Andrea Kienle et al. (Hrsg.): Die 19. Fachtagung Bildungstechnologien (DELFI),
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2021 157

Konzeption und Entwicklung eines interaktiven E-Mail-Interface für Anti-Phishing Lernspiele

Duygu Bayrak[1], René Röpke[2] und Ulrik Schroeder[2]

Abstract: Phishing-Angriffe sind eine ernstzunehmende Bedrohung, die die Internetnutzerinnen und -nutzer zur Preisgabe vertraulicher Informationen verleiten. Der Erfolg dieser Angriffe ist zum Teil auf mangelnde Aufklärung über Phishing zurückzuführen. Anti-Phishing-Lernspiele stellen eine spannende Möglichkeit dar, auf aktive und unterhaltsame Weise über Phishing-Angriffe zu lernen um sich in echten Szenarien davor zu schützen. Eine Analyse existierender Spiele zeigt jedoch erhebliche inhaltliche und methodische Schwächen in der Vermittlung und Überprüfung von Kenntnissen und Fähigkeiten. Dieser Beitrag widmet sich somit der interaktiven Vermittlung und erweiterten Wissensüberprüfung in Anti-Phishing Lernspielen zu E-Mail-Phishing und präsentiert ein interaktives E-Mail-Interface zur modularen Einbindung in Anti-Phishing Lernspiele. Dabei werden definierende Elemente von E-Mail-Clients sowie Sicherheitsindikatoren von E-Mails berücksichtigt, um eine authentische Simulation von E-Mail-Phishing zu realisieren.

Keywords: Lernspiel; Phishing; E-Mail; Interface; Spielbasiertes Lernen

1 Motivation und verwandte Arbeiten

Phishing stellt für Nutzerinnen und Nutzer des Internets eine aktive Bedrohung dar. Allein in 2020 gab es mehr als doppelt so viele Phishing-Angriffe wie in den Jahren zuvor [An21]. Da technische Lösungsansätze allein das Problem nicht lösen, ist die Aufklärung von Nutzerinnen und Nutzern eine aktiv verfolgte, komplementäre Maßnahme. Eine Art der Vermittlung basiert auf Anti-Phishing Lernspielen. Eine Analyse existierender Spiele und verwandter Literatur zeigt inhaltliche und methodische Schwächen in der Vermittlung und Überprüfung von Kenntnissen und Fähigkeiten [Ro20].

Während zahlreiche Spiele zu Themen wie Phishing-URLs existieren [Ro20], gibt es nur wenige Spiele zu E-Mail-Phishing und anderen Phishing-Techniken. Zwei existierende Spiele sind What.Hack [We19] und Bird's Life [WJZ17]. Zur Vermittlung und Erprobung der Erkennungsstrategien für E-Mail-Phishing sind Spielinhalte unterschiedlich aufbereitet. So präsentiert Bird's Life lediglich Screenshots von E-Mails und fordert die Spielerinnen und Spieler zu entscheiden, ob es sich um eine Phishing-E-Mail handelt oder nicht. In What.Hack wird ein E-Mail-Client simuliert, um den Umgang mit Phishing-E-Mail authentisch und realitätsnah abzubilden.

[1] RWTH Aachen, Templergraben 55, 52062 Aachen, duygu.bayrak@rwth-aachen.de
[2] RWTH Aachen, Lehr- und Forschungsgebiet Informatik 9, Ahornstraße 55, 52074 Aachen, {roepke, schroeder}@informatik.rwth-aachen.de

