「マルチメディア，分散，協調とモバイル
(DICOMO2017)シンポジウム」 平成29年6月

個人のセキュリティ意識自動評価技術の提案

山本 匠* 西川 弘毅* 木藤 圭亮* 河内 清人*

概要：様々なセキュリティ施策や対策が導入されているにも関わらず，組織におけるセキュリティ事故は増加の一途をたどっている．この原因の 1 つとして，セキュリティ意識や攻撃のされやすさが組織のスタッフごと異なり，スタッフ全員に一律で同じセキュリティ教育や訓練を実施しても十分な効果が得られないことがあると，著者らは考えている．そこで本研究では，インターネットや社内ネットワークから収集可能な個人の情報と，標的型訓練メール開封数などのセキュリティ事故情報との関係性を分析し，セキュリティ事故を起こしやすいスタッフの共通の人物モデルを作ることで，スタッフごとにセキュリティ意識を評価する仕組みを提案する．これにより，個人に最適なセキュリティ教育の実施やセキュリティリスクの高い人物へのセキュリティ対策導入が可能となると期待される．

Proposal of automated personal security awareness evaluation

TAKUMI YAMAMOTO* HIROKI NISHIKAWA*
KEISUKE KITO* KIYOTO KAWAUCHI*

1. はじめに

組織の機密情報や資産を守るために，サイバー攻撃に対する取り組みが積極的に行われている．その 1 つとして，サイバー攻撃やセキュリティに関する教育や訓練である．セミナーや E-learning でサイバー攻撃やその対策に関する知識を学習するものや，模擬的な標的型攻撃メールの送付による標的型攻撃への対応を訓練するものなどがある．しかし，このような取り組みが行われていながらも，セキュリティ事故は増加の一途をたどっている．

Verizon Business 社が発表した企業の情報流出事件に関する実態調査[1]では，情報流出を経験した企業の 59%が，セキュリティポリシと手順を定めておきながらも，それを正しく実行していなかったと報告されている．この調査結果からも，どれだけセキュリティ対策を導入していたとしても，それを実施する組織やスタッフにセキュリティ対策の効果が強く依存してしまっていることがわかる．

一方，攻撃者の視点に立ってみると，攻撃者は，標的組織に気づかれずに攻撃を成功させるために，その組織の情報を事前に十分調査した上で，攻撃の成功率が最も高いアプローチをとるだろう．組織の情報としては，例えば，組織が利用しているシステムやそのバージョン，外部との窓口，人員の情報，役職，関連組織，組織の取り組み内容などである．人員の情報としては，例えば，上司/同僚/友人などの交友関係，趣味嗜好，ソーシャルネットワークサービス（SNS）などの利用状況などである．

攻撃者は，このような情報を集め，組織における脆弱な人間を見つけ出し，そこから組織に入り込み，徐々に組織の内部に侵入していくと考えられる．

企業を例に考える．一般に，人事や資材などのスタッフは，他のスタッフよりも自組織外の人物とのやりとりが多く（例えば，人事担当であれば就職活動中の学生，資材担当であれば物品の購入先)，これまでにやり取りをしたことの無い人物からメールを受けとる可能性が高い．そのようなメールが多いスタッフであれば，見知らぬメールアドレスから攻撃メールが届いても，不審に思わず開封する可能性が高いと予想できる．

また, Twitter や Facebook などのソーシャルメディアで，組織の情報を不用意に掲載しているスタッフは，セキュリティ意識，特に情報漏えいに関する意識が低いと言える．そのため攻撃者はそのようなスタッフを最初の標的にする可能性が高いと考えらえる．セキュリティ意識が低い人物に共通する特徴はこれ以外にも多数存在すると考えられ，調査が必要である．

このように，組織のスタッフ（人物）によって攻撃のされやすさが異なると考えられ，組織のスタッフ全員に一律で同じセキュリティ教育や訓練を実施しても，十分な効果を得られないと考えられる．セキュリティ意識が最も低いスタッフに合わせたセキュリティ教育や訓練を全てのスタッフに課せば，不必要な作業が増え，業務効率低下に繋がる．

