分散表現を用いたネットワーク通信ログのアノマリ検知

江田 智尊[1,a)] 神原 佑輔[1] 及川 孝徳[1] 古川 和快[1] 海野 由紀[1] 村上 雅彦[1]

概要：自然言語処理技術の一つである分散表現（単語埋め込み）技術は様々なセキュリティログ分析に応用されている．その一般的な分析方法は，単語埋め込みを利用して分析対象の特徴量を抽出し，抽出した特徴量に目的の分析を行う機械学習アルゴリズムを適用する．しかしこの過程で抽出した特徴量は必ずしも目的の分析に適応する保証はなく，分析精度劣化を引き起こす可能性がある．本稿ではネットワーク通信ログにおけるアノマリ IP アドレス検知に着目し，アノマリ検知に適した IP アドレス特徴量を抽出する分散表現技術を提案する．提案手法は，単語埋め込みによる特徴抽出とアノマリ検知の目的関数を調和し，最適化する手法である．これにより検知に適応した IP アドレス特徴量抽出を可能にし，ログに潜むアノマリ IP アドレスを高精度に検知することを実現する．実験では，仮想的な組織ネットワークにおける通信ログの中から，攻撃者の IP アドレスを検知するタスクを実行した．結果，提案手法は従来手法と比較して Area Under the Curve 基準を 0.876 から 0.990 に改善した．

Anomalous IP Address Detection on Traffic Logs Using Novel Word Embedding

Satoru Koda[1,a)] Yusuke Kambara[1] Takanori Oikawa[1] Kazuyoshi Furukawa[1] Yuki Unno[1] Masahiko Murakami[1]

1. はじめに

研究背景：組織が導入するセキュリティ監視機器数の増加に伴い，セキュリティログの容量と SOC (Security Operation Center) オペレータのログ分析業務負担は増大している．このような状況下では，ログの対処優先度を算出するログ分析ツールが必須である．この実現のため，近年では機械学習 (Machine Learning, ML) や自然言語処理 (Natural Language Processing, NLP) 技術を用いたログ分析技術が盛んに研究されている．

本稿も同様に，ML 及び NLP に基づく通信ログ分析に着目する．本稿においては，通信ログは主にフローベース通信ログとアラートログを指す．フローベース通信ログは，ネットワークを流れるパケットの中から共通のプロパティ（送信元/宛先 IP アドレスやポート）を持つパケットを単位時間ごとに集約したログを指す．アラートログは，IDS/IPS (Intrusion Detection/Protection Systems) などのセキュリティ検知機器があげるアラートを指す．本稿が扱う分析は，これらの通信ログに含まれる全ての IP アドレスの中から，振る舞いが最も異常な IP アドレス（アノマリ IP アドレス）を検知することを目的とする．

先行研究と課題：ML や NLP を用いたログ分析技術は，本稿の指す通信ログに限らず，プロキシログ，HTTP ログ，イベントログなど様々なログに対し開発されている．特に NLP における分散表現（単語埋め込み）手法の一つである Word2Vec とその派生手法は，セキュリティログ分析においても近年よく用いられる技術であり，その有効性が度々実証されている [1–11]．Word2Vec に基づくログ分析は基本的に以下のように実装される．まず第一ステップとして，IP アドレスや各 HTTP ログレコードといった分析対象の固定長の実数値特徴量ベクトルを Word2Vec で抽出する．第二ステップで，抽出した特徴量を用いて，目的の分析（判別，アノマリ検知，クラスタリング等）を行う ML アルゴリズムを適用する．しかしこの一般的な分析プロセスは，特徴量が目的の分析に必ずしも適合しないとい

[1] 株式会社 富士通研究所 セキュリティ研究所
Security Laboratory, Fujitsu Laboratories Ltd.
211-8588 川崎市中原区上小田中 4-1-1
[a)] koda.satoru@fujitsu.com

う相性の問題を生じさせる．例えば第二ステップで用いるアルゴリズムに依存して分析パフォーマンスが著しく変化したり，さらに悪い場合には特徴量が目的の分析自体に適合しないといった問題を生じたりする．著者らは特徴量と分析アルゴリズム間にこのような相性問題が実際に存在することを実験的に確認した．

