個人に合わせた巧妙な標的型メールの分析とその対策手法の研究

学 位 論 文 要 約

Summary of Doctoral Thesis

専　　攻：　情報学専攻　　氏　名：西川弘毅

論文題目：個人に合わせた巧妙な標的型メールの分析とその対策手法の研究

論文要約：

特定の企業・組織の機微情報や設備破壊による稼働防止，経済損失の発生を目的とする標的型攻撃はより深刻となっている．近年でも，日本国内の重要政府機関やインフラ企業が標的型攻撃の対象となり，機密情報や個人情報の漏洩など深刻な被害を受けており，対策が重要である．標的型攻撃では，標的に特化したメール（標的型メール）を利用して標的組織をマルウェアに感染させたのち，機密情報の窃取や，データの不正な暗号化によるシステム停止，といった攻撃者の目的を達成する活動を行う．その背景には，情報通信技術（ICT）の発展がある．マルウェアや不正な通信の検知技術，サーバの堅牢化など，サイバー攻撃防御技術の開発は進んでおり，攻撃者は，標的組織での目的を達成することが以前と比べて難しくなっている．そのため攻撃者は，セキュリティ上で最も弱い点(Weakest Link)である「人」を対象に攻撃を実施する．そして ICT の発展が，攻撃者が人に対する攻撃をより巧妙，かつ，より容易に行うことを可能としている．

攻撃は 1 点を突破すれば成功するのに対し,防御はすべての攻撃を防がなければならず,攻防はそもそも攻撃者有利の関係にある．その中でも「人」の価値観や振る舞いは非常に多様であり，その結果，人を対象とした攻撃も多岐に渡ることになる．このため，防御側がそのすべてに対応することは格段に難しく，標的型攻撃は攻高防低が特に顕著となる事案である．この課題に対し，現在の防御策は全ての人に対して一律の対策を実施するに留まっており，巧妙な攻撃を防ぎ切ることができていないという現状にある．本研究の最終目的は，今後，より深刻化していくことが予想される，人に対する脅威の可能性について明らかにし，人ごとに適した対策を取り入れていくことで，より効果的なセキュリティ対策を実現することである．その第一歩として，本論文では，攻撃者が標的を直接揺さぶることができるメールに着目して研究を行った．今後予想される標的型メールの脅威と対策を見据え，攻撃側が擬態精度あるいは心理操作効力の高いメールをどの程度作成することができるのかに関する検討と，その結果を防御側がどのように対策に活用していくことができるのかに関する検討を行った．

まず，擬態精度を高めたメールの脅威と対策を示した．具体的には，攻撃者は標的者の名前のみを手掛かりとして，インターネットに公開されている情報源から標的者に関する様々な情報を Open Source Intelligence（OSINT）ツールによって次々と取得し，これらの

情報をメールに組み入れることによって，擬態精度が高いメールを標的者ごとに作成することが可能であることを示した．このような擬態精度を高める攻撃への対策として，攻撃者による OSINT の進行を状態遷移図によって定式化し，個々の攻撃段階（どこまでの情報が攻撃者に漏洩しているか）ごとに作成され得る標的型メール文面を整理した．各組織は，自らも OSINT を実行することによって攻撃者が自組織の情報をどこまで入手可能であるか確認し，本状態遷移図に照らし合わせることによって，自組織に届く可能性がある標的型メールの文面をあらかじめ把握することが可能となるため，それを踏まえたプロアクティブな対策選定に資することができる．OSINT を起点とした擬態精度と心理操作効力を高める攻撃の全体像を図 1に示す．

