

# セキュリティにおける人的ミス防止のための悪性メール判断支援方式の評価

岡崎 香織 【 コミュニケーション & コラボレーション研究室 】

## Evaluation of a Malicious Email Judgment Support Method for Preventing Human Errors in Security

OKAZAKI, Kaori 【 Communication and Collaboration Lab. 】

### 1 はじめに

近年、悪性メールは巧妙化が進み、流通量も増加傾向にある [1]。このような状況において、ユーザ自身がメールの真偽を適切に判断できるようになることは重要である。これまで、悪性メールをはじめとしたセキュリティインシデント対策としてリテラシー教育や啓発活動が重視されてきた [2, 3]。しかし、十分なセキュリティ知識を有していても、必ずしも正しく判断が行えるとは限らない。

本研究では、悪性メール判断における人間の行動や認知に着目し、知識・メタ認知・判断根拠という3つの観点から、正しい判断がどのような過程で行われているのかを段階的に明らかにすることを目的とする。

具体的には、まずセキュリティ知識と悪性メール判断行動との関係を調査する。次に、教育介入を通じて、認知を含む判断支援方式の効果およびその定着性を検証する。さらに、正誤判断の差を生み出す判断根拠の特徴を分析することで、より効果的な判断支援手法や教育手法の設計に向けた示唆を得ることを目指す。

### 2 セキュリティ知識と悪性メール判断行動の関係分析

本章では、セキュリティ関連用語の理解度と悪意判断行動との関係を分析した。相関係数および人数分布の分析結果から、用語知識が高い実験参加者ほど、悪意があると判断する傾向は見られたものの、その相関はいずれも弱く、知識の有無のみで判断結果を一貫して説明することは困難であることが示された。特にリンク判断においては、用語知識区分の高低にかかわらず判断結果が分散しており、知識として理解している内容が、実際の行動判断へ直結していない可能性が示唆された。

また、判断材料の分析から、実験参加者は送信元アドレスや文章、内容など複数の要素を判断材料として用いていることが確認された。一方で、URLやメールアドレスを確認している実験参加者であっても、判断結果が必ずしも悪意方向に収束するわけではなかった。このことから、確認行動が形式的に行われている可能性や、細部まで十分に確認できていない可能性が示唆された。

### 3 メール判断教育の効果定着分析

本章では、悪性メール判断に対する教育介入が判断精度に与える影響と、その効果が時間経過後に定着するかを検討した。

#### 3.1 実験概要

本分析では、大学生13名のうち、2回目調査まで参加した6名のデータを用いた。実験は自作Webページ上で実施し、1回目調査では、実験前後のアンケートと、事前・教育・事後の3段階に分けた計5ラウンドのメール判断実験を実施した。2回目調査では、一定期間後に同様の手順で、メール判断実験を2ラウンド実施した。

教育方式は、事前に判断ポイントを文章で提示する従来方式、教育ラウンドの判断中にヒントを提示し、ラウンド進行に伴い提示量を減少させる提案方式1、判断時とフィードバック時に自信度を含めることでメタ認知を促す提案方式2の3条件とした。提示メールはECサイト、銀行等を題材とした計28通を用い、メール種別の重複を避けた。

#### 3.2 メール判断実験結果・考察

全方式を総合した正解率を表1に示す。事後調査の結果から、教育直後であっても判断精度が維持されない可能性が示された。一方、一定期間後に実施した2回目調査では正解率が回復する傾向が見られ、教育の影響が即時的に表れるというよりも、時間経過を伴って表出する可能性が示唆された。図1に示す教育方式別に見ても全体としては同様の傾向を示したが、特に提案方式2(meta)では回復率が高くなった。

#### 3.3 アンケート結果・考察

1回目実験後では、普段のメールへの注意意識やフィッシングへの警戒意識、セキュリティ重要性の認識、判断能力向上意欲が高まる傾向が確認された一方、判断精度の即時的改善とは一致しなかった。2回目実験前後では、前回の内容や判断ポイントの記憶、判断根拠の説明可能性に関する回答が得られ、教育内容が記憶として保持され、判断時に想起されることで行動として再利用される可能性が示された。

