hace uso de las contraseñas más comunes, las codifica, y luego pueden ser utilizadas repetidamente sin necesidad de utilizar tiempo en la codificación. Una contramedida es la utilización de un componente aleatorio, conocido como *salt*, que se utiliza, junto a las contraseñas, como una entrada adicional a las funciones *hash* [10].

1.4 La seguridad y la fortaleza de las contraseñas.

Una cuestión muy importante a la hora de utilizar las contraseñas como mecanismo de autenticación lo constituye la seguridad. Varios aspectos influyen en la seguridad de las contraseñas: el almacenamiento seguro, la generación de las contraseñas de manera automática o por los mismos usuarios, y el robo de las contraseñas [11].

El presente trabajo se enfoca en el proceso de generación de las contraseñas, específicamente por parte de los usuarios. Las contraseñas generadas de forma automática se obtienen a partir de una secuencia pseudoaleatoria de caracteres tomados de un conjunto determinado. Cada carácter del conjunto posee la misma probabilidad de ser tomado.

Por su parte, las contraseñas generadas por los usuarios no resultan tan aleatorias. Las personas utilizan contraseñas fáciles de recordar, que incluyan elementos personales, o relacionados con su quehacer.

Es evidente que, ante un ataque de predicción de contraseñas, aquellas generadas por los usuarios tiene más posibilidades se ser descubiertas.

1.4.1 La entropía como medida de la fortaleza de las contraseñas.

Una de las medidas de la fortaleza de las contraseñas más utilizadas es la entropía [12]. Si se asume que el conjunto de caracteres posee *N* elementos, con igual probabilidad de ser tomados, y que la longitud de la contraseña es de *L* caracteres, entonces, a partir de la definición inicial de Shannon [13] se tiene que la entropía *H* es igual a:

$$H = L \log_2 N$$

La entropía así calculada se expresa en bits. Desde este punto de vista, la fortaleza depende de la longitud de la contraseña y del tamaño del conjunto de caracteres utilizado para generarla.

Sin embargo, para el caso de contraseñas generadas por los usuarios no es posible asumir el supuesto de que la selección es completamente aleatoria. En calidad de estimación de la entropía para estos casos se ha propuesto el siguiente esquema [12, p. 118].

- La entropía del primer carácter se toma igual a 4 bits.
- La entropía de los siguientes 7 caracteres se toma igual a 2 bits.
- Del noveno al vigésimo carácter la entropía se toma igual a 1,5 bits por carácter.
- Del vigésimo primero en adelante la entropía se toma igual a 1 bit por carácter.
- Se agrega una bonificación de 6 bits si se utilizan tanto mayúsculas como caracteres no alfabéticos.

- Aquellas contraseñas de longitud menor a 20 caracteres, que no contienen palabras de un extenso diccionario de posibles contraseñas, reciben una bonificación de 6 bits adicionales.

Existen otras propuestas para medir la fortaleza de las contraseñas. Por ejemplo, en [6] se estudia una medida alternativa de fortaleza de las contraseñas que se basa en el tiempo que requiere algoritmo eficiente de predicción de contraseñas en adivinar una contraseña dada, y que los autores denominan *guess number*.

1.5 Estudios relacionados con la fortaleza de las contraseñas.

Como una manera de, por una parte, cumplir con las políticas de contraseñas, y por otra, de poder recordar, muchos usuarios generan contraseñas bastante predecibles, aun cuando en sus opiniones estas pueden ser consideradas fuertes o seguras [14].

El uso de patrones predecibles representa un riesgo de seguridad ya que su conocimiento podría hacer más efectivos los ataques a los sistemas basados en contraseñas.

Los patrones predecibles o los comportamientos que sirvan para transferir información de un cambio de contraseña a otro, o que reduzca el espacio de búsqueda de una contraseña se denominan mecanismos de adaptación. En el estudio presentado por [15] se señala que una parte de los usuarios solo realizan leves transformaciones de la contraseña al realizar cambios de estas. Se comenta, además, que los mecanismos de adaptación son más frecuentes en aquellas políticas que imponen más requerimientos a los usuarios.

Otro de los principales problemas con las contraseñas consiste en que en muchas ocasiones estas son muy cortas para resistir los ataques; los usuarios incluyen dígitos y símbolos al final, y letras mayúsculas al inicio, de las contraseñas [16].

Por otra parte, los usuarios también tienden a recurrir a palabras y/o frases predecibles: nombres, fechas, números de identificación, letras de canciones, y otras de su agrado. Otro esquema utilizado son los patrones de teclado. Por ejemplo, la contraseña 123456 es de las más comunes, así como *qwerty* o *1qazxsw*.

En el estudio realizado [17] se analiza el efecto de las políticas de formación de contraseñas y la fortaleza. La fortaleza de las contraseñas se calcula a partir de la entropía. Como conclusiones se brindan una serie de recomendaciones para la composición de contraseñas.

Un mecanismo de adaptación importante es la reutilización de las contraseñas, ya sea de forma íntegra de una cuenta a otra, o de manera parcial, realizando pequeños cambios. Como alternativa se han recomendado las denominadas contraseñas mnemónicas, formadas según la siguiente regla: “Escoger una frase y memorizarla; formar la contraseña mediante la concatenación de las primeras letras”.

En un estudio reciente [18] se investigan 6 variantes de estrategias para la generación de contraseñas mnemónicas. Se evalúa la capacidad de memorización y se brindan recomendaciones para incrementar la seguridad de las contraseñas resultantes.

En otro sentido, se han realizado estudios acerca de los patrones estructurales de las contraseñas. [19] han realizado un estudio a partir de contraseñas que fueron reveladas de los sitios RockYou y Yahoo!, y concluyen que las contraseñas estudiadas en su mayoría poseen un estructura conformada con sustantivos y adjetivos muy parecidas al idioma inglés.

A los efectos del presente trabajo, resulta importante el método propuesto por [20] para el aprendizaje de patrones estructurales presentes en una lista de contraseñas mediante la utilización de gramáticas probabilísticas libre de contexto (PCFGs – *Probabilistic Context-free Grammars*), así como un algoritmo para generar contraseñas más probables. Esta estrategia ha sido reconocida como el estado del

arte en la revelación de contraseñas [6]. En particular, se utiliza la implementación de esta estrategia que se propone en [21].

Capítulo II: Diseño experimental.

En este capítulo se presenta la metodología de trabajo utilizada. Se incluye la descripción del proceso de recolección de datos, las variables estudiadas, el proceso de análisis propuesto, así como los instrumentos utilizados.