Probleme beider Spiele [We19, WJZ17] sind die fehlende Interaktivität und limitierte Umsetzung der Wissensüberprüfung. Während zwar ein E-Mail-Client simuliert oder ein Screenshot einer echten E-Mail angezeigt wird, müssen Spielerinnen und Spieler jedoch lediglich eine Ja/Nein-Entscheidung treffen, um im Spiel voranzuschreiten. Die Interaktionsmöglichkeiten mit gegebenen E-Mails sind eingeschränkt und es ist nicht erforderlich verdächtige Merkmale einer Phishing-E-Mail zu identifizieren. Die Ratewahrscheinlichkeit ist entsprechend hoch und es kann zu keinem Zeitpunkt im Spiel sichergestellt werden, ob die die vermittelten Merkmale von Phishing-E-Mails erkannt und zur Entscheidung zu Rate gezogen werden. Insbesondere im Kontext digitaler Lernspiele stellen diese fehlenden Aspekte in existierenden Arbeiten eine ungenutzte Chance dar. Digitale Lernspiele bieten eine geeignete Lernumgebung, um interaktive Vermittlung- und Erprobungselemente in den Lernprozess zu integrieren und erweiterte Wissensüberprüfung zur besseren Erfassung der Lernleistung zu nutzen.

Dieser Beitrag widmet sich somit der interaktiven Vermittlung und erweiterten Wissensüberprüfung in Anti-Phishing Lernspielen zu E-Mail-Phishing und präsentiert ein interaktives E-Mail-Interface zur modularen Einbindung in Anti-Phishing Lernspiele.

2 Konzeption

Das Konzept des interaktiven E-Mail-Interface basiert auf den Gestaltungselementen und Funktionalitäten gängiger E-Mail-Clients. Hierfür wurden die weltweit beliebtesten und weitverbreitetsten E-Mail-Clients (Desktop- oder Webanwendungen) betrachtet und definierende Elemente identifiziert. Dazu zählen unter anderem Gmail, Microsoft Outlook, Apple Mail, Yahoo! Mail und Thunderbird von Mozilla. Die Simulation eines E-Mails-Clients in Anti-Phishing-Lernspielen soll eine möglichst authentische und realistische Lernumgebung schaffen und eine Übertragung erlernter Kenntnisse und Fähigkeiten unterstützen. Der Aufbau des Interfaces kann in vier Bereiche unterteilt werden: Menü, Ordnerübersicht, Nachrichtenliste und Vorschau (siehe Abb. 1). Im Folgenden wird auf die Interaktionsmöglichkeiten, die Wissensüberprüfung und den Einsatz von Feedback im E-Mail-Interface eingegangen.

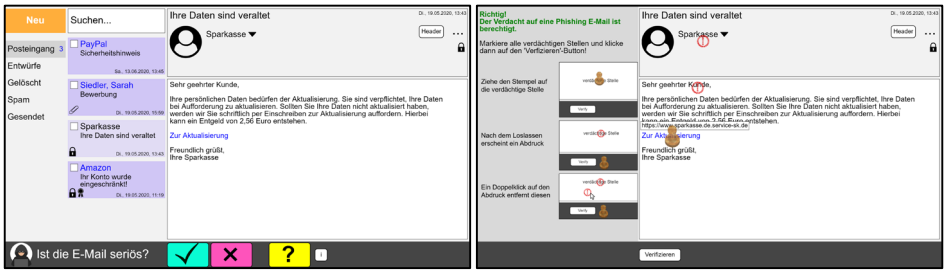

Abbildung 1. Postfachansicht (links) und Markierungsansicht mit platzierten Markierungen (rechts).

2.1 Interaktionsmöglichkeiten

Neben der Gestaltung sind die Interaktionsmöglichkeiten im E-Mail-Interface ebenfalls an gängigen E-Mail-Clients orientiert. Zur späteren Einbindung in Anti-Phishing-Lernspiele ist insbesondere die Interaktion mit den präsentierten E-Mails relevant. Im Fokus stehen hierfür zwei Interaktionsmöglichkeiten:

- **Interaktionsmöglichkeit mit Links und URLs:** Analog zu echten E-Mail-Clients soll es möglich sein, in E-Mails eingebundene Links und URLs mittels Mouse-Over anzuzeigen. Durch die mögliche Verschleierung der echten URL hinter einem Link mittels HMTL oder Rich-Text-Formatierung, können bösartige URLs in Phishing-E-Mails versteckt werden. Damit Nutzerinnen und Nutzer lernen können, Links zu überprüfen und Phishing-URLs zu erkennen, muss die Interaktion analog zu echten E-Mail-Clients realisiert werden.
- **Interaktionsmöglichkeit mit E-Mail-Header-Informationen**: Eine weitere Funktionalität in echten E-Mail-Clients ersetzt die Absender-E-Mailadresse durch den Vor- und Nachnamen, um Nutzerinnen und Nutzer in einem lesbaren, leicht verständlichen Format anzuzeigen, von wem die E-Mail stammt. In Phishing-E-Mails wird diese Funktion jedoch ausgenutzt, um eine bösartige Absender-E-Mailadresse zu verschleiern und stattdessen einen vertrauenswürdigen Namen anzuzeigen. Um die Verschleierung von Absender-E-Mailadresse und weiteren Headerinformationen zu erkennen, muss es Nutzerinnen und Nutzern möglich sein, die notwendigen Informationen einer E-Mail einzusehen und zu überprüfen.

2.2 Wissensüberprüfung

Andere Anti-Phishing-Lernspiele, wie What.Hack [We19], bieten zwar eine Interaktion mit den angezeigten E-Mails an, jedoch mangelt es an Interaktionsmöglichkeiten bei der nachfolgenden Wissensüberprüfung. Oft werden nur Single-Choice-Fragen verwendet, um den Spieler entscheiden zu lassen, ob es sich bei der angezeigten E-Mail um Phishing handeln könnte oder nicht. Da diese Single-Choice-Frage nur nach dem „Was“ fragt, also nach der Klassifizierung der E-Mail, ist es sinnvoll, diese mit einer weiteren Frage nach dem „Warum“ zu kombinieren. Es sind somit zwei Interaktionsmöglichkeiten zur Wissensüberprüfung angedacht: Single-Choice Fragen zur Entscheidung, ob eine E-Mail legitim oder für Phishing verdächtig ist und eine Markierungsfunktionalität verdächtiger E-Mailelemente und -inhalte (siehe Abb. 1).

Für die Markierfunktionalität soll analog zum „Abstempeln“, also etwas kennzeichnen, eine Drag-&-Drop-Interaktion verwendet werden, bei welcher ein Stempel auf die verdächtigen Elemente der E-Mail gezogen und abgelegt wird. Der Stempelabdruck markiert hierbei die Aspekte zur Beantwortung der „Warum“-Frage. Die als verdächtig markierten Stellen geben die Gründe für die Klassifikationsentscheidung an.

2.3 Feedback

Nach Abschluss der Klassifikation einer E-Mail und Markierung verdächtiger Stellen soll elaboriertes Feedback gegeben werden. Hierbei soll zum einen Feedback zur Klassifikationsentscheidung gegeben werden, zum anderen soll Feedback zu den markierten Stellen präsentiert werden. Hierbei sind insbesondere zwei Fälle interessant:

- **Fall 1:** Bei einer **falschen Klassifikation als legitim** (false negative) werden alle verdächtigen Stellen in der E-Mail farblich hervorgehoben, sodass direkt erkennbar ist, welche Stellen den Verdacht auf Phishing begründen. Die Begründung, warum eine Stelle auf Phishing hindeutet, erscheint erst bei einem Mouseover über die farblich hervorgehobene Stelle und nur wenn alle Begründungen durch ein Mouseover angezeigt wurden, erscheint ein Button, um die nächste E-Mail analysieren zu können. Dadurch wird sichergestellt, dass alle Begründungen betrachtet werden, auch wenn nicht garantiert werden kann, dass diese gelesen werden.
- **Fall 2:** Bei einer **korrekten Klassifikation als Phishing** (true positive) und nachfolgender Markierung werden die Textstellen in unterschiedlichen Farben hervorgehoben, je nachdem ob es sich um eine korrekte oder fehlende Markierung handelt. Außerdem werden die Abdrücke, die von dem Stempel erzeugt wurden, durch ein Häkchen- bzw. ein Kreuz-Symbol ersetzt. Dadurch ist in der Auswertung erkennbar, welche Stellen zuvor markiert wurden, auch wenn diese nicht als verdächtig einzustufen sind. Zusätzlich zu der farblichen Hervorhebung wird Rückmeldung über die Anzahl der korrekt markierten Stellen im Verhältnis zu der Anzahl verdächtiger Stellen in der E-Mail angezeigt (z. B. 3 von 5 Stellen). Für alle korrekt markierten Stellen werden die jeweiligen Begründungen in einer Listenansicht gesammelt. Analog zu Fall 1, werden nicht erkannte Stellen hervorgehoben und durch ein Mouseover die Begründung angezeigt, bevor mit der Klassifikation weiterer E-Mails weitergemacht werden kann.

