知識グラフを用いた個人適応型サイバーセキュリティ学習システムの開発

東野 正幸[1,a)]

概要：サイバーセキュリティ教育においては，一般に組織の構成員の理解度が不均質であるため，個々の理解度に応じた学習内容を提供することが望ましい．組織内の講習や研修では，学習に要する時間が業務負担となる場合があり，短時間で学習できる方法が求められる．そこで本研究では，サイバーセキュリティの攻撃手法と対策手法の知識グラフを構築し，これを活用して個々の理解度と学習コンテンツの難易度に応じた学習コンテンツの回答順序を提供し，さらに共通する知識のうち過去の回答から習得済みと判断できる項目を除外することで，効果的かつ効率的な学習を支援するシステムを開発する．

キーワード：サイバーセキュリティ，人的セキュリティ，適応学習，知識グラフ

Development of a Personalized Cybersecurity Learning System Using Knowledge Graphs

Masayuki Higashino[1,a)]

Abstract: In cybersecurity education, the level of understanding among organizational members is generally heterogeneous, making it desirable to provide learning content that corresponds to individual levels of comprehension. In organizational in-house training sessions and workshops, the time required for learning may disrupt regular operations, creating a need for methods that support efficient learning within a limited time frame. To address this issue, this paper constructs a knowledge graph of cybersecurity attack techniques and countermeasures and develops a system that supports effective and efficient learning. The system uses the knowledge graph to determine the order of answering learning content according to both individual comprehension levels and the difficulty of the learning content. Moreover, the system excludes learning content, which can be inferred as already acquired from past responses, from the list of provided learning content.

Keywords: Cybersecurity, Human Factor in Cybersecurity, Adaptive Learning, Knowledge Graph

1. はじめに

サイバーセキュリティ攻撃は，システムセキュリティや物理セキュリティだけで完全に防ぐことが困難である．そのため，教育や訓練といった人的セキュリティも重要である．例えば，標的型攻撃メールは，メールゲートウェイやメールソフトウェアのセキュリティ機能といったシステムセキュリティだけでは完全に防ぐことは難しく，すり抜けた標的型攻撃メールへの対応として，訓練や教育等の人的セキュリティの重要性が高い [1].

また，近年では生成 AI を悪用したサイバーセキュリティ攻撃のリスク [2] や，生成 AI の活用に伴うリスク [3], [4], [5] が懸念されている．例えば，攻撃者によるフィッシングメールの作成においては，V-Triad [6] と呼ばれるフィッシングメールを設計するためのルールセットと生成 AI を併用することで，より多くの人を騙すメールを容易に生成できることが示唆されている [7]．そのため，人的セキュリティ対策に必要な知識や技術が高度化及び複雑化すること

[1] 鳥取大学工学部
Faculty of Engineering, Tottori University, 4-101 Koyama Minami, Tottori 680–8552, Japan
[a)] higashino@tottori-u.ac.jp

が予想される．

一方で，組織におけるサイバーセキュリティ教育においては，構成員のサイバーセキュリティに関する理解度が一様ではなく，それぞれの役割や習熟度に応じた学習内容を提供することが求められる [8]．また，組織内の研修においては，学習にかかる時間が業務負担となる場合があり，短時間で効果的に学習できる方法が必要とされている．

そこで本研究では，知識グラフを用いた個別適応型のサイバーセキュリティ学習システムを開発する．サイバー攻撃の手法と対策手法を体系化した知識グラフを構築し，学習者の理解度と学習コンテンツの難易度に応じた学習順序を提示する．また，学習者が過去の回答からすでに習得済みと判断できる知識を除外して不要な学習を削減する．

提案システムにより，個人に合ったサイバーセキュリティに関する学習が行えるだけでなく，学習に取り組む時間の短縮も可能とし，効果的で効率的なサイバーセキュリティ教育の実現を目指す．

なお，本稿ではサイバーセキュリティ攻撃の例として，フィッシング攻撃を題材として取り扱う．

2. 提案手法

サイバーセキュリティ攻撃の例として，フィッシング攻撃が挙げられる．一般的にフィッシング攻撃の対策としてメールの差出人が正しいことの確認やメールに記載されたURLが正しいものであることの確認などが挙げられる．しかし，そのような対策方法を伝えて対応できる人もいれば，どのように対応すれば良いか分からない人もいる．この差はフィッシング対策に関する基礎的な知識の差だと考えられる．つまり，学習者が学習内容に関する基礎知識を持っていればそれを理解できる．また，攻撃手法への対策に必要な知識の数が多いほど理解が難しくなる．このことから，個人の理解度に合った学習難易度の問題を優先する．