提案手法：上記問題を受け，本稿では Word2Vec に基づく効率的な特徴量抽出手法を提案する．さらに本手法を通信ログにおけるアノマリ IP アドレス検知に適用し，アノマリ検知に適合した IP アドレス特徴量を抽出することを試みた．提案手法は，Word2Vec に基づく特徴抽出と ML アノマリ検知の目的関数最小化を同時に実行する．従来手法はこれらを独立に実行する点で異なる．本手法の主要素技術は，特徴抽出のための Word2Vec と，ML ベースアノマリ検知手法の Support Vector Data Description (SVDD) から成る．さらに，非線形構造をもつアノマリ検知に対処するため，random Fourier feature 技術を採用する．これらの技術を調和し新たな目的関数を定式化する．これにより，アノマリ IP アドレス検知の精度を向上するような IP アドレス特徴量の抽出を実現する．数値実験では，仮想的な組織ネットワークにおけるフローベース通信ログの中から，攻撃者の IP アドレスを検知する実験を実行した．結果，Area Under the Curve (AUC) 評価基準において，提案手法は従来手法と比較して AUC を 0.876 から 0.990 に改善した．

貢献：本研究の貢献を以下のようにまとめる．

- 数値実験を通して，特徴抽出とアノマリ検知が独立に実行された場合に，これらの間に精度に大きく作用する相性の問題が生じることを実証した．
- 上記相性の問題を解決する，Word2Vec に基づく IP アドレス特徴量抽出手法を提案し，数値実験でその有効性を示した．

論文構成：本稿は以下の章で構成される．2 章は提案手法に関する基礎事項を記述する．3 章は提案手法を記述する．4 章は数値実験の結果を記述する．5 章は結論を述べる．

2. 準備

本章では分析の目的，単語埋め込みの基礎事項，先行研究について記述する．

2.1 分析の目的

まず初めに本稿が扱う分析の目的を述べる．目的はネットワーク通信ログの中からアノマリ IP アドレスを高精度に検知することである．アノマリ IP アドレスとは，良性 IP アドレスに対し異なる振る舞いで通信を行った IP アドレスを指す．例えば攻撃者や偵察者，内部不正者などが用いる IP アドレスを想定する．分析適用例として，ある組織ネットワークで従業員が業務で利用するような良性 IP アドレスに潜み不審な行動をする外部攻撃者や内部不正者が用いる IP アドレスを検知する状況を想定する．目的実現のため，我々は従来より高精度にアノマリ IP アドレスを検知する手法の開発を試みる．本手法により，SOC オペレータによる不審ログ・IP アドレスの絞り込みといったログ分析業務を支援し，サイバーインシデントの早期検知を実現する．

2.2 単語埋め込み，Word2Vec

単語埋め込みは，文章内の各単語を単語間の関係を保持しつつ実数値ベクトル（分散表現）に変換する技術であり，Word2Vec はそれを実現する代表的なアルゴリズムである [1]．Word2Vec においては，各単語は前後の単語（コンテクストと呼ぶ）との共起関係を基にベクトルに埋め込まれる．埋め込まれた単語ベクトルは単語間の本質的な関係を保持することが知られている．例えば同意語の単語ベクトルは，コサイン類似度の意味で類似度が高くなることなどが実証されている．ログ分析における Word2Vec は，IP アドレスやログレコードの一行といった分析対象の実数値特徴量の抽出に用いられる．この際のコンテクストは，それらと共起する情報（例えば IP アドレスがアクセスした宛先 IP アドレス・ポートや，ログに含まれる情報）などで定義される．次節にて具体的な Word2Vec のログ分析への先行適用事例について記述する．