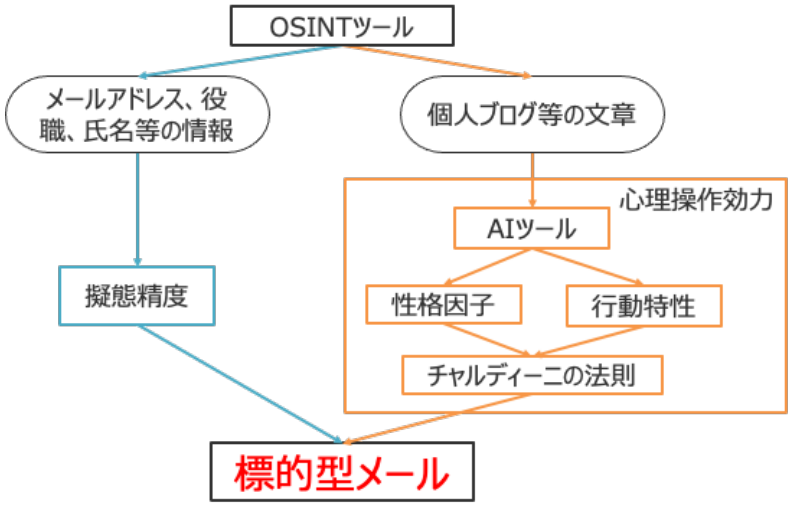

図 1　OSINT ツールと AI ツールを用いた標的型メール攻撃の全体像

次に，心理操作効力を高めたメールの脅威と対策を示した．攻撃者は，AI ツールによって推定した標的者の特性に基づいて，その標的者に有効なチャルディーニの法則を推定することで，心理操作効力の高い標的型メールを標的者ごとに作成することができる．本研究では，将来起こりうる脅威である「攻撃者が, OSINT ツールと AI ツールを駆使し，標的者の性格因子や行動特性を収集して，標的に合った標的型メール文面を作成する」という脅威について，その脅威が現実的であることをアンケート実験により示す．100 人規模のアンケート調査を実施し，攻撃者が, OSINT ツールと AI ツールを駆使し，標的者の性格因子や行動特性を収集して，標的型メールの効果を高めるために応用しうる可能性について，関連研究の紹介を交えながら説明する．まず，人の性格因子と心理操作に関して，ビッグファイブとチャルディーニの法則の説明を行う．さらに，ソーシャルエンジニアリングにおける心理操作に関して，チャルディーニの法則とフィッシングメールとの関係を説明する．続いて，ビッグファイブとチャルディーニの法則との関係と，チャルディーニの法則とフィッシングメールとの関係とを基に，性格に基づいて効果的なチャルディーニの法則を標的型

メールの文面に反映することができる可能性について説明する．次に，行動特性とチャルディーニの法則（説得されやすさ）との関係性と，セキュリティ（プライバシ）意識の個人差は，性格因子よりも行動特性によって引き起こされることから，標的型メールへの引っかかりやすさは，むしろ行動特性に左右される可能性が示唆されることを説明する．これまで，種々の既存研究において説得のされやすさ（チャルディーニの法則に対する感受性）と性格因子との相関関係，あるいは，説得のされやすさ（詐欺師の説得を受け入れてしまいやすさ）と行動特性との相関関係がそれぞれ個別に調査されてきたのに対し，説得のされやすさ（チャルディーニの法則に対する感受性）を性格因子と行動特性のコンビネーションによって分析したことが，本研究の第一の貢献である．

攻撃者は OSINT ツールと AI ツールを活用することで，インターネットに公開されている情報をもとに，標的者の性格因子と行動特性を推定することが可能となる．今後，インターネットと AI の進歩により，OSINT ツールで収集可能な情報が増殖することで，ソーシャルエンジニアリングの脅威がますます深刻化することを暗示している．

そこで本章では，個々人の性格因子や行動特性が，OSINT ツールと AI ツールにより収集することが可能となることを前提として，「標的型メールの文面に用いられたチャルディーニの法則」と「個人の性格因子および行動特性」との間の関係性を，ユーザ実験により調査する．

ここで,フィッシングメールとチャルディーニの法則には関係性があることは示されており，また標的型メールにおいても，その一部の法則については関係性が示されている．しかし，標的型メールにおいて，6 つのチャルディーニの法則全てに対する関係性は検証されていない．そのうえで,チャルディーニの法則に対する反応度は，個々人の性格因子と行動特性の両者に影響されることが予想される．すなわち，標的型メールの開封率も個々人の性格因子と行動特性の両者に影響されると考えられる．

以上より，本論文のリサーチクエスチョンとして，以下を掲げ、100 人規模のユーザー実験を実施した。

RQ1：チャルディーニの法則を標的型メールの文面作成に応用することで，標的型メールが開封されやすくなるのか.