そこで本研究では，スタッフのプロファイル情報（性別，年齢，所属，上司，メールの送受信頻度，インターネット利用頻度，出社/退社時間，インターネット上に公開している個人情報の量，趣味など）とセキュリティ事故情報（訓

* 三菱電機株式会社 情報技術総合研究所，〒247-8501 神奈川県鎌倉市大船 5-1-1，Information Technology R&D Center, Mitsubishi Electric Corporation, 5-1-1 Ofuna, Kamakura, Kanagawa, 247-8501, JAPAN

練メール開封数，悪質サイト訪問数，マルウェア検知数，ポリシー違反の回数，実行ファイルダウンロード数，ファイルダウンロード数，インターネット利用量など）とを自動的に収集・分析し，セキュリティ事故を起こしやすいスタッフの共通の人物モデルを作ることで，スタッフごとにセキュリティ意識を評価する仕組みを提案する．

このように，スタッフごとセキュリティ意識を評価し，攻撃のされやすいスタッフに，適切なセキュリティ教育や訓練を実施することで，組織全体の作業効率を落とすことなくセキュリティを向上することができると考えられる．

2. 関連研究

個人のセキュリティ意識を評価するための既存研究としては，文献[2][3][4]が知られている．本章では既存研究について概説し，本研究との差異を示す．

中澤らの研究[2]では，性格，経験，環境の 3 要因を基に，個人に最も適したセキュリティ対策を提示するシステムの実現を目指している．研究の第一ステップとして，性格と本人認証技術に関するセキュリティ意識との相関に焦点を当て 400 人規模のアンケート調査を実施し，いくつかの性格特性と種々の認証技術に関するセキュリティ意識との間に 関係性があることを確認している．

片山らの研究[3]では，IT 被害に遭いやすいユーザや組織の特性を明らかにし，リスクの高いユーザや組織の早期検知やリスクに応じた攻撃対策につなげる技術の提案を行っている．アンケート調査で抽出した心理特性や行動習慣との相関関係を調べ，ユーザの業務中の行動を PC 操作ログから観測することで，ユーザの行動ログの観測のみから IT 被害に遭う可能性（IT リスク）が高いユーザを検知し対策する．規約表示時間の短いユーザはウィルス感染被害に遭いやすいといった関係が導き出されている．

守屋らの研究[4]では，「標的型攻撃の被害に遭うリスクの高い人物に共通する特徴を明らかにすること」と，「標的型攻撃の被害に遭いやすい人物の予測の能否」の 2 つの課題を設定し，標的型攻撃訓練時のデータを解析し，攻撃に対し不適切な行動をとった人に共通する特徴の分析，ならびに機械学習の適用により，そのような人を事前に予測できるかについて評価を行っている．評価の結果，質問紙調査から実験で不適切な行動をとった人物には内向的な傾向があったことが報告されている．また，機械学習を用いた予測では比較的高い精度で標的型攻撃の被害に遭うリスクが高い人物を抽出できることも明らかにされている．

[2-4]の研究は，個人や組織ごとに最適なセキュリティ対策を実施するために，個人や組織のセキュリティ意識や IT リスクを導き出すという画期的な研究である．しかし心理や性格といった定量化の難しい特徴を利用しているため，特徴を取得する際の環境にモデルの精度が影響を受ける可能性がある．また，心理や性格とセキュリティとの関係性を導き出す初期段階の取り組みということもあり，アンケートベースの情報収集になっており，手間暇を要してしまう．

3. 提案方式

本研究のモチベーションも既存研究[2-4]と同じく，個人や組織ごとに最適なセキュリティ対策を実施するために，個人や組織のセキュリティ意識や IT リスクを導き出すことである．既存研究との大きな違いは，定量化のしやすい情報を利用する点と自動化に主眼を置いている点である．