2.1 Recolección de datos.

Para realizar el análisis de la fortaleza de las contraseñas se diseñó un esquema de recolección de datos que involucró un conjunto de estudiantes voluntarios. Como parte del estudio, se registran seis cambios de contraseñas de cada usuario. A cada usuario se le asigna una política de contraseñas. Cada contraseña utilizada se guarda junto a la fecha y la hora con el objetivo de establecer un orden cronológico. Una vez culminado este proceso, el usuario debe responder un conjunto de preguntas relacionadas con las contraseñas que ha utilizado, tanto en el ejercicio, como fuera de él.

Para implementar el esquema de recolección de datos se desarrolló una aplicación con los siguientes requisitos funcionales:

R1: Registrar.

R2: Iniciar Sesión.

R3: Cerrar Sesión.

R4: Modificar el perfil.

R5: Ver el perfil.

R6: Listar las contraseñas utilizadas.

R7: Cambiar la contraseña.

R8: Llenar el cuestionario.

R9: Eliminar un usuario.

R10: Cambiar contraseña de un usuario.

R11: Modificar los datos de un usuario.

R12: Ver los datos de un usuario.

R13: Crear un usuario.

R14: Listar los usuarios.

R15: Buscar usuarios.

R16: Exportar a Excel los datos de los usuarios.

R17: Eliminar un registro de contraseña.

R18: Listar las contraseñas de los usuarios.

R19: Buscar contraseñas.

R20: Exportar a Excel los registros de las contraseñas de los usuarios.

R21: Ver los cuestionarios.

R22: Listar los cuestionarios.

R23: Eliminar un cuestionario.

R24: Exportar a Excel los cuestionarios.

El diagrama de casos de uso del sistema se muestra a continuación.

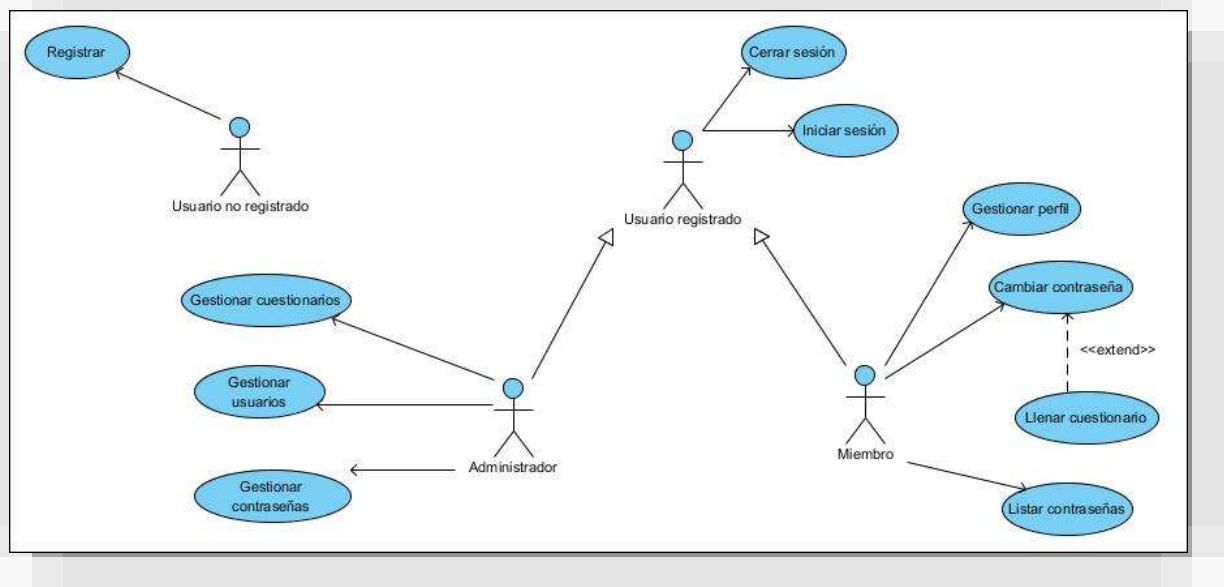


Figura 1 Diagrama de casos de uso del sistema.

Según los casos de uso del diagrama, los requisitos funcionales quedan agrupados de la siguiente manera:

R1: (Registrar).

R2: (Iniciar Sesión).

R3: (Cerrar Sesión).

R4, R5 (Gestionar Perfil).

R6: (Listar Contraseñas).

R7: (Cambiar Contraseña).

R8: (Llenar Cuestionario).

R9, R10, R11, R12, R13, R14, R15, R16: (Gestionar Usuarios).

R17, R18, R19, R20: (Gestionar Contraseñas).

R21, R22, R23, R24: (Gestionar Cuestionarios).

Cuando un usuario se registra en el sistema, debe ingresar sus datos personales. En este momento, se le asigna una de las tres políticas previstas, descritas en la sección 1.2.1 del Capítulo I, por la cual se debe regir para especificar su contraseña y realizar los cambios posteriores de la misma. Las políticas utilizadas son: Básica16; Exhaustiva8; Blacklist Hard.

Un usuario registrado, además de iniciar y cerrar sesión, puede gestionar su perfil, listar las contraseñas utilizadas, cambiar la contraseña por una nueva no utilizada con anterioridad, así como responder el cuestionario una vez realizados los seis cambios de contraseñas previstos. Si en el momento de ingresar al sistema el usuario ha olvidado la contraseña, después de tres intentos fallidos, tiene la posibilidad de recibir, vía correo electrónico, una nueva contraseña generada por el sistema con la cual puede iniciar una sesión. Estos restablecimientos de contraseña por olvido se registran para su posterior análisis.

Por su parte, el administrador puede gestionar los usuarios, las contraseñas y los cuestionarios. De esta manera se recopilan los datos que se toman como muestra para realizar el estudio previsto.

La aplicación desarrollada se encuentra alojada en pass.ucf.edu.cu. En la Figura 2 se muestra la página de inicio de la aplicación.

Figura 2. Página de inicio del sitio pass.ucf.edu.cu.

2.2 Variables utilizadas.

A los efectos del estudio, se analizan los mecanismos de adaptación que utilizan los usuarios en el momento de realizar cambios de contraseñas. En particular, se tienen en cuenta los siguientes mecanismos:

1. Utilización de información conocida, por ejemplo, nombres propios o números de identidad.
2. Repetición parcial, es decir, por ejemplo, reutilización total o parcial de contraseñas anteriores con leves modificaciones.
3. Patrones fácilmente identificables (por ejemplo, incrementos de números de años y patrones de teclado).

La relación entre los cambios de contraseñas por olvido y la utilización de mecanismos de adaptación es también analizada. Por ejemplo, ¿tienden los usuarios a utilizar patrones fácilmente recordables después del olvido de una contraseña?