RQ2：個々人で有効なチャルディーニの法則は，性格因子と行動特性の両者によって異なる傾向を示すのか.

ユーザ実験の結果，標的型メールの文面において，チャルディーニの法則と，個人の性格因子および行動特性との間には関係性があることが判明した．すなわち攻撃者は，OSINT ツールと AI ツールを用いて，標的者の性格因子および行動特性に有効な文面を推定可能であると結論付けることができる．

まず，攻撃者が標的者の名前を知っている場合，OSINT ツールによって，たとえばメールアドレスと所属組織を取得し，図 2のような擬態精度を高めた標的型メールを作成することができる．自分の名前がメールの本文に含まれている，差出人欄が所属組織のメールアド

レスである，などから，標的者は正規のメールであると感じやすくなる．

さらに攻撃者は，OSINT ツールと AI ツールによって標的者の性格因子や行動特性を調べることができる．図 2の標的型メールに対し，心理操作効力を高めるために，チャルディーニの法則である希少性と権威を使用したメールをそれぞれ図 3と図 4に示す．プレーンメールを開きにくい傾向にあり，かつ誠実性スコアが高い人間であれば希少性は有効であるため，標的者の行動特性と性格因子から希少性が有効であると判断できた場合には，図 3のような標的型メールを標的者に送ることによって，攻撃者はさらに標的型攻撃の効果を高めることができる．プレーンメールを開きにくい傾向にあり，かつ外向性スコアが高い人間であれば権威は有効であるため，標的者の行動特性と性格因子から権威が有効であると判断できた場合に，権威を用いることでより効果的な標的型メール文面を作成することができる．

その他の法則についても，その法則の特性を活かして標的型メールを作成することが可能である．少し文章を追加する程度で十分であることから，攻撃者の費用対効果は非常に高いと考えられる．

差出人	*******@corp.co.jp （所属組織に詐称したアドレス）　所属組織
件名	ウイルスソフト導入のお願い
宛先	*******@corp.co.jp　メールアドレス
本文	○○様　名前 技術部です。 下記サイトより、新しいアンチウィルスソフトの導入を行ってください。 http://hogehoge.com（悪性URL） 以上、よろしくお願い致します。

図 2　標的型メール例

差出人	*******@corp.co.jp （所属組織に詐称したアドレス）
件名	ウイルスソフト導入のお願い
宛先	*******@corp.co.jp
本文	○○様 技術部です。 下記サイトより、新しいアンチウィルスソフトの導入を行ってください。 http://hogehoge.com（悪性URL） ※ファイルは会社規定により本日のみ閲覧可能です。 以上、よろしくお願い致します。

図 3　希少性を利用した標的型メール例

差出人	*******@corp.co.jp （所属組織に詐称したアドレス）
件名	ウイルスソフト導入のお願い
宛先	*******@corp.co.jp
本文	○○様 技術部です。 XX部長より、新しいアンチウィルスソフトの導入の要請がありました。 下記サイトより、ソフトの導入を行ってください。 http://hogehoge.com（悪性URL） 以上、よろしくお願い致します。

図 4　権威を利用した標的型メール例

このような心理操作効力を高める攻撃への対策として，標的者が標的型メールを受けた際には，メールに含まれている説得のフレーズを通知することで標的者に注意を促すことが有効であると考えられる．

メール文面からチャルディーニの法則をヨーロッパの金融機関における Security Operation Center(SOC)で収集したデータを実際のフィッシングメールを基に，フィッシングメールの本文中で用いられているチャルディーニの法則を検知する識別器を作成した研究がある．本研究において作成した識別器は，メールがフィッシングメールかそうではないかを判断するのではなく，そのメール文面が，ユーザにとって引っ掛かりやすいメール本文であるかを推定する識別器である．本研究では，データに対して識別器を適用した結果を基に統計分析をした結果，一つのメール文中で用いられるチャルディーニの法則の種類の数が増えるほど，ユーザに影響を与えやすいことが示唆される結果を得た．本研究結果は，フィッシングドメインのテイクダウンや，ブラックリストの作成，顧客への注意喚起に活用できる．また，メールに用いられているチャルディーニの法則の度合を表示することで，インシデントレスポンスを効率化するといった新しい方法が期待できると述べている．攻撃者の能力が向上した場合における防御策としては，防御側も自組織内に所属する人物の性格因子や行動特性を把握することで，個々人に応じた標的型メール対策を実施することが考えられる．

