JSiSE Research Report
vol.35,no.1(2020-5)

適応的フィッシングメール判断トレーニング課題出題手法の評価

程 秋涛[*1]，長谷川 忍[*2,1]，太田 光一[*2,1]
[*1] 北陸先端科学技術大学院大学 先端科学技術研究科
[*2] 北陸先端科学技術大学院大学 情報社会基盤研究センター

Evaluation for Adaptive Training Generation in Discrimination of Phishing Emails

Qiutao Cheng[*1], Shinobu Hasegawa[*2,1] , and Koichi Ota[*2,1]
[*1] Graduate School of Advanced Science and Technology, JAIST
[*2] Research Center for Advanced Computing Infrastructure, JAIST

The purpose of this research is to verify the effectiveness of our adaptive training generation algorithm to improve the discrimination skill of phishing emails. We conducted a comparative experiment between random and adaptive training conditions with 20 participants. The results showed that their skill and awareness to identify the type of emails were improved in both conditions.

キーワード: セキュリティ教育，フィッシングメール判別トレーニング，適応的課題生成，評価

1. はじめに

ベライゾンの「2019 年度データ漏洩/侵害調査報告書」によると，データ漏洩/侵害における攻撃のうち最も多いものが，送信者を偽装するメールにより，個人情報の窃取やシステムの破壊を行うフィッシングメールによるものであった(1)．フィッシングメールの内容は年々巧妙化しており，新たに起こる全ての種類の攻撃を技術的に防ぐことは困難である．

こうした背景に基づいて，我々はセキュリティに関する意思決定スキルの向上の中でも特に，フィッシングメール，スパムメール，通常のメールの判別スキルを対象として，(a) トレーニング課題の難易度の定義，(b) 適応的なトレーニング課題の出題，から構成されるトレーニング環境を提案している(2). 本稿では，提案手法の概要について説明するとともに，その有用性を調べるために行ったケーススタディの方法および結果，考察について述べる．

2. 適応的課題生成手法の概要

本稿では，我々が提案している適応的課題生成手法の概要について述べる．詳細については先行研究を参考にされたい(2)．

2.1 トレーニング項目の設定

本研究では，メール判別スキルに対するトレーニングの対象として，メールの(1) 件名，(2) アドレス，(3) 送信者，(4) 送信日時，(5) 添付ファイル，(6) ハイパーリンク，(7) 本文，をそれぞれトレーニング項目として採用することとした．

2.2 トレーニング課題の難易度の設定

トレーニング課題の難易度を定義するために，まず実際の(F) フィッシングメール 96 通，(S) スパムメール 54 通，(U) 通常のメール 51 通を収集した．なお，収集したメールの割合は一般に送信されるメールの割合とは異なっているが，フィッシングメール判別のト

レーニングを行う観点から，フィッシングメールの割合を多く収集した．

次に，収集したメールについて各トレーニング項目に対して表 1 に示すような典型的なラベルを人手で付与した．

表 1 件名に関するラベルとその分布

ラベル	F	S	U
業務に関連する件名	61 通	2 通	41 通
業務に関連しない件名	8 通	37 通	9 通
即座の行動が必要な件名	11 通	2 通	0 通
件名なし	16 通	0 通	1 通

トレーニング課題の難易度は以下のように定義する．トレーニング課題 X がフィッシングメールである条件付確率を P(a|X)，スパムメールである条件付確率を P(b|X), 通常のメールである条件付確率を P(c|X)とそれぞれしたとき, (1)式の D の値が大きいほどトレーニング課題として難しいと定義する．

$$D = \frac{1}{2}(1 - (|P(a|X) - P(b|X)| + |P(a|X) - P(c|X)| + |P(b|X) - p(c|X)|) \quad \ldots (1)$$

なお，$P(a|X)$，$P(b|X)$，$P(c|X)$はそれぞれ(2)式のナイーブベイズより求める．ここで，$X_1, \ldots, X_n$はメールの各項目におけるラベルであり，Yはメールの種類（フィッシングメール，スパムメール，通常のメール）である．

$$P(Y|X_1, \ldots, X_n) = \frac{P(Y)P(X_1, \ldots, X_n|Y)}{P(X_1, \ldots, X_n)}$$

ただし，$Y = \{a, b, c\}$. …(2)

これにより，新たなフィッシングメールやスパムメールの事例を収集すれば，そのトレンドを反映した難易度の設定を行うことが可能となる．

2.3 適応的トレーニング課題出題

プレテストとして，全ての項目についてレベル 2 に対応するトレーニング課題を出題し，正解した場合はより難易度の高い問題，不正解の場合はより難易度の低い問題から開始し，同じ問題を出さないように 5 問セットで出題する．