2.3 先行研究

Ring et al. (2017) は，Word2Vec に基づく IP アドレス特徴量抽出の枠組みを初めて提案し，IP2Vec と名付けた [2]．著者らはコンテクストとして宛先 IP アドレス・ポートと通信プロトコルを用いて IP アドレスの特徴量抽出を行った．提案手法は IDS ログにおける IP アドレスのクラスタリングに適用され効果を実証した．Carrasco and Sicilia (2018) は Word2Vec を用いて，ネットワーク通信ログの各行を数値ベクトルで表現する手法を示した [3]．コンテクストとして送信元・宛先 IP アドレス，宛先ポート，プロトコルを用いた．攻撃検知実験において，著者らの手法は SOTA アルゴリズムを上回る精度を実現した．Cui et al. (2018) は Word2Vec を用いて，通信パケットのペイロードからパケット特徴量を抽出する分散表現技術を実装した [4]．ペイロードを 8bit 単位でトークン化して文字列とみなし，Word2Vec を適用した．特徴量はニューラルネットワークに入力され，攻撃検知に利用された．Bertero et al. (2017) はイベントログの各レコードの数値ベクトルを抽出するために Word2Vec を適用した [5]．ログ内の単語が Word2Vec によって分散表現され，各ログの特徴量はログ内の単語ベクトルの重心として定義された．特徴量は ML 判別器に入力され，システムのストレス状態を判別するタスクに適用された．Mimura and Tanaka (2017) は，

Word2Vec の派生技術の一つである Doc2Vec を用いて，プロキシサーバのログから悪性通信の攻撃種別を判定する手法を実装した [6]．プロキシログは Doc2Vec によって数値ベクトルに埋め込まれ，ML 判別器に入力された．Pande and Ahuja (2017) は HTTP ログにおけるアノマリを検知する検証に際し，各 HTTP ログレコードを Word2Vec により数値ベクトルに変換する手法を実装した [7]． 同じように Zhuo et al. (2017) は KDD99 データセットの各レコードを Word2Vec により特徴量に変換し，攻撃種別判定に用いた [8]．通信ログに限らないセキュリティ保守を目的とした Word2Vec 適用事例として，URL からの Web アノマリ検知事例も存在する (Yuan et al. (2017), Ming et al. (2018), Li et al. (2019)) [9–11]．著者らは Word2Vec を用いて URL シーケンスから URL 特徴量を算出した．特徴量は攻撃種別判定のためにクラスタリングアルゴリズムに入力された．

上記先行研究は様々な分析タスクにおいて最新技術を上回る精度を実現している．しかし共通の課題として，特徴抽出と目的の分析を別々のステップで実行しているため，抽出した特徴量が必ずしも目的の分析に適合しない可能性をはらんでいる．我々の先行研究では上記ステップを同時に最適化するプロトタイプ手法を提示したが，非線形構造をもつアノマリ検知に対処できないために精度が低く，実用的ではなかった [12]．本稿の提案手法はここで挙げた課題を全て解決することで，分析パフォーマンスの向上を図る．

3. 提案手法

本章では我々が提案する手法を定式化する．

3.1 提案手法概要

提案手法の概要を図 1 に示し，以下で分析の流れを説明する．本分析では通信ログが与えられた際に，ログ内のアノマリ IP アドレスを検知する状況を想定する．このために，全送信元 IP アドレスは IP2Vec を用いて固定長特徴量ベクトルに変換される．同時に，特徴量はアノマリ検知アルゴリズム SVDD に適合した特徴量へと変換される．この変換過程では，非線形なアノマリ検知に対処するために，random Fourier feature 技術を適用し，埋め込まれた特徴量を別の高次元空間に射影する．この変換により射影後のベクトルが SVDD に適合するような IP アドレス特徴量が成型される．提案手法の最適化後，各 IP アドレスに対してアノマリスコアを算出しアノマリを決定する．

3.2 記号

N 個の自然数で表記される送信元 IP アドレスを含む通信ログが与えられているとする．$i\,(\in \{1,...N\})$ 番目の送信元 IP アドレスを $w_i\,(=i)$ と表記する．w_i のコンテクストを $C(w_i)$ と表記し，w_i と共起する情報として定義する．例えば通信ログの場合，w_i がアクセスした宛先 IP アドレスやポート，用いた通信プロトコルを含む．アラートログの場合は検知用のシグネチャ ID なども含む．コンテクスト $C(w_i)$ の要素を $\{w_{c,i}\}_c$ と表記する．コンテクストの各要素は，取り得る全コンテクスト要素の集合 $\mathcal{C}$ 間でユニークな自然数で表記されているとする．