3.1 コンセプト

提案方式の大まかな流れは，情報の収集（3.2），関係性のモデルの作成（3.3），セキュリティ意識の推定（3.4）である．まず，セキュリティ意識を調査したい組織のスタッフの情報（プロファイル情報）を収集する．セキュリティ事故情報との関係性を分析し，どのようなスタッフがどのようなセキュリティ事故を起こしやすいか（セキュリティ意識）を表した関係性（モデル）を特定する（図 1）．最後に，特定した関係性をもとに，新たに調査したい対象（例えば，新入社員）が起こす可能性のあるセキュリティ事故を推測する（図 2）．各ステップの詳細については次節以降に記載する．

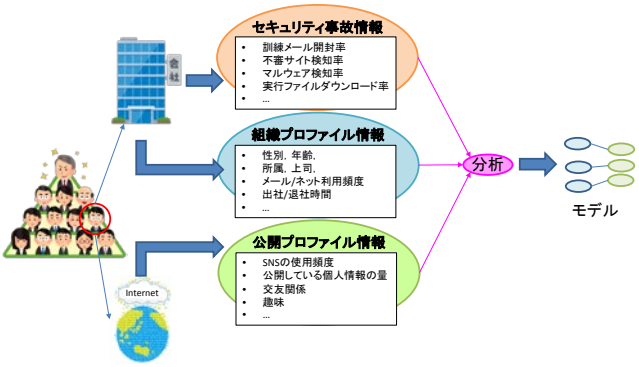

図 1　提案方式のコンセプト（モデルの作成）

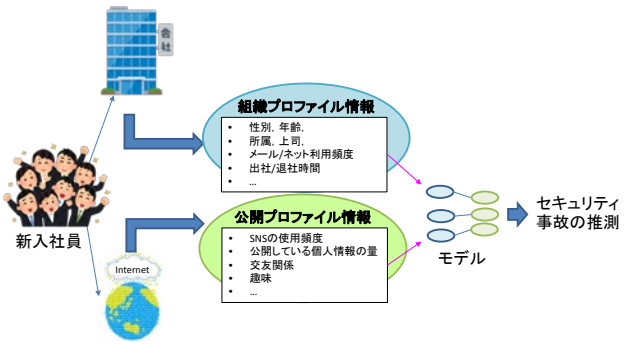

図 2　提案方式のコンセプト（セキュリティ事故の推測）

3.2 情報の収集

提案方式で利用する情報とその収集方法について記載する．

3.2.1 収集する情報

① プロファイル情報

プロファイル情報は，インターネット上で公開されている公開プロファイル情報と，組織の管理職や IT 管理者であればアクセスできる組織プロファイル情報の 2 つからなる．図 3 にプロファイル情報のイメージを記載する．

●公開プロファイル情報

公開プロファイル情報は，インターネットサービスの使用頻度や公開されている個人情報の量，などの情報である．クローリングやスクレイピングを許可しているインターネット上のサービスから，公開プロファイル情報を収集する．クローリングした情報を解析し，個人の興味に関する情報についても抽出する．個人の氏名やメールアドレスを含むページをインターネット上のサービスをから収集する．

TF・IDF などの自然言語処理技術を活用し，収集されたページ内でキーとなる単語をピックアップし，個人の興味に関する情報の DB を作成する．既存技術の，Martego CE[5]や theharvester[6]を組合せて公開プロファイル情報を収集することも可能である．

●組織プロファイル情報

組織プロファイル情報は，性別，年齢，所属，上司，メールの送受信頻度，インターネット利用頻度，出社/退社時間などの情報である．組織の管理職や IT 管理者であればアクセスできる情報であり，自動的に収集することが可能である．