El estudio descriptivo de los datos incluye los siguientes elementos:

- La distribución de la longitud de las contraseñas por políticas.
- Patrones sintácticos utilizados y su distribución de frecuencias.
- El conjunto de caracteres no alfabéticos utilizado, es decir,
    - Caracteres numéricos.
    - Símbolos especiales.

Los patrones sintácticos permiten conocer cuáles son las estructuras más comunes utilizadas por los usuarios para crear sus contraseñas. Por ejemplo, el patrón LLLLSDD, describe a las contraseñas que incluyen cuatro caracteres alfabéticos, un símbolo y dos números. Para la obtención de estos patrones se utiliza la implementación en Python de [21] y [20].

Para el caso de la detección de patrones de teclado se utiliza la biblioteca *zxcvbn*, “Estimación de la fortaleza de contraseñas”, cuya implementación en Python está disponible en https://github.com/dwolfhub/zxcvbn-pyth.

Las variables utilizadas en el presente estudio son:

1. Prevalencia de los mecanismos de adaptación.
2. La entropía como medida la fortaleza de las contraseñas.

Para el cálculo de la entropía se utiliza el modelo NIST (ver sección 1.4.1 del Capítulo I). Se incluye, además, una variante de cálculo de la entropía con adaptación secuencial (*post coping*), que fue propuesta por [15], y que tiene en cuenta la dependencia secuencial que tienen los mecanismos de adaptación en la fortaleza de las contraseñas a medida que se utilizan en el tiempo. Esta variante de cálculo básicamente penaliza cualquier reutilización, total o parcial, y/o modificación

leve de una contraseña anterior, provocando, en general, una disminución de la entropía.

Veamos un ejemplo.

Iteración 1

P	a	l	a	b	r	a	x
4	2	2	2	2	2	2	2

La entropía es de 18 bits.

Iteración 2

P	a	l	a	b	r	a	z
4							2

En esta iteración se penaliza la reutilización, manteniendo el primer carácter. Para el resto se aplica el mismo método de cálculo. De esta manera, después de la reutilización la entropía es de 6 bits.

El procedimiento para el cálculo de la entropía con adaptación secuencial tiene en cuenta los siguientes elementos:

1. Repetición de dígitos dentro de la misma contraseña. Disminuir la entropía dividiendo por la cantidad de repeticiones.
2. Repetición de contraseñas en el tiempo. Restar la entropía de la porción que se repite.
3. Incremento numérico en el tiempo. Restar la bonificación de 6 bits por usar caracteres no alfabéticos.
4. Repetición de mayúsculas o caracteres no alfabéticos. Restar la bonificación de 6 bits.
5. Cambio de minúscula a mayúscula manteniendo la misma palabra. Restar la entropía de la palabra, pero mantener la bonificación 6 bits por mayúscula.

6. Mayúscula al inicio o número o carácter especial al final de la misma palabra. Restar la palabra y mantener la bonificación de 6 bits.

2.3 Proceso de análisis.

En el proceso de análisis se establece el comportamiento de los mecanismos de adaptación desde el punto de vista de su prevalencia y su influencia en la fortaleza de las contraseñas.

Las dos primeras preguntas de investigación son las siguientes:

1. ¿Cuáles mecanismos de adaptación utilizan los usuarios antes los cambios de contraseñas?
2. ¿Cómo se comporta la utilización de los mecanismos de adaptación en el transcurso de los cambios según la política de contraseña?

Las respuestas a estas preguntas se obtienen a partir del análisis de los mecanismos de adaptación.

En primer lugar, se analiza la cantidad de mecanismos detectados por cada política e iteración. La prevalencia de los mecanismos de adaptación se calcula como el cociente de la cantidad de usuarios que acuden a mecanismos de adaptación entre el total de contraseñas en cada caso.

En segundo lugar, la prevalencia de los mecanismos de adaptación se analiza también de manera individual en cada política con el objetivo de determinar la incidencia de cada tipo de mecanismo. Aquí, la prevalencia se calcula como el cociente de la cantidad de mecanismos de un tipo dado entre el total de todas las de contraseñas en la iteración.

Para el análisis de la entropía y sus posibles efectos en las políticas, el análisis comienza por la entropía NIST promedio total del estudio. Además, se calcula la

entropía NIST promedio por cada política. El mismo análisis se realiza para la entropía con adaptación secuencial.

Para cada una de las políticas examinadas los valores mínimos de entropía son los siguientes [6]:

- Exhaustiva 8: 30 bits.
- Blacklist Hard: 24 bits.
- Básica 16: 30 bits.

Tomando estos valores como base, es posible analizar si los valores promedios obtenidos son significativamente mayores que dichos valores. Con este fin se puede utilizar una prueba de hipótesis o prueba de significación. En primer lugar, se plantea una hipótesis denominada hipótesis nula o H_0, que es la que se quiere contrastar, y una hipótesis alternativa H_1, que se contrapone a la hipótesis nula.

Para un nivel de significación α dado, se calcula el estadígrafo de la prueba, así como el valor teórico, según la distribución asociada a la hipótesis nula. La distribución nula, describe la distribución que siguen los valores en caso de que la hipótesis nula sea cierta. Adicionalmente, es posible calcular la probabilidad *p* de obtener un valor calculado del estadígrafo al menos tan pequeño como el obtenido a partir de la distribución.

De esta manera existen dos reglas de decisión:

1. Si el estadígrafo calculado es mayor que el valor teórico, entonces se rechaza la hipótesis H_0.
2. Si el valor calculado de *p* es menor que el nivel de significación α especificado, entonces se rechaza la hipótesis H_0.

La tercera pregunta de investigación es la siguiente:

¿Disminuye la fortaleza de las contraseñas con la utilización de mecanismos de adaptación a un nivel que pueda afectar las políticas de seguridad?

Su respuesta puede ser obtenida contrastando los valores obtenidos con los valores mínimos descritos con anterioridad. En este caso se utiliza una prueba *t* para una

muestra tomando como base de comparación contra las entropías mínimas según la política. Cuando los niveles calculados son significativamente menores se considera que existe una afectación a la política en cuestión.

Sea $X = \{X_1, X_2, \cdots, X_n\}$ los valores que toma la variable y α el nivel de significación. Se calculan entonces los siguientes valores.

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}$$

$$ErrorStd = \frac{S}{\sqrt{n}}$$

$$T = \frac{(\bar{X} - \mu_0)}{ErrorStd}$$

Para una prueba de dos colas, es decir,

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

Se calcula

$$pValue = 2 * P\{T_{n-1} \leq -|T|\}$$

Si $pValue \leq \propto$, se rechaza la hipótesis H_0.