学習においては苦手な問題を優先することで教育効果が高い事例が報告されている [9]．まだ習得できていない基礎知識が多く含まれる問題ほど苦手であり，さらに，1回の回答でより多くの知識を習得できると考えられる．そのため，未習得の基礎知識が多く含まれる問題を優先する．

以上より，学習者に合った難易度の問題を優先し，さらにより多くの知識を習得できる問題を優先して学習を進める．

2.1 知識グラフの構築

フィッシング対策として何を学習するべきなのかを明らかにし，その知識間の関係性を知る必要がある．そこで，NIST Phish Scale User Guide (TN 2276) [1] を情報源とした知識グラフを構築する．

攻撃手法に関するノードとして，Phish Scaleで定義されている「フィッシングメール自体の観察可能な特徴」の「Cue Type」（手がかりの種類）のノードを作成する．このノードを起点に「Cue Name」（手がかりの名称）のノードを作成して接続する．

NIST Phish Scale User Guide (TN 2276) [1] に記載されている「フィッシングメール自体の観察可能な特徴」のうち，「Technical indicator」（技術的指標）に関する「Cue Type」（手がかり種類）と「Cue Name」（手がかりの名称）を以下に引用する．

- Technical indicator
  - Attachment type
  - Sender display name and email address
  - URL hyperlinking
  - Domain spoofing

これらに続くノードは独自に作成する必要があり，攻撃手法に関する学習項目であるフィッシングに用いられる具体的な攻撃手法のノードを独自に作成して接続する．

さらに，各攻撃手法の学習に必要な基礎知識や攻撃手法ノード間の知識の共有関係を知るために，具体的な攻撃手法のノードの下にサイバーセキュリティに関連する知識ノードを独自に作成して接続する．

図 1 に構築した知識グラフを示す．Phish Scaleの「Cue Type」のうち，茶色の「Technical Indicator」を起点として，黄色の「Cue Name」にリンクする．さらに薄い赤色の具体的な攻撃手法に細分化し，その周囲に濃い赤色の基礎知識をリンクする．Phish Scaleで定義されているものは黄色の「Cue Name」までであり，それ以降は独自に作成した内容に基づいている．

2.2 学習項目選択関数

式 (1) の目的関数を最小にするのに最も効果的な学習項目から順に学習する．

$$\sum_{n=1}^{N} a_n \cdot \frac{\alpha - |L_{a_n} - L_i|}{\alpha} \cdot \frac{1 + K_n + \sum_{m=1}^{N} a_m P_{nm}}{\beta} \quad (1)$$

α は学習項目に設定する難易度の最大値，β は $1 + K_n + \sum_{m=1}^{N} a_m P_{nm}$ の取り得る値の最大値とする．

N は学習項目の個数とする．$a_n, (1 \leq n \leq N)$ は学習状況を表し，$a_n = 1$ であれば学習が終わっていない項目（初期値）であり，$a_n = 0$ であれば学習が完了した項目を表す．K_n は学習項目 a_n のみに関係する知識の個数とする．P_{nm} は2つの学習項目 a_n と a_m が共通して持っている知識を1つ経由して繋がる経路の数とする．$n = m$ のときは0とする．

L_{a_n} は学習項目 a_n の難易度とし，必要な知識の数が少ないものから順番に $1, 2, 3, \cdots$ と付番する．$L_i (1 \leq L_i \leq \alpha)$ は学習者の習熟度とする．