3.3 IP2Vec

IP アドレスの特徴量を抽出するために，Word2Vec に基づく分散表現技術である IP2Vec を適用する [2]．その目的関数は以下の通り与えられる：

$$L_{i2v}(U,V) = -\sum_{i=1}^{N} \sum_{w_{c,i}\in C(w_i)} \log P(w_{c,i}|w_i). \qquad (1)$$

ここで $P(w_{c,i}|w_i)$ は

$$P(w_{c,i}|w_i) = \frac{\exp(\boldsymbol{v}_{w_{c,i}}^T \boldsymbol{u}_i)}{\sum_{k\in\mathcal{C}} \exp(\boldsymbol{v}_k^T \boldsymbol{u}_i)} \qquad (2)$$

で与えられる．ベクトル $\boldsymbol{u}_i\,(=\boldsymbol{u}_{w_i}) \in \mathbb{R}^d$ は w_i の特徴量ベクトルに相当する d 次元のベクトルであり，ベクトル $\boldsymbol{v}_{w_{c,i}} \in \mathbb{R}^d$ は，コンテクストのある要素 $w_{c,i}$ に対応するベクトルである．行列 U, V は各列にそれぞれ $\{\boldsymbol{u}_i\}_{i=1}^N$, $\{\boldsymbol{v}_{w_{c,i}}\}_{w_{c,i}}$ を並べた行列である．目的関数 (1) の最小化は，IP アドレス w_i がそのコンテクスト $C(w_i)$ と共起する尤度を最大化することに対応する．最適化後のベクトル $\boldsymbol{u}_i\ (i=1,...N)$ が IP アドレス w_i の特徴量となる．

3.4 Random Fourier Features

実用面では，複雑な（非線形な）検知にも対応可能なアノマリ検知アルゴリズムが望ましい．例えばカーネルトリックを用いた support vector machine (SVM) は非線形なアノマリ検知を容易に実現するため実用によく用いられる．しかし IP2Vec はミニバッチ形式で最適化されるため，カーネルトリックを直接的に適用することが容易ではない．つまり単純には IP2Vec とアノマリ検知アルゴリズムの目的関数を調和することができない．そこで提案手法は，IP2Vec が抽出する IP アドレス特徴量を更に高次元空間に射影し，射影後の特徴量間の内積が以下のガウスカーネル

$$K(\boldsymbol{u}_i, \boldsymbol{u}_j) = \exp\left(-\frac{\gamma}{2}||\boldsymbol{u}_i - \boldsymbol{u}_j||_2^2\right) \qquad (3)$$

を近似するような関数 $\phi(\cdot) : \mathbb{R}^d \to \mathcal{H}$ を明示的に定義することを試みる．これは実際に Rahimi and Recht (2008) で提案された Random Kitchen Sinks と呼ばれる技術で実現される [13]. 本技術は Bochner の定理に基づき，これを実現する射影を以下のように構築する：