También es posible utilizar una prueba de una cola utilizando, por ejemplo, las siguientes hipótesis:

$$H_0: \mu = \mu_0 \text{ o } H_0: \mu \leq \mu_0$$

$$H_1: \mu > \mu_0$$

Se calcula

$$pValue = P\{T_{n-1} \leq -T\}$$

Para las siguientes hipótesis,

$$H_0: \mu = \mu_0 \text{ o } H_0: \mu \geq \mu_0$$

$$H_1: \mu < \mu_0$$

se calcula

$$pValue = P\{T_{n-1} \leq T\}$$

En ambos casos, Si $pValue \leq \propto$, se rechaza la hipótesis H_0.

Capítulo III: Presentación y análisis de los resultados.

En este capítulo se muestran y analizan los resultados obtenidos durante el estudio y se presentan las posibles acciones derivadas de estos análisis.

3.1 Resultados generales.

Los resultados que se presentan a continuación se obtuvieron al procesar cuatro iteraciones de la seis planificadas. La explicación está dada por el hecho de que no todos los involucrados culminaron el experimento, incluyendo el llenado de los cuestionarios. De esta manera, se lograron recopilar ciento sesenta y tres contraseñas, registrándose cerca de un veinte por ciento de olvidos de contraseñas de una iteración a otra.

La participación en cada política por iteración se resume en la siguiente tabla.

Tabla 1. Participación de usuarios.

Cantidad de contraseñas			
	Exhaustiva 8	**Blacklist Hard**	**Básica 16**
Iteración 1	**23**	**24**	**24**
Iteración 2	**14**	**11**	**10**
Iteración 3	**13**	**10**	**9**
Iteración 4	**12**	**7**	**6**

La entropía promedio obtenida del estudio según el modelo NIST y el de adaptación secuencial o *post coping* se refleja en el gráfico que sigue.

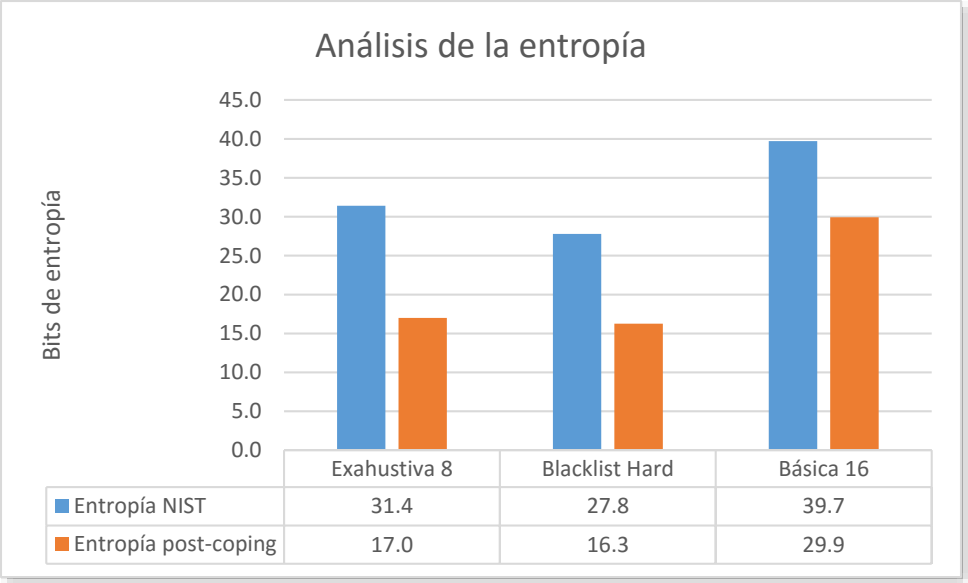


Figura 3. Análisis de la entropía.

Como se puede apreciar, la política más fuerte según el modelo NIST y el modelo de adaptación secuencial es la Básica 16, seguida de la Exhaustiva 8. También se puede observar las diferencias entre ambos modelos. Este análisis se detalla más adelante.

3.2 Análisis de patrones estructurales.

Las contraseñas utilizadas en el estudio están restringidas por los requerimientos de cada política. La distribución general de las longitudes de las contraseñas se refleja a continuación.

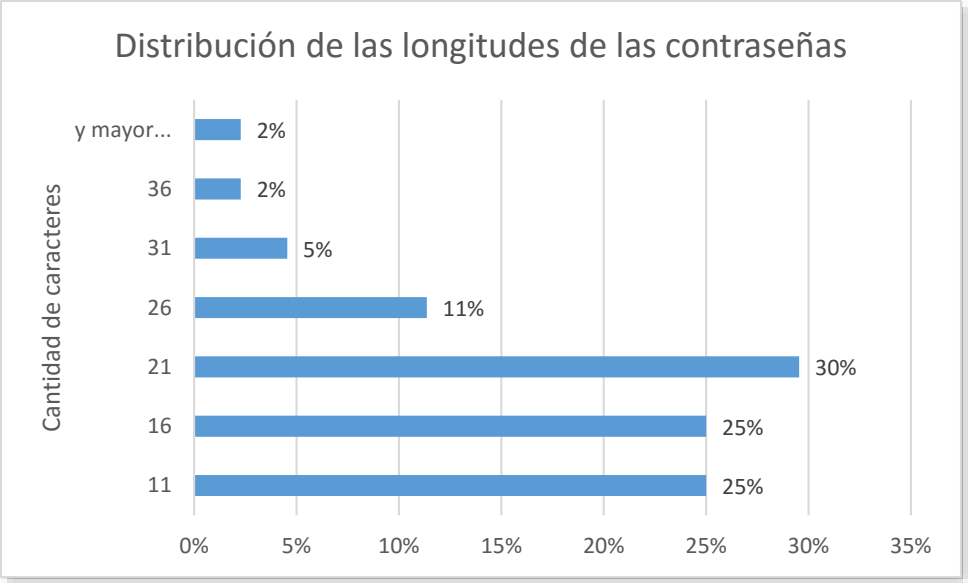


Figura 4. Histograma de las longitudes de las contraseñas.

Se puede apreciar que un alto porcentaje de los usuarios (el 80%) utilizaron contraseñas entre 11 y 21 caracteres.

En cuanto a la estructura de las contraseñas, se constató que las obtenidas se rigen, en orden de frecuencia, por los siguientes patrones:

Tabla 2. Gramática simple.

Gramática simple	
D	**0.44**
LSD	**0.173333**
LD	**0.12**
L	**0.12**
LDLD	**0.04**
LDS	**0.026667**
LSLSLSLSLSLSL	**0.013333**
LSLSLSLSL	**0.013333**
LSLD	**0.013333**
LSDLD	**0.013333**
LDSLD	**0.013333**
LDSL	**0.013333**

El patrón más frecuente se corresponde con las contraseñas asociadas al olvido que son generadas por el sistema. Estas contraseñas constan de dígitos solamente y no se consideran como generadas por los usuarios.