② セキュリティ事故情報

セキュリティ事故情報は，サイバー攻撃に関わるセキュリティ事故の兆候の数である．例えば，訓練メール開封数（訓練メールの添付ファイルを開封した割合，訓練メール中の URL をクリックした割合），悪質サイト訪問数（悪質サイト検知システムで警告を受けた数），マルウェア検知数，ポリシー違反の回数など，組織の IT 管理者やセキュリティ管理者であればアクセスできる情報であり，自動的に収集することが可能なものを利用する．また潜在的に事故に関係する可能性が高い情報として，実行ファイルダウンロード数，インターネット利用数などの情報も利用しても良い．図 4 にセキュリティ事故情報のイメージを記載する．

3.2.2 収集方法

① プロファイル情報の収集

step1. 予め用意した組織のスタッフ名簿に記載されている識別子 I を 1 つずつ順番に調査する．名簿にはスタッフの氏名やメールアドレスなどの識別子 I が含まれている．識別子 I をインターネット上から探し，識別子 I を含むページの情報から，インターネットサービスの使用頻度，公開されている個人情報の量，個人の興味に関する情報などを収集し，得られた情報をプロファイル情報 DB に登録する．

step2. 識別子 I を組織内のシステム（例えば イントラネット）から探し，得られた情報をプロファイル情報 DB に登録する．例えば，識別子 I に関連する部署，上司，部下，スケジュールなどを情報として収集する．収集したプロファイル情報は以下のような多次元のベクトルで管理される．

$$p_{ij} \in ProfileInfoDB \quad (1 \le i \le N, 1 \le j \le P)$$
Nはサンプルの数、Pは特徴の種類

インターネットサービスの使用頻度としては，例えば，一定期間中にソーシャルネットワークサービスでコメントをポストしている回数を利用することができる．公開されている個人情報の量としては，例えば，ソーシャルネットワークサービスで公開している自らの個人情報の種類（例えば，氏名，知人関係，組織名，連絡先，住所などに関する情報）を利用することができる．趣味などの個人の興味に関する情報としては，例えば，ソーシャルネットワークポストしている記事の内容や識別子 I を含むページから，Bag of Words や TF･IDF などの自然言語処理技術を活用し，出現頻度の高い単語や重要な意味を持つ単語をピックアップすることで抽出することができる．

また同じページに別の人物の情報（識別子 I’）も記載されている場合，識別子 I と識別子 I’は関係性があるとし，知人関係として情報を抽出する．

② セキュリティ事故情報の収集

step1. スタッフ名簿に記載されている識別子 I を組織内のセキュリティ事故に関するログ（訓練メール開封数，悪質サイト訪問数，マルウェア検知数，ポリシー違反の回数など）から探し，得られた情報をセキュリティ事故情報 DB に登録する．実行ファイルダウンロード数やインターネット利用数などの潜在的に事故に関係する情報も各種ログから収集する．ログは，組織の IT 管理者やセキュリティ管理者であればアクセス可能な情報であるとする．

step2. 収集したセキュリティ事故情報は以下のような多次元のベクトルで管理される．

$$s_{ik} \in SecurityAccidentInfoDB \quad (1 \le i \le N, 1 \le k \le S)$$
Nはサンプルの数、Sは特徴の種類

組織プロファイル情報　　公開プロファイル情報

	年齢	所属	…	上司	…	部下	…	予定A	…	氏名	所属	連絡先	住所	友人名	…	興味	…	趣味	…
p_1																			
p_2																			
…																			
p_i	p_{i1}	p_{i2}	…	…	…	…	…	p_{ij}	…	…	…	…	…	…	…	…	…	…	…
…																			

図 3　プロファイル情報

セキュリティ事故情報

潜在的な事故情報（リスク低）

	訓練メール開封数	…	マルウェア検知数	…	悪質サイト訪問数	…	ポリシー違反数	…	実行ファイルダウンロード数	…	ファイルダウンロード数	…	インターネット利用数	…
s_1														
s_2														
…														
s_i	s_{i1}	s_{i2}	…	…	…	…	s_{ik}	…	…	…	…	…	…	…
…														

図 4　セキュリティ事故情報

3.3 関係性のモデルの作成

プロファイル情報とセキュリティ事故情報との関係性のモデルを作成する．モデルは，プロファイル情報にどういう傾向を持つ人物がどういうセキュリティ事故を起こしやすいかという関係性を表している．事前にプロファイル情報とセキュリティ事故情報の相関を求め，無相関な項目を除外しても良い．