i を学習回数とし，R_i を学習結果とする．学習を行うたびに，習熟度を $L_{i+1} = L_i + R_i$ により更新する．学習回数

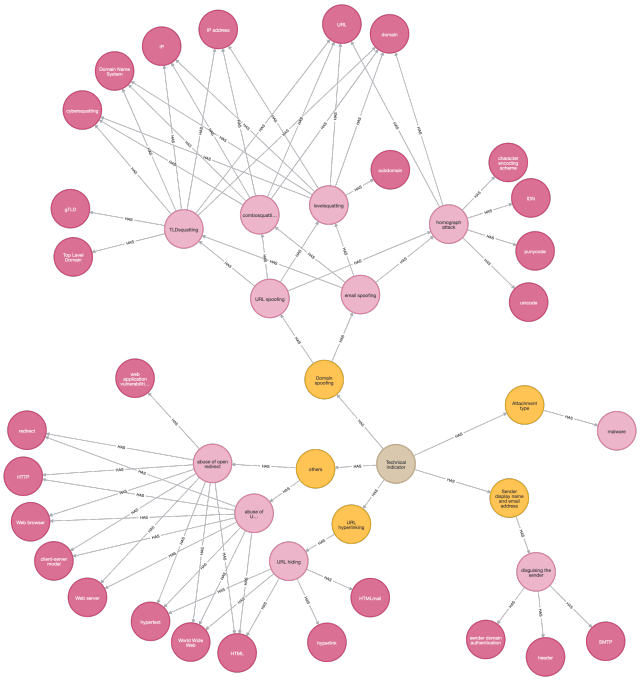

図 1 NIST Phish Scale User Guide (TN 2276) の Technical indicator を起点とした知識グラフ

が i 回目の学習項目について理解できたと判断したとき，$R_i = 1$ とし，習熟度を $L_{i+1} = L_i + R_i$ により更新し，学習状況を $a_n = 0$ として学習済みとする．学習回数が i 回目の学習項目について理解できていないと判断したとき，$R_i = -1$ とし，習熟度を $L_{i+1} = L_i + R_i$ により更新し，学習状況を $a_n = 1$ として学習済みとする．

3. 実験

3.1 実験の目的

提案手法における個人適応に関する評価として，実験参加者ごとの学習項目の回答解順を記録し，実験参加者の習熟度（L_i）と出題した学習項目の難易度（L_{a_n}）がどの程度一致しているかを確かめる．また，アンケートによる評価として，実験終了後に「難易度が調整される感覚はあったか.」等を項目とする主観評価を実施する．

3.2 実験の方法

実験はゲームブック形式による筆記試験で行う．提案手法に基づき，各学習項目の番号，問題文，正答，及び解説文を作成する．また，提案手法の知識グラフと評価関数が実装されたプログラムを用意する．実験参加者はこのプログラムが提示する番号の問題に取り組み，その回答をプログラムに入力すると，正誤判定が行われ，その結果に応じて次の問題の番号が提示される．各学習項目の問題に全て正答するまで回答を繰り返す．

プログラムは，学習項目として定めた攻撃手法に関する問題を評価関数によって選ばれた順番に出題する．出題された問題に正解すると，その学習項目について学習が完了したとみなし，それ以降の出題候補から除外する．間違えた場合は，その問題の解説文と関係する知識項目の説明文を提示することでフィードバックとする．また，間違えた問題は学習がまだ完了していないとみなし，それ以降の出題候補として残す．このように学習を繰り返し，全ての問題に正解することで学習終了とする．また，学習終了後にはアンケートを行う．

3.3 実験の結果と考察

提案手法を用いて 10 名の実験参加者が学習を行った．提案手法を用いたゲームブックによる学習結果について，実験参加者の習熟度（Proficiency）と学習項目の難易度（Question Difficulty）の平均と標準偏差の推移を **図 2** に示す．学習の序盤から中盤については，出題された学習項目の難易度が実験参加者の習熟度に近い値を維持していることから，個人適応した学習ができていると言える．また，学習の終盤については，本実験では全ての学習項目について正答するまで学習を続けることとしていたため，最後まで残った難易度の低い学習項目が出題されている．このように，実験参加者の習熟度に対して，難易度の低い学習項目しか残っていない場合，それ以降の学習を打ち切ることで，総学習時間を短縮する判断も可能であると考えられる．また，各問題の難易度と回答に要した時間の平均と標準偏差を**表 1** に示す．問題と難易度と回答に要した時間には相関がなく，簡単な問題であっても長時間要する場合と要しない場合が観察された．

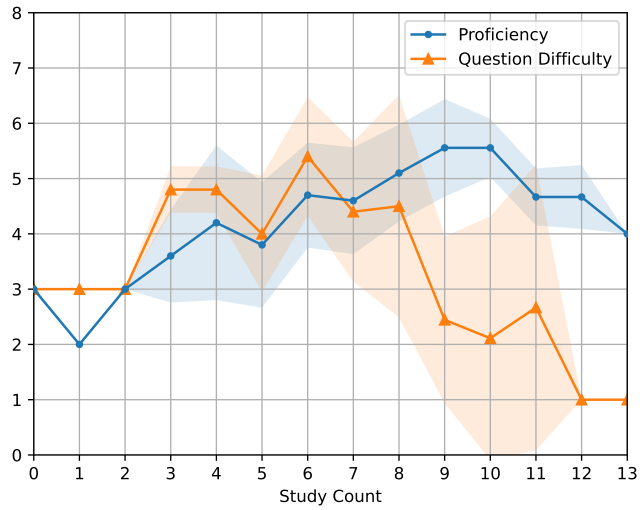

図 2 実験参加者の習熟度（Proficiency）と学習項目の難易度（Question Difficulty）の平均と標準偏差の推移

表 1 問題の難易度と回答に要した秒数の平均と標準偏差

問題番号	難易度	平均	標準偏差
1	1	31.2	26.6
2	2	19.2	8.9
3	5	13.3	10.4
4	3	21.5	36.1
5	4	14.9	10.9
6	3	14.8	10.8
7	5	21.5	11.6
8	6	28.9	23.6