El patrón LSD fue el más común y se corresponde con aquellas contraseñas que comienzan con letras, seguidas de símbolos y que terminan en dígitos. Le siguen en orden la combinación de letras y dígitos; letras solamente; las formadas por letras, dígitos, letras y dígitos. Los patrones estructurales detallados pueden ser consultados en el Anexo 1.

De igual manera, los símbolos más utilizados, y sus combinaciones, fueron los siguientes.

Tabla 3. Patrones de símbolos.

Símbolos	
.	0.357143
+	0.285714
	0.178571
@	0.035714
,	0.035714
=	0.035714
.#	0.035714
+-	0.035714

Como se aprecia, el punto y el signo de adición fueron los más comunes en las contraseñas obtenidas, seguido del espacio y el símbolo de arroba.

Por su parte, en el caso de los números las combinaciones más frecuentes fueron las siguientes.

Tabla 4. Combinaciones de números.

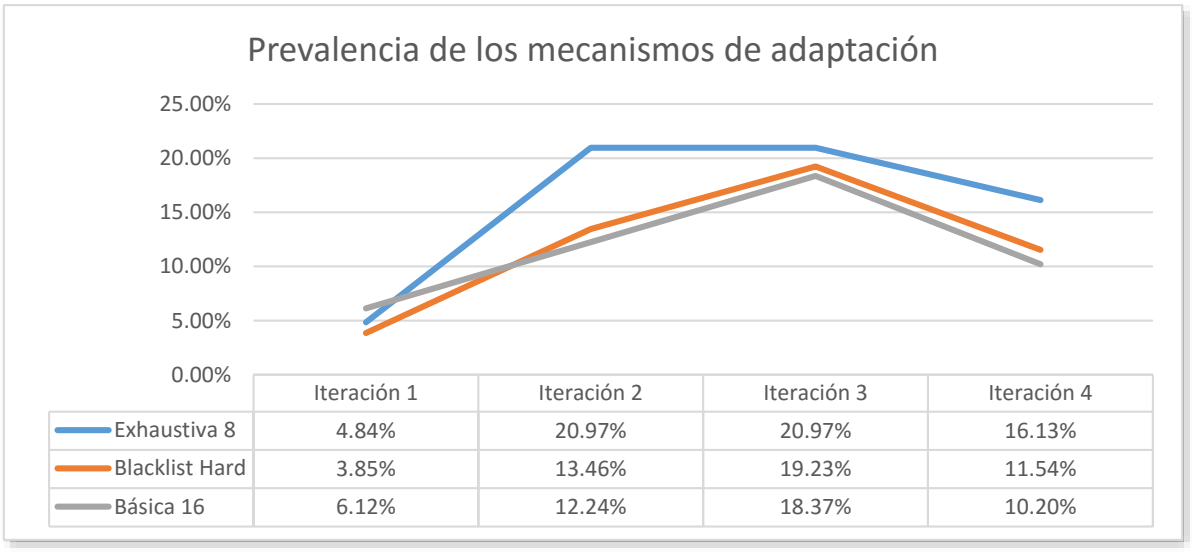
Números	
2017	0.043478
123456	0.043478
1	0.028986
09	0.028986
2034	0.028986
12345678	0.028986

El número del año actual, y el patrón de teclado 123456 fueron los más comunes. El reporte detallado se puede consultar en el Anexo 2.

3.3 Análisis de la prevalencia de los mecanismos de adaptación.

Como se describió en el Capítulo II, las preguntas de investigación 1 y 2 se responden mediante el análisis de la presencia de los mecanismos de adaptación a nivel general, de política y de iteración.

A nivel de política, la prevalencia de los mecanismos de adaptación fue la siguiente.



	Iteración 1	Iteración 2	Iteración 3	Iteración 4
Exhaustiva 8	4.84%	20.97%	20.97%	16.13%
Blacklist Hard	3.85%	13.46%	19.23%	11.54%
Básica 16	6.12%	12.24%	18.37%	10.20%

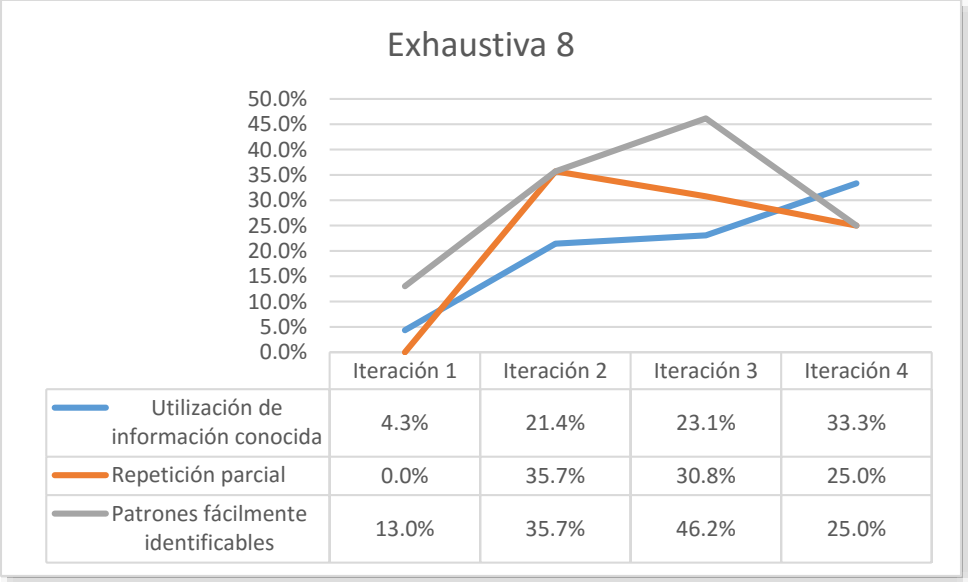
Figura 5. Prevalencia de los mecanismos de adaptación.

En el gráfico se aprecia que, ya desde la primera iteración, los usuarios acudieron a alguna estrategia de adaptación según la política a la que fueron asignados.

Se nota que la política Exhaustiva 8 exhibe una mayor prevalencia de mecanismos de adaptación que el resto. Se debe recordar que la reutilización de partes de las contraseñas o patrones de incremento no pueden contabilizarse en la primera iteración.