Die verbleibenden Fälle umfassen die Klassifikation legitimer E-Mails als legitim (true negative) oder verdächtig (false positive). In beiden Fällen wird entsprechendes Feedback ohne weiteren Interaktionsbedarf gegeben und es kann mit der Analyse der nächsten E-Mail weitergemacht werden.

3 Implementierung

Das Interface wurde mithilfe des Multi-Touch-Learning-Game-Frameworks (MTLG), welches von der RWTH Aachen entwickelt wurde, implementiert. Es ermöglicht die Erstellung von webbasierten Lernspielen, basierend auf HTML5 Canvas-Element und nativem JavaScript. Außerdem bietet es eine Ablaufsteuerung für Levelstrukturen und Analyse des Nutzerverhaltens mittels Learning Analytics.

Das Interface wurde am Revealing Module Pattern[3] orientiert modular entwickelt, um flexibel in spätere Lernspiele integriert zu werden. Dabei wird es an das global sichtbare Singleton *MTLG* angebunden, worüber nachher im Spiel auf die Modulfunktionen zugegriffen werden kann. Diese Funktionen ermöglichen unter anderem die Initialisierung des Interface und das Laden von E-Mails in einer aufbereiteten JSON Struktur.

3.1 Verwendung echter E-Mails

Neben der Gestaltung des E-Mail-Interface sollen auch die zu anzuzeigenden E-Mails realitätsnah und authentisch sein. Dazu wird der Import von echten E-Mails im eml-Format[4] unterstützt, welche in einem Pre-Processing-Schritt in JSON-Objekte überführt werden. Die Informationen aus dem JSON-Objekt werden für die Anzeige unterschiedlicher Teile der E-Mail verwendet. So werden das Absendedatum, der Betreff sowie die Absender- und Empfängeradressen in der Kopfzeile der E-Mail angezeigt. Die Headerinformationen werden dagegen erst auf Wunsch der Nutzerinnen und Nutzer angezeigt. Für die Nachrichtenanzeige wird die Text- oder HTML-Version der E-Mail verwendet. Zuletzt werden Anhänge mit ihrem Namen und einem Dateityp-Icon dargestellt.

3.2 Sicherheitsindikatoren

Anhand der Informationen des JSON-Objekts kann die E-Mail im Interface dargestellt werden, doch um die E-Mail klassifizieren und gegebenenfalls die verdächtigen Stellen markieren zu können, sind zusätzliche Schlüsselinformationen notwendig. Dazu zählen die Klassifizierung der E-Mail (z. B. legitim oder Phishing), das Vorhandensein einer S/MIME Verschlüsselung bzw. Signatur und Hinweise zur Anzeige im Spiel. Falls es sich um eine legitime E-Mail handelt, wird zusätzlich noch die Begründung dafür benötigt, welche dann bei dem Feedback angezeigt wird. Falls es sich um eine Phishing E-Mail handelt, werden Informationen über die verdächtigen Stellen benötigt, die im Interface markiert werden sollen. Für alle zu markierenden Stellen muss ein Grund als Textbaustein ergänzt werden, um diese später im Interface anzeigen zu können. Diese Informationen werden in einem weiteren, aktuell manuellen Pre-Processing-Schritt dem JSON-Objekt hinzugefügt. Die Automatisierung stellt eine zukünftige Herausforderung dar.