3.3.1 学習用データのクラスタリング

step1. プロファイル情報 $p_j\ (1 \le j \le P)$ の各特徴 とセキュリティ事故情報の各特徴 $s_k\ (1 \le k \le S)$ との相関をとる．

相関係数 $corr_{jk}(1 \le j \le P, 1 \le k \le S)$ は以下の式により計算する．

$$corr_{jk} = \frac{\sigma_{p_j s_k}}{\sigma_{p_j}\sigma_{s_k}}$$

$\sigma_{p_j s_k}$: p_j と s_k の共分散

σ_{p_j} : p_j の標準偏差

σ_{s_k} : s_k の標準偏差

$p_j\ (1 \le j \le P)$: プロファイル情報の j 種類目の特徴列（ベクトルの次元数は N）

$s_k\ (1 \le k \le S)$: セキュリティ事故情報の k 種類目の特徴列（ベクトルの次元数は N）

step2. セキュリティ事故情報のどの特徴とも相関係数の絶対値が閾値以下であるプロファイル情報の特徴（$p_j : \forall k(|corr_{jk}| \le \theta_{c1})$）を除外し，セキュリティ事故情報と相関のあるプロファイル情報を作成する． プロファイル情報は以下の多次元のベクトルで表現される．閾値 θ_{c1} はあらかじめ決めておく．

$p'_{ij} \in ProfileInfoDB' \quad (1 \le i \le N, 1 \le j \le P')$
Nはサンプルの数、P'は特徴の種類

step3. 同様に，プロファイル情報のどの特徴とも相関係数の絶対値が閾値以下であるセキュリティ事故情報の特徴（$s_k : \forall j\left(\left|corr_{jk}\right| \le \theta_{c2}\right)$）を除外し，プロファイル情報と相関のあるセキュリティ事故情報を作成する．セキュリティ事故情報は以下の多次元のベクトルで表現される．閾値θ_{c2}はあらかじめ決めておく．

$s'_{ik} \in SecurityAccidentInfoDB' \quad (1 \le i \le N, 1 \le k \le S')$
Nはサンプルの数、S'は特徴の種類

step4. $ProfileInfoDB'$と$SecurityAccidentInfoDB'$のサンプルを，$ProfileInfoDB'$の特徴情報をもとに，クラスタリングを行い，N個のサンプルを，C個のクラスタ$c_l \in Clusters \quad (1 \le l \le C)$に分類する．各々のクラスタは，クラスタリングされたサンプルのプロファイル情報とセキュリティ事故情報のペア$c_l = \{(p_i, s_i) \mid i \in CI_l\}$の集合で表現される．ここで$p_i$は$P'$種類の特徴情報からなるベクトルである．$s_i$は$S'$種類の特徴情報からなるベクトルである．また$CI_l$は$c_l$にクラスタリングされたサンプルのインデックスの集合である．クラスタリングのアルゴリズムは K-means 法などの一般的なものや独自のアルゴリズムを利用可能である．

step5.

3.3.2 学習用データのラベル付

step1. クラスタ$c_l = \{(p_i, s_i) \mid i \in CI_l\}$における，セキュリティ事故情報のそれぞれの特徴の平均$SecAccInfoAve(c_l)$と標準偏差$SecAccInfoStdv(c_l)$を求める．$|CI_l|$はクラスタ にクラスタリングされたサンプル数を表す．

$$SecAccInfoAve(c_l) = (ave(s_1), ave(s_2), \cdots, ave(s_k), \cdots, ave(s_{S'}))$$
$$ave(s_k) = \frac{\sum_{i \in CI_l} s_{ik}}{|CI_l|}$$
$$SecAccInfoStdv(c_l) = (stdv(s_1), stdv(s_2), \cdots, stdv(s_k), \cdots, stdv(s_{S'}))$$
$$stdv(s_k) = \sqrt{\frac{\sum_{i \in CI_l} (s_{ik} - ave(s_k))^2}{|CI_l|}}$$