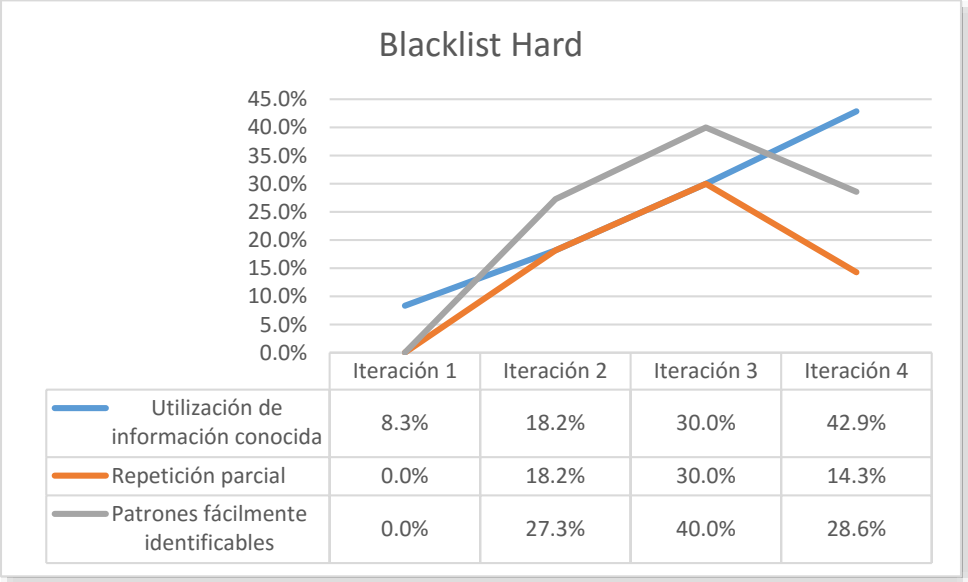
Con el objetivo de precisar más el comportamiento de los mecanismos de adaptación, se realiza un análisis a nivel de cada política que tiene en cuenta la ocurrencia de los mecanismos de adaptación en cada iteración.

Los siguientes gráficos muestran la prevalencia de los mecanismos de adaptación en cada política. La prevalencia se calcula tomando como base la cantidad de contraseñas de cada iteración.



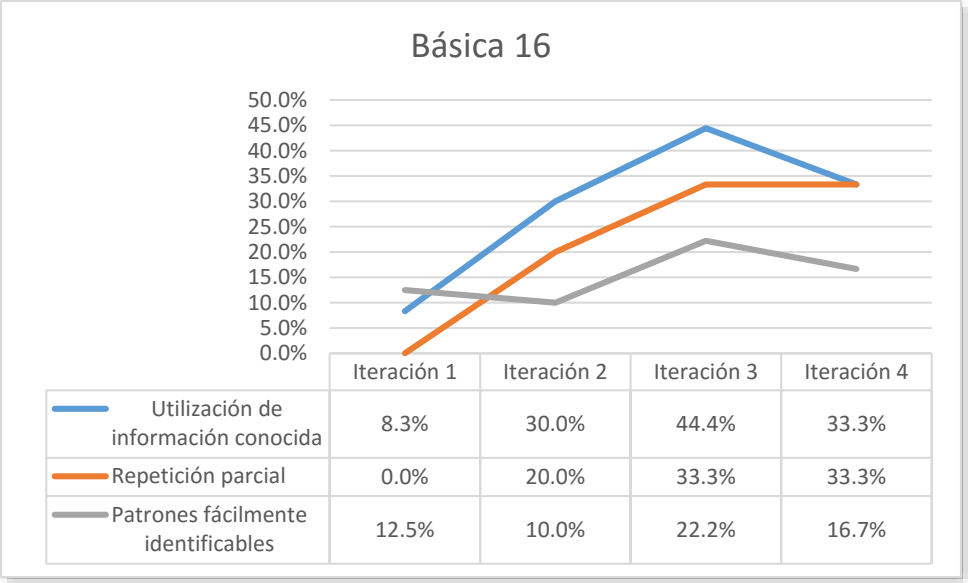
	Iteración 1	Iteración 2	Iteración 3	Iteración 4
Utilización de información conocida	4.3%	21.4%	23.1%	33.3%
Repetición parcial	0.0%	35.7%	30.8%	25.0%
Patrones fácilmente identificables	13.0%	35.7%	46.2%	25.0%

Figura 6. Prevalencia de los mecanismos de adaptación en la política Exhaustiva 8.



	Iteración 1	Iteración 2	Iteración 3	Iteración 4
Utilización de información conocida	8.3%	18.2%	30.0%	42.9%
Repetición parcial	0.0%	18.2%	30.0%	14.3%
Patrones fácilmente identificables	0.0%	27.3%	40.0%	28.6%

Figura 7. Prevalencia de los mecanismos de adaptación en la política Blacklist Hard.



	Iteración 1	Iteración 2	Iteración 3	Iteración 4
Utilización de información conocida	8.3%	30.0%	44.4%	33.3%
Repetición parcial	0.0%	20.0%	33.3%	33.3%
Patrones fácilmente identificables	12.5%	10.0%	22.2%	16.7%

Figura 8. Prevalencia de los mecanismos de adaptación en la política Básica 16.

En los gráficos anteriores se observan las diferencias entre las políticas en cuanto a la prevalencia de los diferentes tipos de mecanismos de adaptación.

Así, las políticas Exhaustiva 8 y Blacklist Hard son más susceptibles al uso de patrones fácilmente identificables que la política Básica 16, ya desde las primeras iteraciones.

En el caso de la política Básica 16 se observa una prevalencia predominante del uso de información conocida, seguida de la reutilización de información a medida que transcurren las iteraciones.

Sin embargo, se debe notar que en las políticas Exhaustiva 8 y Blacklist Hard se aprecia una tendencia al incremento del uso de información conocida a medida que avanzan las iteraciones.

3.4 Análisis de la entropía.

La tercera pregunta de investigación es la siguiente:

¿Disminuye la fortaleza de las contraseñas con la utilización de mecanismos de adaptación a un nivel que pueda afectar las políticas de seguridad?

Para su respuesta se utilizó como base de cálculo la entropía de las contraseñas, como una medida de su fortaleza.

Para el cálculo de la entropía de las contraseñas se utilizó el modelo NIST de entropía y el modelo de cálculo de entropía con dependencia secuencial, según se explicó en el Capítulo II.

La idea del segundo modelo consiste en tener en cuenta el uso de patrones de adaptación en la creación de nuevas contraseñas. Su motivación viene del hecho de que, si un atacante logra obtener la contraseña de un usuario, la comprobación de variantes de adaptación sobre la contraseña conocida puede llevar a descubrir las nuevas contraseñas, bajo el supuesto del uso de patrones de adaptación.