アンケートの結果については，「難易度が調整されている感覚は有ったか.」という設問に対して，4 名からは「有った.」，6 名からは「無かった.」という趣旨の回答が得られた．また，「有った.」群と「無かった.」群との学習項目の正誤率に大きな違いは無かった．また，「特に URL とドメインに関連する問題が難しく感じた.」と言う意見も得られた．提案手法では，学習項目の難易度として，その学習項目の習得に必要な基礎知識の数を用いている．専門知識については十分簡単な基礎知識まで分解していることを前提としているが，例えば URL やドメインは基礎知識といっても複雑な仕様が定められており，初学者にとっては難しい内容であった可能性がある．このため，基礎知識についてはさらに細分化を行う必要があると考えられる．

4. おわりに

本稿では，知識グラフを用いた個人適応型のサイバーセキュリティ学習システムの検討を行った．標的型攻撃メールを題材として NIST Phish Scale User Guide (TN 2276) [1] からフィッシングに関する攻撃手法と防御手法に関する知識グラフを構築し，それらの知識グラフに基づき作成した学習項目の正答に必要な基礎知識の数から学習項目の難易度付けを行なった．また，学習者の習熟度，学習項目の難易度，及び関連する基礎知識の関係から学習項目の学習優先度を算出し，それらに基づき回答する学習項目の順番を決定する手法を提案した．

提案手法のプロトタイプとしてゲームブックを作成し，実験参加者 10 名による学習を実施した．その結果，実験参加者の習熟度に応じた難易度の問題を出題できることを確認した．また，実験参加者の習熟度が十分に高く，かつ全ての未着手の学習項目の難易度が低い場合に，それ以上の学習を打ち切ることで，総学習時間を短縮するための判断材料を示した．

今後は，提案手法を用いた場合と用いなかった場合での比較実験を行う予定である．また，今後の課題として，本稿では扱わなかった NIST Phish Scale User Guide (TN 2276) [1] におけるフィッシングメール判別の評価基準（Criteria for Counting Cues）の他の手がかりの種類（Cue Type）や，標的型攻撃メールにおけるフィッシング以外の脅威に対する学習に提案手法を適用可能かどうかを検証することが挙げられる．また，大規模言語モデルを用いて知識グラフと問題，回答，及び解説の自動生成を行うことで，本稿では扱わなかった他の攻撃手法と対応手法にも容易に対応可能とすることを検討している．

謝辞 本研究の一部は公益財団法人電気通信普及財団の研究調査助成を受けたものである．

参考文献

[1] Dawkins, S. and Jacobs, J.: NIST Phish Scale User Guide, Nist series Technical Note (NIST TN) 2276, National Institute of Standards and Technology (NIST) (2023).
[2] Neupane, S., Fernandez, I. A., Mittal, S. and Rahimi, S.: Impacts and Risk of Generative AI Technology on Cyber Defense (2023).
[3] Humphreys, D., Koay, A., Desmond, D. and Mealy, E.: AI hype as a cyber security risk: the moral responsibility of implementing generative AI in business, *AI Ethics*, Vol. 4, pp. 791–804 (online), DOI: 10.1007/s43681-024-00443-4 (2024).
[4] The Open Worldwide Application Security Project (OWASP) : OWASP Top 10 for LLM Applications 2025. Version 2025, Guide (2024).
[5] The Open Worldwide Application Security Project (OWASP) : Agentic AI - Threats and Mitigations: OWASP Top 10 for LLM Apps & Gen AI Agentic Security Initiative. Version 1.0, Guide (2025).
[6] Vishwanath, A.: *The Weakest Link: How to Diagnose, Detect, and Defend Users from Phishing*, The MIT Press (2022).
[7] Heiding, F., Schneier, B., Vishwanath, A., Bernstein, J. and Park, P. S.: Devising and Detecting Phishing: Large Language Models vs. Smaller Human Models (2023).
[8] Ministry of Economy, Trade and Industry (METI) and Information-technology Promotion Agency, Japan (IPA): Cybersecurity Management Guidelines for Japanese Enterprise Executives Ver. 3.0, Guide (2023).
[9] 和久友親，田村哲嗣，川瀬真弓：学習者の理解度に応じた自動問題生成 AI システムの開発，日本教育工学会論文誌 (2024).