step2. 平均$SecAccInfoAve(c_l)$と標準偏差$SecAccInfoStdv(c_l)$をもとに，クラスタを表現するラベル$LABEL(c_l)$を作成する．$LABEL(c_l)$は以下のように定義する．

$$LABEL(c_l) = (label(s_1), label(s_s), \cdots, label(s_k), \cdots, label(s_{S'}))$$
$$label(s_k) = \begin{cases} ave(s_k) & (\theta_{k1} \le stdv(s_k) \le \theta_{k2}) \\ None & (otherwise) \end{cases}$$

kはセキュリティ事故情報の特徴のインデックスである．θ_{k1}およびθ_{k2}は，セキュリティ事故情報の特徴ごとにあらかじめ定義した範囲の境界値を示す．

3.3.3 モデルの学習

作成した全てのクラスタ$c_l = \{(p_i, s_i) \mid i \in CI_l\}$に対して，プロファイル情報$p_i$を学習データ，$LABEL(c_l)$をラベルとして識別器の学習を行う．学習時に扱いやすいよう$LABEL(c_l)$に適当な定数を割り当てておく．

3.4 セキュリティ意識の推定

step1. 3.2.2 と同様に，セキュリティ意識を評価したい対象の人物のプロファイル情報を収集する．

step2. 次に 3.3.1 で除外した特徴を，プロファイル情報から除外する．

step3. 得られたプロファイル情報を 3.3.3 で学習した識別器に入力し，識別器が推定するラベル$LABEL(c_l)$を取得する．

step4. $LABEL(c_l)$から，人物 A が起こしやすいセキュリティ事故を表示する．$LABEL(c_l)$を構成する要素$label(s_k)$が規定の閾値以上（非 None）の場合，人物 A はs_kに関するセキュリティ事故を起こしやすいと定義する．セキュリティ事故情報の特徴ごとあらかじめϑ_{k1}とϑ_{k2}を決めておく．

$\begin{cases} s_k\text{に関する事故を起こしやすい} & (llabel(s_k) \ge \vartheta_{k1}) \\ s_k\text{に関する事故を起こしにくい} & (label(s_k) \le \vartheta_{k2}) \\ s_k\text{に関しては特になし} & (otherwise) \end{cases}$

4. 考察

提案方式により，組織のスタッフごとどういうセキュリティ事故を起こしやすいかを推定し，推定した結果をもとに，スタッフごと適した対策を導入することができると期待される．プロファイル情報とセキュリティ事故情報との関係性のモデルが，統計的にどの程度妥当なものになるかについては，実データを用いた検証が今後必要である．

またセキュリティ管理者があらかじめセキュリティ事故ごとに検討し用意したセキュリティ対策（例えば，表 1 や表 2）の DB（対策 DB）を使うことで，対策の選定についても自動化をすることができると考えられる．

表 1　不正な添付ファイル開封による被害の対策例

対策番号	対策案
1	ウィルス対策ソフトの導入
2	標的型攻撃メール訓練の実施
3	Windows の AppLocker を導入

表 2　悪質サイト訪問による被害の対策例

対策番号	対策案
1	ウィルス対策ソフト
2	標的型攻撃メール訓練の実施
3	URL フィルタ
4	ブラウザやその他アプリケーションのアップデート

なお，これまでの説明では，全てのスタッフに対して一定の方法でセキュリティ事故情報を収集することを前提としていた．しかしこの方法では，セキュリティ事故情報を適切に収集できない可能性がある．例えば，訓練メール開封数に関して言えば，訓練メールのコンテンツによって開封しやすい人もいればそうではない人もいるだろう．そこで，スタッフのプロファイル情報をもとに，訓練メールのコンテンツを変えてやることで，より効果的にセキュリティ事故情報を収集することも可能になると考える．スタッフごとにセキュリティ事故情報の収集を変更するイメージを図 5 に記載する．