En los análisis que siguen se tiene en cuenta que lo requerimientos mínimos de entropía para cada política son 30, 24 y 30 bits para la política Exhaustiva 8, Blacklist Hard, y Básica 16, respectivamente.

Las pruebas de hipótesis realizadas contrastan los valores obtenidos contra estos requerimientos mínimos con el fin de establecer si se pueden señalar afectaciones a las políticas. En todos los casos se utilizó un nivel de significación del 5%.

En la tabla que sigue se muestran los valores de entropía NIST calculados para cada política y cada iteración.

Tabla 5. Valores de la entropía según el modelo NIST.

Modelo NIST	Exhaustiva 8			Blacklist Hard			Básica 16		
	N	Entropía	Desv. Estándar	N	Entropía	Desv. Estándar	N	Entropía	Desv. Estándar
Iteración 1	23	30.2	3.5	24	27.4	4.2	24	39.8	4.1
Iteración 2	14	31.1	3.2	11	28.3	4.6	10	40.1	5.2
Iteración 3	13	32	3.7	10	27.3	3.9	9	40.2	4.5
Iteración 4	12	32.3	4.1	7	28.1	4.1	6	38.7	5.1

Los resultados que se obtienen muestran que, bajo el modelo NIST, las entropías se mantienen por encima de los requerimientos mínimos de cada política.

Esta situación se explica, fundamentalmente, por el hecho de que los usuarios generalmente utilizaron contraseñas cuyas longitudes excedían los mínimos de las políticas (8, 8 y 16 respectivamente), tal y como se refleja en la sección 3.2. Según el modelo NIST, esto contribuye a aumentar la entropía.

El mero aumento de la longitud, aunque puede conllevar a un aumento de la entropía, se debe asumir con cautela.

Si el usuario reutiliza partes de contraseñas anteriores, patrones fácilmente identificables, o información conocida, la fortaleza de las contraseñas se puede ver afectada y, por ende, la política.

Con el fin de analizar esta situación se realizó el cálculo de la entropía usando el modelo de adaptación secuencial descrito en el Capítulo II.

En la tabla que sigue se muestran los resultados obtenidos para cada política en cada iteración.

Tabla 6. Valores de la entropía según el modelo de adaptación secuencial.

Modelo con adaptación secuencial	Exhaustiva 8			Blacklist Hard			Básica 16		
	N	Entropía	Desv. Estándar	N	Entropía	Desv. Estándar	N	Entropía	Desv. Estándar
Iteración 1	23	19.1	3.3	24	16.7	3.3	24	27.2	5.1
Iteración 2	14	15.3	3.1	11	14.6	3.6	10	26.2	3.2
Iteración 3	13	18.4	3.4	10	15.4	3.4	9	27.5	2.1
Iteración 4	12	15.2	3.6	7	18.5	4.2	6	26.6	2.1

Como se puede apreciar en la tabla anterior, los valores promedios son significativamente menores ($p < 0.05$) que los mínimos de cada política.

Las disminuciones en las entropías ocurren desde las primeras iteraciones. Esto está dado por el hecho de que los usuarios hacen uso de información conocida o patrones de teclado. En las siguientes iteraciones se comienza, además, a reutilizar información de contraseñas anteriores.

Estas disminuciones de los valores promedio de las entropías afectan las políticas desde el punto de vista de su seguridad bajo el supuesto de que se develen contraseñas de los usuarios.

De acuerdo a estos resultados y aún con la disminución de la entropía por debajo de los límites mínimos, la política Básica 16 protege mejor a los usuarios que el resto desde el punto de vista del modelo de entropía con adaptación secuencial.

Conclusiones

Como resultado del estudio realizado, y atendiendo a las preguntas de investigación planteadas se llegan a las siguientes conclusiones.

Pregunta 1.

- Los mecanismos de adaptación son más prevalentes en la política Exhaustiva 8 que en el resto. Desde las primeras iteraciones los usuarios utilizan información conocida y patrones fácilmente identificables. A medida que avanzan las iteraciones el comportamiento se mantiene. Esto tiene relación con la cantidad de requerimientos que impone cada política: cinco requerimientos en la política Exhaustiva 8 (mínimo ocho caracteres, una mayúscula, una minúscula, un dígito y un carácter especial), dos requerimientos para la política Blacklist Hard (mínimo ocho caracteres y no usar palabras de diccionario), y un requerimiento para la política Básica 16 (mínimo dieciséis caracteres).
- Los usuarios tienden a utilizar contraseñas cuyas longitudes sobrepasan los límites mínimos de las políticas, lo que influye en el comportamiento de las entropías desde el punto de vista del modelo NIST.

Pregunta 2.

- Los participantes en las políticas Exhaustiva 8 y Blacklist Hard se involucran más en el uso de patrones fácilmente identificables, por ejemplo, patrones de teclado y combinaciones predecibles de símbolos y números. Este hecho está asociado fundamentalmente a los requerimientos impuestos por estas políticas y a la necesidad de los usuarios de recordar sus contraseñas.
- La política Básica 16 se caracteriza por un mayor uso de información conocida.

Pregunta 3.

- Los mecanismos de adaptación afectan las contraseñas como una medida de seguridad en todas las políticas.
- El uso de mecanismos de adaptación es común a todas las políticas desde las primeras iteraciones.
- A pesar del uso de mecanismos de adaptación, la política Básica 16 resulta la más segura debido a la menor pérdida de entropía a través de las iteraciones.

Recomendaciones

Como culminación del trabajo se plantean las siguientes recomendaciones.

1. Si se toma en cuenta la utilización de mecanismos de adaptación, se recomienda incluir la menor cantidad de requerimientos en la política de contraseñas, en particular, el requisito de longitud mínima, por ejemplo, una política de contraseñas largas como Básica 16.

2. Para disminuir el uso de información conocida y la reutilización de información, se recomienda implementar un mecanismo para evitar el uso del nombre de usuario en las contraseñas, así como almacenar un historial de contraseñas y utilizarlo como una lista negra de contraseñas, que no se pueden reutilizar.

3. Divulgar entre los usuarios los resultados del trabajo, en particular, los patrones estructurales más comunes, así como los mecanismos de adaptación más usados, como una vía para contribuir a su educación en los temas de seguridad informática.

Referencias bibliográficas

[1] Ministerio del Interior, «Resolución No. 6», 1996.

[2] Ministerio de la Informática y las Comunicaciones, «Resolución 127. Reglamento de Seguridad Informática para las Tecnologías de la Información», Ministerio de la Informática y las Comunicaciones, La Habana, 2007.

[3] Bishop, M., *Introduction to Computer Security*. Boston: Pearson Education, Inc., 2005.

[4] Escalona Tillet, D., «Plan de Seguridad Informática», Universidad de Cienfuegos, 2015.