個々人に応じた標的型メール対策として，自分にとって「騙されてしまいやすい」チャルディーニの法則が用いられているメールを受信した際に“肩を叩いてもらう”という方法が有効であると考えられる．これは，説得心理学における脅威アピールの考えに該当する．そのためには，メールを受け取った際に，メール文面からチャルディーニの法則が使われている箇所を特定することと，ユーザの個人特性を踏まえてどのように提示するかを判断するシステムが必要となる．しかし，脅威アピールの既存研究においても，脅威の通知に効果がある場合と，そうではない場合があり，本研究で検討しているシステムにおいても，個々人の特性に応じて効果が異なると考えられる．例えば，アラートが必要以上に多い場

合，人によってはアラートに慣れてしまい無視することがあり得る．

チャルディーニの法則が使われたメールが必ずしも標的型メールであるとも限らないので，ユーザに提示するアラートを絞ることで誤アラートを減らすことや，真に有効なアラートのみを提示することが必要となる．たとえば，権威の法則が使われたメールである場合に，そのことをアラートで通知することが有効である人物に対しては，アラートを提示し，権威の法則に対するアラートが有効ではない人物に対してはアラートを提示しない方法が効果的であると考えられる（図 5）．具体的には，個々人の性格因子や行動特性に応じて，アラートを挙げる閾値を調整する形となる．ここで，個々人において，4 章で示した「標的型メールにおいて有効なチャルディーニの法則」と，本章で示す「アラートを通知する際に提示すると有効なチャルディーニの法則」とでは，それぞれコンテキストが異なる．そのため，チャルディーニの法則の効果として，異なる結果が得られると考えられる．

ここで，対象者に対して直接チャルディーニの法則の効果を尋ねずに，性格因子や行動特性を媒介とするかを説明する．チャルディーニの法則を直接尋ねると，回答者が，良い解答をしようとするため正しく反応を測定することができない．そのため，性格因子や行動特性を通じて推定できるようにすることで，回答者によるバイアスを除いて，提示する効果を推定することが期待される．