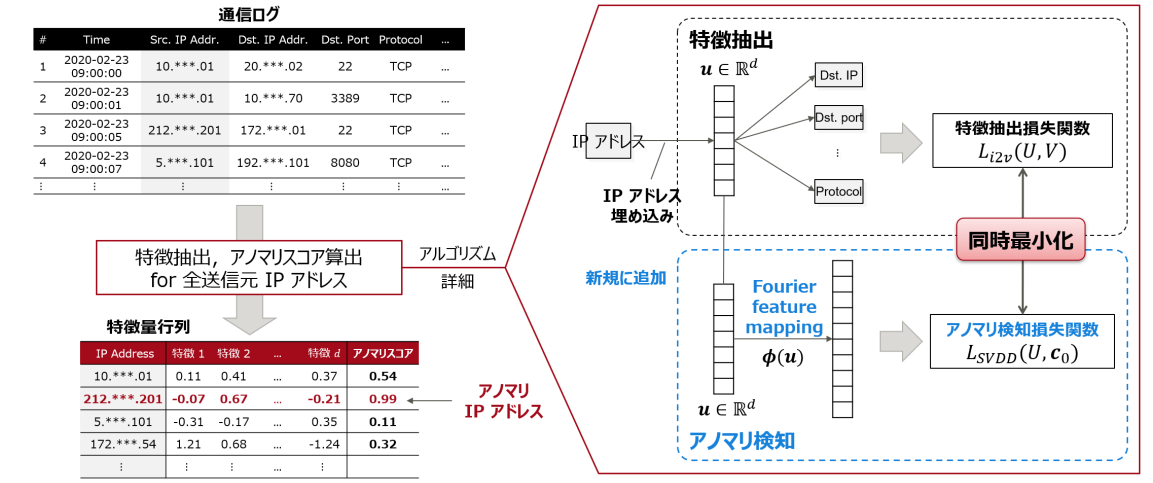


図 1: 提案アノマリ検知手法概要

$$\phi(\boldsymbol{u}) = \frac{1}{\sqrt{L}}\big[\cos(\boldsymbol{\omega}_1^T\boldsymbol{u}),\cdots,\cos(\boldsymbol{\omega}_L^T\boldsymbol{u}), \sin(\boldsymbol{\omega}_1^T\boldsymbol{u}),\cdots,\sin(\boldsymbol{\omega}_L^T\boldsymbol{u})\big]^T. \quad (4)$$

ここで，$\boldsymbol{\omega}_1,\cdots,\boldsymbol{\omega}_L \in \mathbb{R}^d$ はガウス分布 $N(0,\gamma I_d)$ から独立同分布にサンプルされた定数であり，L は random Fourier feature の個数を表す．この関数による射影で，特徴量射影後の空間での内積はガウスカーネルを以下のように近似する：

$$K(\boldsymbol{u}_i,\boldsymbol{u}_j) \approx \phi(\boldsymbol{u}_i)^T\phi(\boldsymbol{u}_j). \quad (5)$$

以上の操作により，射影後の空間で線形アノマリ検知を実行することで，射影元の空間でガウスカーネルを用いたような非線形アノマリ検知を実現する．

3.5 Support Vector Data Description

続いて，射影後の IP アドレス特徴量セット $\{\phi(\boldsymbol{u}_i)\}_{i=1}^N$ は，ML ベースのアノマリ検知アルゴリズムである SVDD に入力される [14]．SVDD の基本的なアイデアは，与えられたサンプルを全て包含するような最小の超球面を求める問題に帰着される．超球面の中心から離れて分布するサンプルがアノマリとなる仕組みである．ここでは Ruff et al. (2018) を基に，元の線形 SVDD を簡略化した以下の目的関数を用いる [15]：

$$L_{SVDD}(U,\boldsymbol{c}_0) = \sum_{i=1}^{N}||\phi(\boldsymbol{u}_i)-\boldsymbol{c}_0||_2^2. \quad (6)$$

上記目的関数の最小化は，IP アドレス $\{w_i\}_{i=1}^N$ の特徴量 $\{\boldsymbol{u}_i\}_{i=1}^N$ が，射影後の空間で超球面の中心 $\boldsymbol{c}_0 \in \mathbb{R}^{2L}$ 近くに分布するようにパラメータ U を学習する．このため，振る舞いに共通点のある大多数の良性 IP アドレスの特徴量は射影後の空間で中心付近に分布する一方で，少数のアノマリ IP アドレスの特徴量は中心から離れた点に分布するようになる．これにより良性とアノマリの分離が可能になる．

3.6 最適化

上記の議論から，提案手法における目的関数を以下のように定義する：

$$L(U,V,\boldsymbol{c}_0) = L_{i2v}(U,V) + \lambda L_{SVDD}(U,\boldsymbol{c}_0). \quad (7)$$

ここで $\lambda\,(>0)$ は二項を調整するトレードオフパラメータである．第二項が特徴抽出の際の正則化項の役割を果たす．この目的関数の最小化は，振る舞い（コンテクスト）に基づき IP アドレスの特徴量を抽出しつつ，同時に特徴量が SVDD ベースのアノマリ検知に適合するように仕向ける．更に random Fourier feature 技術により，複雑なデータにおけるアノマリ検知も可能にする．この結合により，従来技術の課題を解決する．