[5] Mujeye, S., «An Experimental Study on the Role of Password An Experimental Study on the Role of Password Productivity.», Doctoral Dissertation, Nova Southeastern University, 2016.

[6] Kelley, G. P. *et al.*, «Guess Again (and Again and Again): Measuring Password Strength by Simulating Password-Cracking Algorithms», en *IEEE Symposium on Security and Privacy*, 2012.

[7] Roa Buendía, J. F., *Seguridad informática*. Madrid: McGraw-Hill/Interamericana de España, S. L., 2013.

[8] Grassi, P. A., Garcia, M. E., y Fenton, J. L., «Draft NIST Special Publication 800-63-3: Digital Identity Guidelines». National Institute of Standards and Technology, US Department of Commerce, 2017.

[9] Bonneau, J., Herley, C., van Oorschot, P. C., y Stajano, F., «The Quest to Replace Passwords: A Framework for Comparative Evaluation of Web Authentication Schemes», en *Proc. IEEE Symp. on Security and Privacy*, Oakland, 2012.

[10] Woschek, M., *OWASP Cheat Sheets*. OWASP, Open Web Application Security Project, 2015.

[11] Chanda, K., «Password Security: An Analysis of Password Strengths and Vulnerabilities», *I. J. Computer Network and Information Security*, vol. 7, pp. 23-30, 2016.

[12] Burr, W. E. *et al.*, «NIST Special Publication 800-63-2: Electronic Authentication Guideline». National Institute of Standards and Technology, US Department of Commerce, 2013.

[13] Shannon, C. E., «A Mathematical Theory of Communication», *Bell System Technical Journal*, vol. 27, pp. 379-423, 1948.

[14] Blase Ur, J. B., Segreti, S. M., Bauer, L., Christin, N., y Cranor, L. R., «Do Users' Perceptions of Password Security Match Reality?», presentado en CHI'16, San Jose, CA, USA, 2016.

[15] Curnett, B. T., «Password strength analysis: User coping mechanisms in password selection.», Open Access Theses. 536, Purdue University, 2015.

[16] Bonneau, J., «The science of guessing: Analyzing an anonymized corpus of 70 million passwords», en *Proc. IEEE Symposium on Security and Privacy*, 2012.

[17] Komanduri, S. *et al.*, «Of Passwords and People: Measuring the Effect of Password-Composition Policies», en *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11). ACM*, New York, NY, USA, 2011, pp. 2595-2604.
[18] Yang, W., Li, N., Chowdhury, O., Xiong, A., y Proctor, R. W., «An Empirical Study of Mnemonic Sentence-based Password Generation Strategies», en *Proceedings of CCS'16. ACM.*, Vienna, Austria, 2016.
[19] Blase Ur, J. B. *et al.*, «Poster: The Art of Password Creation», en *Proc. of the IEEE Symp. on Security and Privacy*, 2013.
[20] Weir, M., Aggarwal, S., Medeiros, B. D., y Glodek, B., «Password cracking using probabilistic context-free grammars», en *Proc. of the IEEE Symp. on Security and Privacy*, 2009, pp. 391–405.
[21] Glodek, B., «Using a Specialized Grammar to Generate Probable Passwords», Master's of Science Thesis, Florida State University, 2008.

Anexo 1

Gramática completa	
DDDD	0.373333
LLLLLLLLLLLLLLLLL	0.04
LLLLLDDDD	0.04
LLLLLLLSDDDD	0.026667
LLLLLLLLLLLLLLLLLLDDDD	0.026667
LLLLLLLLLLLLLLLL	0.026667
LLLLLLDDDDLLLLLDDDD	0.026667
LLSDLLLDDDD	0.013333
LLLLSLLSLLSLLLLLLLLLSLLSLLLL	0.013333
LLLLSLLSLLLLLLLSLLLSLLLLLLLLLSLLSLLSLLLL	0.013333
LLLLSLLDDDDDDDDDDD	0.013333
LLLLSDDDD	0.013333
LLLLLSDDDDDDDDD	0.013333
LLLLLLSDDDDDD	0.013333
LLLLLLLSDDD	0.013333
LLLLLLLLSSDD	0.013333
LLLLLLLLSDDDD	0.013333
LLLLLLLLLSDDD	0.013333
LLLLLLLLLLSDD	0.013333
LLLLLLLLLLLSDDDD	0.013333
LLLLLLLLLLLLSDDDD	0.013333
LLLLLLLLLLLLLLLLLLSSDDDD	0.013333
LLLLLLLLLLLLLLLLLLLLLLLLLLLLLLL	0.013333
LLLLLLLLLLLLLLLLLLLLLLLLLL	0.013333
LLLLLLLLLLLLLLLLLLLLLLLLL	0.013333
LLLLLLLLLLLLLLLLLLLLLLLLLL	0.013333
LLLLLLLLLLLLLDDDDDDDD	0.013333
LLLLLLLDDDDDD	0.013333
LLLLLLLDDS	0.013333
LLLLLLLDDDLLLLLDDDD	0.013333
LLLLLLDDDDDDSS	0.013333
LLLLLLDD	0.013333
LLDDDDDDDDDDDSLL	0.013333
LLDDDDDD	0.013333
LDSLLDDDDDD	0.013333
DDDDDDDDDDDDDDDDDD	0.013333
DDDDDDDD	0.013333
DDD	0.013333
DD	0.013333
D	0.013333

Anexo 2

Patrones numéricos	
1	0.028986
8	0.014493
09	0.028986
61	0.014493
25	0.014493
23	0.014493
605	0.014493
203	0.014493
123	0.014493
106	0.014493
2017	0.043478
2034	0.028986
9737	0.014493
9601	0.014493
9508	0.014493
9237	0.014493
9034	0.014493
8824	0.014493
8661	0.014493
8047	0.014493
7976	0.014493
7775	0.014493
7717	0.014493
6458	0.014493
6241	0.014493
5710	0.014493
5184	0.014493
5050	0.014493
5036	0.014493
4655	0.014493
4537	0.014493
4490	0.014493
4314	0.014493
3811	0.014493
3012	0.014493
2708	0.014493
2571	0.014493
2531	0.014493
2386	0.014493

2276	0.014493
2208	0.014493
2160	0.014493
2020	0.014493
2013	0.014493
2012	0.014493
2003	0.014493
2001	0.014493
2000	0.014493
1999	0.014493
1998	0.014493
1711	0.014493
1374	0.014493
1298	0.014493
123456	0.043478
960105	0.014493
527873	0.014493
12345678	0.028986
123456789	0.014493
98030212097	0.014493
97071812441	0.014493
1234567891234567 89	0.014493