トレーニング中は，学習者がトレーニング課題を解答する毎にシステムが正誤判定を行い，不正解の場合には学習者に注目すべき項目をフィードバックする．

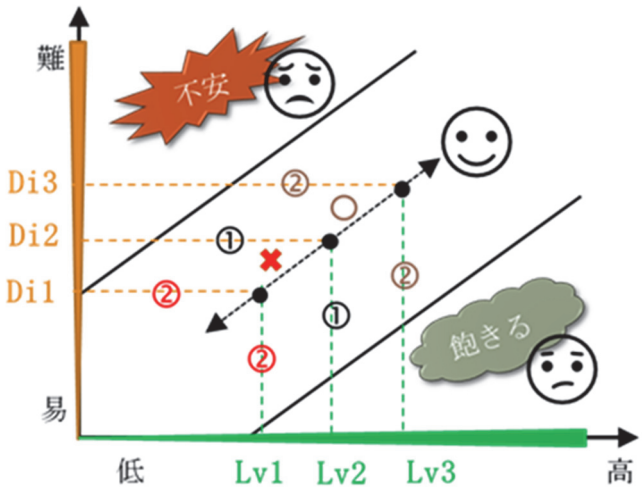

図 1. 初期レベルの設定

3. ケーススタディ

提案した適応的トレーニング課題出題手法の有効性を確認するために，2019 年 12 月にケーススタディを行った．被験者は情報セキュリティを専門としておらず，メール判別トレーニングを体験したことのない大学院生 20 名（日本語話者 6 名，日本語を読み書きできる留学生 14 名）を対象とした．

3.1 ケーススタディ用トレーニング課題の作成

ケーススタディ用に 51 通のフィッシングメール, 45 通のスパムメール，42 通の通常メールをトレーニング課題として作成した．トレーニング課題の難易度とプレテスト・ポストテストにおける配点の分布は表 2 の通りである．

表 2. トレーニング課題の難易度分布

レベル	Lv.1	Lv.2	Lv.3	Lv.4	Lv.5
配点	1 点	2 点	3 点	4 点	5 点
難易度	<0.025	< 01.4	< 0.36	< 0.55	< 1
問題数	32	35	29	22	20

3.2 方法

ケーススタディは，提案手法による課題出題とランダム順の課題出題の差を比較することを目的として，被験者内計画で実施した．具体的には，2 つのグループにそれぞれランダムに 10 名ずつ割り当てた．グルー

プ A はランダム出題→適応的出題，グループ B は適応的出題→ランダム出題の順にトレーニングを行うことでカウンターバランスをとった．実際の手順は以下のとおりである．

1. プレテスト（計 15 点）
   被験者に Lv.1 から Lv.5 の課題 1 問ずつからなるプレテストを実施した．また，被験者の基礎的情報や，日常生活におけるフィッシングメールおよびスパムメールについての注意の程度についてアンケート調査を行った．
2. 一回目のトレーニング（30 分間）
   ランダム出題条件（グループ A）：被験者に 5 問のランダムな難易度のトレーニング課題を 5 問提示し，回答および答え合わせを行う活動を一セットとし，30 分間繰り返した．
   適応的出題条件（グループ B）：学習者のレベルに応じて 5 問の課題を提示し，回答および答え合わせを行う活動を一セットとし，30 分間繰り返した．
3. 中間テスト（計 15 点）
   一回目のトレーニング完了後に中間テストを実施した．これにより，一回目のトレーニングを経た被験者のメール判別能力について確認した．中間テストに用いる課題は，プレテストと同様に Lv.1 から Lv.5 の課題をそれぞれ 1 問ずつで構成した．また，フィッシングメールやスパムメールへの注意の程度についてもアンケート調査を行った．
4. 二回目のトレーニング（30 分）
   二回目のトレーニングは，一回目のトレーニングから 1 日以上空けて実施した．グループ A は適応的出題条件，グループ B はランダム出題条件でそれぞれ一回目の出題方法と同じ形式で行った．
5. ポストテスト（計 15 点）
   二回目のトレーニングを行った後にポストテストを実施した．これにより，二回目のトレーニングを経た被験者のメール判別能力について確認した．ポストテストに用いる課題は，プレテストと同様に Lv.1 から Lv.5 の課題をそれぞれ 1 問ずつで構成した．また，フィッシングメールやスパムメールへの注意の程度に加えて，各項目の視点から被験者のメール判別能力が向上したかについてアンケート調査を行った．

3.3 結果

表 3，表 4 にプレテスト，中間テスト，ポストテスト（それぞれ最高点は 15 点）の結果をまとめたものを示す．なお被験者 ID が A から始まるものがグループ A，B から始まるものがグループ B である．

表 3. テスト結果

被験者 ID	プレ得点	中間得点	ポスト得点
A1	4	13	13
A2	9	13	15
A3	13	13	15
A4	7	9	13
A5	5	8	13
A6	6	9	15
A7	4	13	15
A8	6	12	12
A9	6	13	15
A10	9	10	15
B11	3	13	15
B12	8	13	15
B13	9	10	15
B14	3	15	15
B15	6	13	15
B16	3	12	15
B17	11	15	15
B18	9	13	15
B19	9	8	15
B20	6	9	13