最適化には通常の Word2Vec と同じく一般的な確率的勾配降下法 (Stochastic Gradient Descent, SGD) を適用する．最適化中の各エポックで，まず $\boldsymbol{c}_0$ を固定して埋め込みに関するパラメータ (U,V) を更新する．エポック k 終了後，中心 $\boldsymbol{c}_0^{(k)}$ を以下のように更新する：

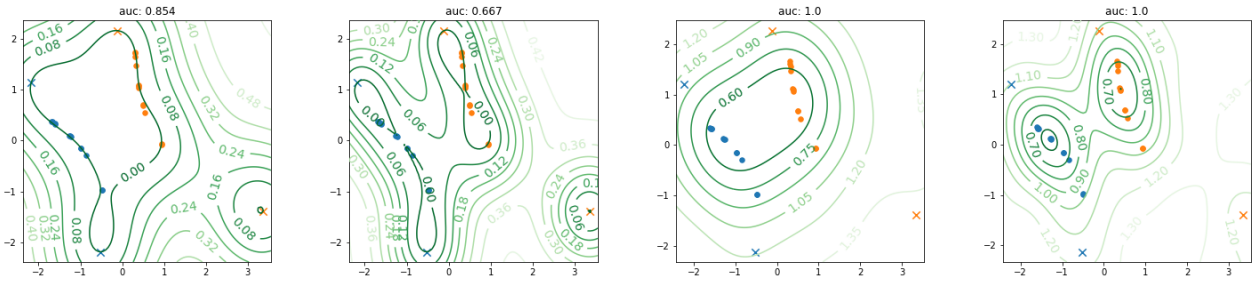
$$\boldsymbol{c}_0^{(k+1)} \leftarrow \frac{1}{N}\sum_{i=1}^{N}\phi\left(\boldsymbol{u}_i^{(k)}\right). \quad (8)$$

ここで，$\boldsymbol{c}_0^{(k)}$, $\boldsymbol{u}_i^{(k)}$ はエポック k における各パラメータ値を表す．この更新は $\frac{\partial L}{\partial \boldsymbol{c}_0} = 0$ から導かれる．

更新終了後，各 IP アドレス w_i のアノマリスコア $S(w_i)$ は以下のように算出される：

$$S(w_i) = ||\phi(\boldsymbol{u}_i)-\boldsymbol{c}_0||_2^2. \quad (9)$$

この値が大きいほどアノマリ度が高いことを表す．



(a) 従来 ($\gamma = 0.5$ AUC $= 0.85$) (b) 従来 ($\gamma = 1.0$, AUC $= 0.67$) (c) 提案 ($\gamma = 1.0$, AUC $= 1.0$) (d) 提案 ($\gamma = 1.5$, AUC $= 1.0$)

図 2: Toy ログを用いた実用例：(a,b) ガウスカーネルを用いた one-class SVM，(c,d) 提案手法． 両手法ともパラメータ γ は，ガウスカーネル内のパラメータを表し，AUC 値はアノマリ検知結果の定量的評価を示す．等高線はアノマリスコアの値によって描かれる．異なるクラスタのサンプルはそれぞれの色 (青，橙) で描かれる．マーカー 'o' と 'x' はそれぞれノーマル・アノマリ IP アドレスを表す．

3.7 Toy Example への適用例

本格的な数値実験の前に toy example で提案手法の効果を確認する．実験に用いた toy ログは付録に示す．本ログは 2 つの IP アドレスクラスタから構成される．各クラスタは 12 個のノーマル IP アドレスと 2 個のアノマリ IP アドレスを含む．提案手法の比較として，IP2Vec で抽出した特徴量にガウスカーネルを用いた one-class SVM (OCSVM) を適用する．本実験では，従来・提案手法が IP アドレスのクラスタ構造を捉えたアノマリ IP アドレス検知が可能かを視覚的に検証する．詳細な実験設定は省略する．