これを実現するには，訓練メールのコンテンツを管理するための DB（訓練メールコンテンツ DB）が必要となる．訓練メールコンテンツ DB には，コンテンツのトピック（ニュース，趣味，仕事，転職など）ごとにいくつかの内容の訓練メールが用意される．例えばトピックがニュースの場合，訓練メールは，経済，国際，国内，エンタメなどに関係する内容となる．スタッフのプロファイル情報（特に，仕事の情報や興味）に関係する，訓練メールをトピックごと取得し，訓練メールのデータセットを作成する．

データセットにある訓練メールを定期的にスタッフに送信し，トピックごとの訓練メール開封数をセキュリティ事故情報に登録していく．訓練メールの送信については，例えば，「標的型メール訓練サービス[7]」などの既存技術や既存サービスを利用することができる．

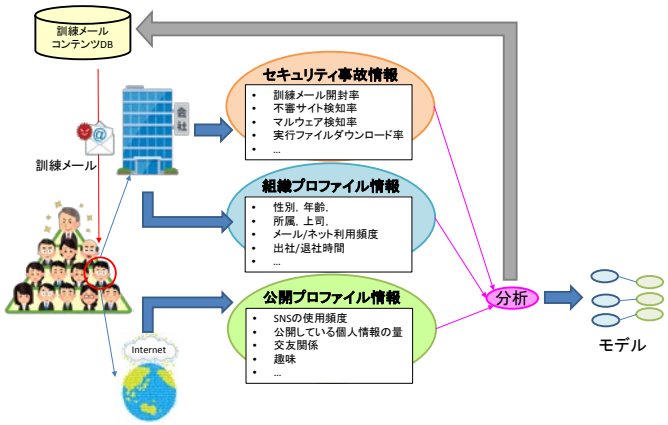

図 5　スタッフごとにセキュリティ事故情報の収集を変更する場合のモデルの作成

5.　おわりに

本稿では，スタッフのプロファイル情報とセキュリティ事故情報とを自動的に収集・分析し，セキュリティ事故を起こしやすいスタッフに共通の人物モデルを作ることで，スタッフごとにセキュリティ意識を評価する仕組みを提案した．

提案方式により，組織のスタッフごとどういうセキュリティ事故を起こしやすいかを推定し，推定した結果をもとに，スタッフごと適した対策を導入することができると期待される．

今後は，実データを用いた検証実験を行い，プロファイル情報とセキュリティ事故情報との関係性のモデルが，統計的にどの程度妥当なものかを確認していく．

参考文献

[1] Verizon Business, 2008 Data Breach Investigations Report, http://www.verizonenterprise.com/resources/security/databreachreport.pdf（2017 年 4 月 20 日確認）

[2] 中澤 優美子，加藤 岳久，漁田 武雄，山田 文康，山本 匠，西垣 正勝，Best Match Security－性格と本人認証技術のセキュリティ意識との相関に関する検討－，情報処理学会研究報告，Vol.2010-CSEC-48，No.21.

[3] 片山 佳則，寺田 剛陽，鳥居 悟，津田 宏，ユーザー行動特性分析による個人と組織の IT リスク見える化の試み，SCIS2015 暗号と情報セキュリティシンポジウム，4D1-3

[4] 守屋 潤一，孫博，森 達哉，後藤 滋樹，標的型攻撃の被害者となる人を予測することは可能か？SCIS2017 暗号と情報セキュリティシンポジウム，1F1-2.

[5] Edge-security, theHarvester, http://www.edge-security.com/index.php（2017 年 4 月 20 日確認）

[6] Paterva, Maltego CE ，https://www.paterva.com/web7/buy/maltego-clients/maltego-ce.php（2017 年 4 月 20 日確認）

[7] NTT ソフトウェア，標的型メール訓練サービス，https://www.ntts.co.jp/products/apttraining/index.html（2017 年 4 月 20 日確認）