表 4. テスト結果の統計量

条件	差の平均	差の分散	t 検定
ランダム	3.6	2.8	t = -.436
適応的	4.1	3.4	p = .668

ランダム条件および適応的条件のトレーニング効果を比較するために各条件の前後のテストの点数の差分に対して，Shapiro-Wilk による正規性の検定を行ったところ，どちらの条件も正規性が確認できたため，対応のある t 検定を行った．その結果，双方の結果には有意差が見られなかった．

トレーニングを通じてメールの判断能力が向上した実感とトレーニングを通じてメール判別能力を向上させる目的が達成できたかどうかについて質問したアンケート結果を図 2 に示す．これらより，いずれの手法においても 8 割以上の被験者がトレーニングを通じてメール判別の能力の向上や目標を達成できたことがわかる．

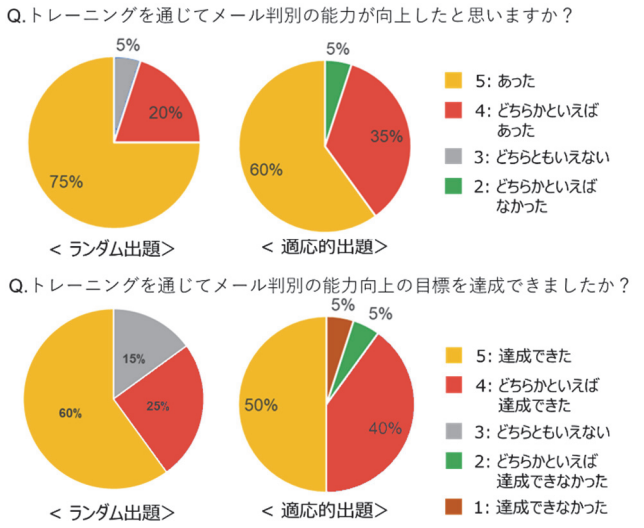

図 2. アンケート結果

これらのことから，今回のケーススタディではトレーニングの方式に関わらず，メール判別のトレーニングにより被験者のメール判別能力が向上したことが観測された．トレーニング前後のテストの点数の差は両条件で有意な差は見られなかった．

今回のケーススタディでは，30 分間のメール判別トレーニングを行ったが，1 回目のトレーニング後の中間テストでどちらの条件でもかなり高得点となり，2 回目のトレーニング後には 7 割以上の被験者が満点となった．これらのことから，トレーニング時間が長かったことやトレーニング課題が簡単すぎたことなどが考えられる．

4. おわりに

本稿では，情報セキュリティ，特にフィッシングメールの判別に関する意思決定スキルのトレーニングを対象として，学習者の理解状況に応じたトレーニング課題を出題することで，集中的な学習を実現するトレーニング課題生成手法について紹介した．提案手法は柔軟性があり，今後，新しい攻撃手法に対するデータセットが構築できれば，トレーニング内容を容易に更新することができる．

さらに，トレーニング課題生成手法の有効性に関するケーススタディを通じて，学習者が経験や失敗を繰り返すことがスキル改善にどのように役立つかについて調査した．その結果，トレーニングのタイプに関わらず学習者のメール判別能力が向上したことがわかった．

今後の課題としては，本来はトレーニング項目のうち学習者の感度が低い（学習者が気付きにくい）項目を重点的にトレーニングすることを目指していたが，課題作成時に特定の項目のみをトレーニングする課題を設定することが難しく，今回のケーススタディではトレーニング課題の全体の難易度により課題を設定した．今後は，この点について改善し，より効果的な課題提示手法を確立したい．

また，今回提案した出題手法では，学習者がトレーニング課題を理解していながら不正解であった場合やトレーニング課題を理解せずに正解する場合を考慮していない．今後，この問題への改善対策として，Corbett らが提案した Bayesian Knowledge Tracing（BKT）学習者モデルの利用を検討している[3]．

謝辞

本研究の一部は，JSPS 科研費基盤研究(B) (No. 17H01992)，基盤研究(C) (No. 17K00479)の助成による．

参 考 文 献

(1) Verizon: “2019 年度データ漏洩/侵害調査報告書”, https://enterprise.verizon.com/ja-jp/resources/ reports/dbir/ (2020 年 3 月 25 日確認)

(2) CHENG Qiutao，長谷川 忍 : “意思決定スキル向上のためのトレーニング課題生成”，教育システム情報学会研究報告 Vol.32, No.4, pp1-8 (2019)

(3) Corbett, Albert T., and John R. Anderson. "Knowledge tracing: Modeling the acquisition of procedural knowledge." User modeling and user-adapted interaction 4(4), pp.253-278(1994).