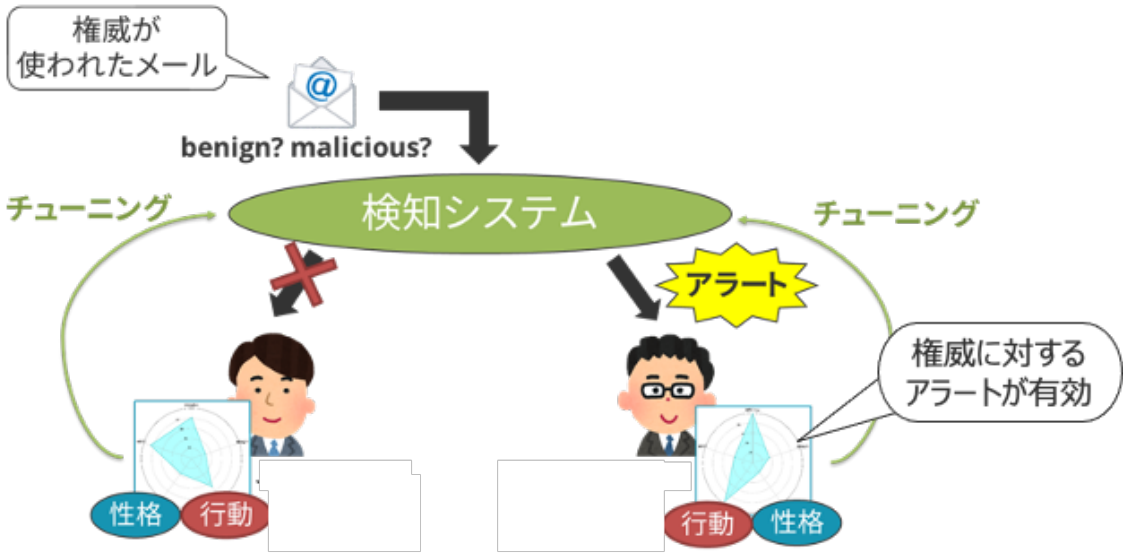

図 5　検知システムのチューニング

本研究では，次のふたつのリサーチクエスチョン（RQ）を設定し、個々人の性格因子や行動特性が，組織内でのアンケート等によって収集し，そのデータを組織内部で活用可能であることを前提とし，「標的型メール文面にて使われているチャルディーニの法則への注意を促すアラート」と「個人の性格因子および行動特性」との関係性を，ユーザ実験により調査する．

RQ1：単純にアラートを提示することと，各種のチャルディーニの法則への注意を促すアラートとではメールを開くという行動の傾向は異なるのか.

RQ2：個々人で有効なチャルディーニの法則への注意を促すアラートは，性格因子と行動特性の両者によって異なる傾向を示すのか.

400 人規模のアンケート調査を実施し，効果的な注意喚起の方法（メール内のどの説得フレーズに対してアラートを提示すると，標的者は標的型メールへの不信感を高めるか）も，標的者の性格因子と行動特性の両者によって異なることを示した.

これまでの既存研究においては，標的型メールの脅威分析（攻撃側が擬態精度あるいは心理操作効力の高いメールをどの程度作成することができるのか）に主眼が置かれていた．これに対し本研究では，標的型メールへの対策（防御側が脅威分析の結果をどのように対策に活用していくことができるのか）にまで足を踏み込んだ．攻撃者による脅威をあらかじめ把握することで，標的者に応じた有効な対策を配備することができることを示したことが，本研究の第二の貢献である.

本研究は，著者が知る限り，個々人の特性に応じてセキュリティシステムの防御方法を変える考えを具体化した初めての研究である．企業においては現在もメールは多く利用されており，標的型メール攻撃の脅威は依然として考慮する必要がある．一方で昨今ではチャットツールを導入する企業もあり，大手企業においてのコミュニケーションのうち，14.2%がチャットツールによって行われているという調査結果がある．スピーディにコミュニケーションができるようになった，という回答が 41.9%あるなど，チャットツールの利便性が示唆されており，チャットツールを利用したコミュニケーションは今後増えていくと考えられる．攻撃者もメールだけではなくチャットツールを駆使することが考えられるため，今後はメールだけではなく，チャットツールにおいてもソーシャルエンジニアリングを考慮する必要があると考えられる.

さらに，対話型のチャットボットなど，AI の進歩により，あたかも人間とやりとりをしているかのようなツールの開発が進んでいる．今後このようなツールの発展により，ソーシャルエンジニアリングの対話部分すらも自動化される危険性があり，「人」を守ることがますます重要になると考えられる.

さらに今後，リモートワークなどによって，従業員の端末を利用する BYOD が進むことも考えられる．そのため，従業員がばらまき型の攻撃に対しても引っかからないように注意する必要がある．ばらまき型の手法の一つとして，フィッシングメールにおけるアラートのようなメール（Amazon のアカウントが停止しました，など）が考えられる．このようなメールが到達した際にも，アラートに対する反応をあらかじめ特定しておくことで，偽アラートに対して注意を促すことで，フィッシングメールに気づくことが期待される.

今後，標的型攻撃において人を対象とした攻撃がなくなることはなく，その手口も発展していくと考えられる．本研究によって，攻撃者による脅威をあらかじめ把握することで人への有効な対策を提案することに貢献することができた．本研究は，著者が知る限り，

個々人の特性に応じてセキュリティシステムの防御方法を変える考えを具体化した初めての研究である．本研究が，先に述べた，Weakest Link である人に着目した研究の足掛かりとなることを期待する．