段階	表1 判断段階正解率			
	事前	教育	事後	2回目
正解率	0.875	0.653	0.417	0.667

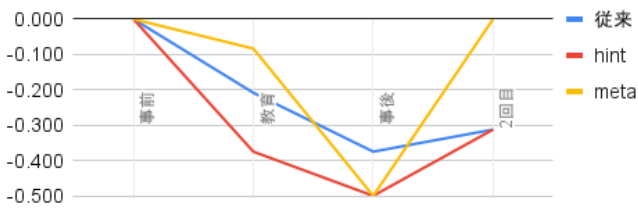


図1 教育方式別の判断段階正解率

## 4 悪性メール判断に用いられる判断根拠の特徴分析

### 4.1 実験概要

本章では、悪性メール判断においてユーザが用いる判断根拠の特徴を明らかにすることを目的として、メール判断課題に加えて、性格・行動特性アンケートおよびヒアリングを実施した。実験参加者は学生12名、社会人5名の計17名とし、用意したPC上のGmailを用いて6通のメール（悪性4通、正しい2通）を確認させ、各メールについて怪しいか判断させた。判断中は注目箇所をマウスカーソルで示させ、画面録画により記録した。

### 4.2 正誤判断の要因分析

悪性メール判断においては、判断根拠としてメールアドレスが最も多く挙げられ、それ以外の判断根拠にはばらつきが見られた。比較的多く挙げられた要素としては、URL、個人情報への欠如、リンク誘導、焦らせる文言、日本語の違和感などが確認された。正しいメールに対しては、URL等に加え、「怪しい点が見当たらない」こと自体を判断根拠とする回答も見られた。

誤判断の要因としては、心当たりや身に覚えがあるといった主観的判断や、怪しく感じない、怪しい点がないといった消極的判断が多く確認された。これに対し、全問正解者は複数の判断根拠を併用し、日本語表現やメールアドレスなどを段階的に確認する傾向があった。

また、性格特性との対応付けから、せっかち・判断切り上げ傾向が高い実験参加者は件名や冒頭文、焦らせる文言など早期に確認可能な情報を根拠としやすく、確認行動・精読傾向が高い実験参加者は複数要素を組み合わせて判断する傾向が示唆された。

## 5 議論

本章では、これまでに得られた結果を踏まえ、悪性メール判断に関する示唆について議論する。

まず、セキュリティ知識が高い実験参加者ほど悪意があると判断する傾向は見られたものの、相関はいずれも弱い結果となった。知識や確認行動そのものではなく、どのように解釈し、判断に結びつけるかというプロセス

が重要であり、知識付与や行動促進のみが十分な判断支援とはならないことが示唆された。

次に、教育介入の効果定着分析の結果から、本教育は即時的効果よりも、注意意識向上など前段階を形成し、時間経過後に判断行動として表出する可能性があると考えられる。特に、メタ認知促進方式では回復傾向が比較的大きく、支援による振り返りの重要性が示唆された。また、すべての教育方式において、事前から事後にかけて正解率が低下し、再測定で向上する傾向が見られたことから、判断精度の変動にはメール難度の影響が含まれている可能性も考えられる。加えて、参加者数が限られている点には留意が必要である。

さらに、判断根拠の特徴分析から、単一の判断根拠への依存や消極的判断は誤判断につながる可能性が示された。一方、全問正解者は複数の判断根拠を併用する傾向があり、複数の観点からの確認が正答につながりやすいことが示された。

## 6 まとめ

本研究では、悪性メール判断における人間の行動や認知に着目し、知識・メタ認知・判断根拠の3観点から、正しい判断がどのような過程で行われるかを段階的に検討した。

その結果、セキュリティ知識の有無のみでは判断結果を十分に説明できず、確認行動も形式的に行われる場合があることが示された。また、教育介入は判断精度を即時的に向上させるとは限らないものの、注意意識や振り返りを形成し、時間経過後に判断行動として表出する可能性が示唆された。さらに、判断根拠の分析から、単一根拠への依存や消極的判断は誤判断を招く可能性があり、複数の観点から確認を行う行動が正答につながりやすいことが示された。

以上より、悪性メール判断支援には、判断根拠の偏りを補正し、複数視点の確認とメタ認知的振り返りを促す設計が重要であることが示された。

## 参考文献

- [1] 内閣府政府広報室. 巧妙化するフィッシングから身を守るには. <https://www.gov-online.go.jp/tokusyuu/phishing/>, 2023. 参照日: 2026-01-26.
- [2] 経済産業省. サイバーセキュリティ人材の育成促進に向けた検討会 最終取りまとめ. <https://www.meti.go.jp/press/2025/05/20250514002/20250514002-2.pdf>, 2025. 参照日: 2026-01-26.
- [3] 藤原, 晴, 敷田, 幹文. 悪性メール提示による脅威認知がセキュリティ対策意欲へ与える影響の分析と啓発利用への検討. 情報処理学会インターネットと運用技術研究会, Vol. 2022-IOT-58, No. 4, pp. 1-6, 2022.