4 Fazit und Ausblick

Im Rahmen dieses Beitrags wurde ein interaktives E-Mail-Interface präsentiert, welches auf Basis einer umfänglichen Analyse existierender Lernspiele zu E-Mail-Phishing entwickelt wurde. Dabei standen die interaktive Vermittlung und erweiterte Wissensüberprüfung im Fokus, um identifizierte Schwächen verwandter Arbeiten [Ro20] zu

[3] https://addyosmani.com/resources/essentialjsdesignpatterns/book/, zuletzt besucht am 17.06.2021
[4] Spezifiziert in RFC-5322, https://datatracker.ietf.org/doc/html/rfc5322, zuletzt besucht am 17.06.2021

vermeiden. Zur Analyse und Klassifikation möglicher Phishing-E-Mails wurden verschiedene Interaktionsmöglichkeiten, wie z.B. das Mouseover-Event über URLs, die Analyse des E-Mailheaders und die Überprüfung der Absenderinformationen implementiert. Die erweiterte Wissensüberprüfung umfasste neben der bereits etablierten Single-Choice-Fragestellung zur Klassifizierung der E-Mail auch eine Markierungsfunktionalität, um verdächtige Stellen der E-Mail zu markieren. Anders als in vorangegangen Arbeiten kann dadurch explizit erfasst werden, ob Nutzerinnen und Nutzer die richtigen Merkmale bei der Erkennung von Phishing-E-Mails identifizieren. Zusätzlich wurde elaboriertes Feedback in den unterschiedlichen Klassifikationsfällen implementiert. Die Implementation erfolgte auf Basis des MTLG-Frameworks und unter Verwendung echter E-Mails, welche in einem zweistufigen Pre-Processing in eine erweiterte JSON-Struktur überführt und um Schlüsselinformationen erweitert wurden.

Das entwickelte interaktive E-Mail-Interface wurde im Rahmen einer ersten Nutzerstudie mit einer Gruppe von Nutzerinnen und Nutzern (n=5) in den Aspekten der Usability und Funktionalität evaluiert. Rückmeldungen und Feedback aus der Studie konnten so unmittelbar in den Entwicklungsprozess integriert und zur Verbesserung des Interfaces verwendet werden. So wurde beispielsweise die Bedeutung der Klassifizierungsbuttons missverstanden, welche die Nichtausführung der eigentlichen Aufgabe bedingt hat. Im jetzigen Interface ist die Aufgabe nun als Frage formuliert, sodass eine Antwort die Benutzung der Buttons voraussetzt. Eine vollumfängliche Studie zum Einsatz des interaktiven E-Mail-Interface im Rahmen eines digitalen Lernspiels soll im nächsten Schritt erfolgen. Auch die Bewertung der Realitätsnähe und Authentizität steht noch aus.

Zusammenfassend wurde in diesem Beitrag ein interaktives E-Mail-Interface für digitalen Anti-Phishing Lernspiele präsentiert. Durch die Integration echter E-Mails und Simulation gängiger Funktionen beliebter E-Mail-Clients birgt das Interface ein großes Potential für Personalisierung. Die Integration in ein sich aktuell in Entwicklung befindendes personalisierbares Lernspiel und die umfassende Evaluation bilden die nächsten Schritte.

Literaturverzeichnis

[An21] Anti-Phishing Working Group: Phishing Activity Trends Report 4th Quarter 2020. https://docs.apwg.org/reports/apwg_trends_report_q4_2020.pdf, 01.04.2021

[Ro20] Roepke, R. et al.: A Pond Full of Phishing Games - Analysis of Learning Games for Anti-Phishing Education. In Model-Driven Simulation and Training Environments for Cybersecurity. Springer, Cham, 2020. https://doi.org/10.1007/978-3-030-62433-0_3.

[We19] Wen, Z. A. et al.: What.Hack: Engaging Anti-Phishing Training Through a Role-Playing Phishing Simulation Game. In Proceedings of the Conference on Human Factors in Computing Systems. CHI ’19. ACM, New York, USA, 2019. https://doi.org/10.1145/3290605.3300338 .

[WJZ17] Weanquoi, P.; Johnson, J.; Zhang, J.: Using a Game to Teach About Phishing. In Proceedings of the 18th Annual Conference on Information Technology Education, 75. SIGITE ’17. ACM, New York, USA, 2017. https://doi.org/10.1145/3125659.3125669.