Resumo do Traballo de Fin de Grao

(Describa brevemente o traballo a desenvolver, xustificando o interese do mesmo, e indicando obxectivos, descrición técnica, proceso de desenvolvemento, e medios empregados. Engada tantas liñas como sexa necesario)

El término "***phishing***", es un modelo de abuso informático caracterizado por intentar adquirir información confidencial de forma fraudulenta, realizado de forma muy habitual en el servicio de correo electrónico (aunque también se puede producir mediante SMS, sistemas de mensajería instantánea, redes sociales, etc.) y que tiene un gran impacto en la seguridad de los usuarios de Internet, dado que los datos personales obtenidos de manera ilegal, serán utilizados en suplantaciones de identidad o nuevas campañas de *phishing*. Este fenómeno, se basa en el uso de estrategias de ingeniería social para solicitar las credenciales de acceso a distintos portales de los usuarios. Entre las claves habitualmente solicitadas, figuran las de la banca electrónica (que permite al atacante obtener directamente dinero) o de acceso a sistemas de correo electrónico (permitiendo al atacante disponer de cuentas de correo electrónico adicionales para lanzar nuevos ataques de *phishing*).

A pesar de la importancia del *phishing* y las grandes pérdidas relacionadas con este tipo de ataques, no existen, en la actualidad, herramientas diseñadas para su detección, marcado y eliminación de forma automática, lo cual, motiva el desarrollo de este trabajo. El objetivo, es desarrollar una herramienta capaz de realizar un análisis de contenido sobre un correo electrónico y en caso de ser necesario, marcar a este como mensaje con contenido *"phishing"* para su posterior eliminación.

En el análisis de contenido del correo electrónico a estudiar, se pretende analizar las características mas habituales de los mensajes *phishing*: (*i*) Inclusión de patrones de lenguaje demarcando la urgencia, (*ii*) petición de datos específicos como usuario y contraseña, (*iii*) no disponer del nombre del receptor y emplear, por tanto, formas no específicas para referirse a él (como por ejemplo: estimado usuario, querido cliente...), e (*iv*) inclusión de URLs cuyo host no coincide con el propio de la institución o persona que supuestamente envía el mensaje.

Para la detección de estas características, se empleará la herramienta SpamAssassin como framework de base, el cual permitirá el análisis e identificación de los contenidos de los mensajes y facilitará su integración con distintas plataformas de servicio e-mail existentes y despliegue tanto en entornos empresariales en producción como en hogares.

El proceso de desarrollo de software estará guiado por la metodología SCRUM. Esto se debe al carácter investigador que motiva este trabajo, el cuál requiere gran flexibilidad ante los cambios que puedan ser necesarios debido a los resultados obtenidos, adaptando así el desarrollo de la herramienta a las necesidades surgidas y obteniendo tras cada iteración una herramienta mas productiva y de mayor calidad.

Para desarrollar el proyecto, se empleará el editor de textos ATOM, el cual cuenta con extensiones específicas de ayuda al desarrollo de lenguaje Perl, que será el lenguaje de programación utilizado para una mejor integración con SpamAssassin, por ser el lenguaje en el que está desarrollado. La herramienta de diagramación será Visual Paradigm y la herramienta de documentación será LaTeX. El hardware empleado para el desarrollo del trabajo constará simplemente de un ordenador portátil.