図 2 は両手法の検知結果を示す．可視化のため，IP アドレスの特徴量を 2 次元に埋め込み，特徴量空間のサポート上でアノマリスコアに基づく等高線を引いた．図 2 (a,b) が示すように，従来手法はクラスタ構造を捉えられず，検知結果 (AUC) も低い結果となった．この結果は，特徴抽出とアノマリ検知を独立に実行した場合，検知精度劣化に影響を与える相性の問題が生じることを例示する．一方で，図 2 (d) が示すように，提案手法は意図した通りにクラスタ構造を捉えつつ，完全な検知を実現した．等高線が各クラスタの重心付近を頂上とする二峰性の山（実際は谷）のようになっていることから，クラスタ構造を捉えたことが示唆される．また図 2(c) とその AUC が示すように，クラスタ構造を十分捉えられない場合でも，少なくとも検知と相性の良い特徴量を抽出可能なことを示した．

4. 数値実験

本実験ではオープンデータセットを用いた実験結果について記述する．

4.1 データセット

本実験では *CIDDS-001* データセットを用いた [16]. 本データセットは 2017 年に作成された，正解ラベル付きのフローベース通信ログデータセットである．通信ログは小規模組織のネットワーク環境を十分に模した仮想環境で採取された．本実験では，攻撃者によるログを含む External Server week2-week4 データセットを週毎に利用した．本データセットには PortScan 攻撃と BruteForce 攻撃に用いられた IP アドレスが含まれる．本実験の目的は，アノマリ検知アプローチが，全 IP アドレスの中から攻撃者の IP アドレスを高精度で検知可能かを検証することである．

実験の前にいくつかの前処理を施した．まずサーバと攻撃犠牲者の送信元 IP アドレスを除外した．実験で用いるコンテクスト情報として，ログの中から Dst Pt (宛先ポート), Packets (パケット総数), Duration (通信時間), Flags (TCP フラグ) のカラムを抽出した．Duration カラムは値に応じて 15 段階に量子化した．Dst Pt は 49,152 以上は同一とみなした．異なる攻撃キャンペーンに用いられた IP アドレスは，例え同一のものでも別々の IP アドレスとみなした（例えば IP1 が 2 度攻撃した場合，IP1_1, IP1_2 と別々のものとみなした）．前処理後のデータセット統計を表 1 に記す．

4.2 実験設計

4.2.1 評価基準

実験評価に 2 つの評価基準を用いた．一方は AUC で，もう一方は偽陽性率 5%点での precision (PRC) を用いた．PRC は実用面で重要な偽陽性が低い領域での検知精度を定量評価するために用いる．

4.2.2 比較手法とハイパーパラメータ設計

4.2.2.1 比較手法

比較のため，特徴抽出とアノマリ検知を独立に実行する手法を適用した．特徴抽出には IP2Vec (Ring et al. 2017) を適用し，アノマリ検知には 2 つの ML アノマリ検知アルゴリズム (OCSVM と LOF (Local Outlier Factor)) を適用

表 1: データセット統計

Dataset	week2		week3		week4	
	normal	attack	normal	attack	normal	attack
# of logs	56,172	2,389	55,405	9,225	56,788	616
# of IP addresses	169	6 (2, 4)*	167	12 (5, 7)*	159	4 (1, 3)*
# of logs min.	10	84	10	100	10	20
# of logs med.	23	100	28	100	28	60
# of logs max.	11,840	1,004	13,870	3,275	12,739	476

* 括弧内の数字はそれぞれ PortScan，BruteForce 攻撃の回数を記す．

表 2: 実験結果

	week2		week3		week4	
Method	AUC	PRC	AUC	PRC	AUC	PRC
IP2Vec + SVM	0.876	0.00	0.890	0.250	0.728	**0.500**
IP2Vec + LOF	0.870	0.333	0.819	0.417	0.753	0.250
提案手法	**0.990**	**1.00**	**0.967**	**0.917**	**0.936**	**0.500**

した．OCSVM はガウスカーネルを用いた．ガウスカーネルのパラメータ γ（式 (3)　）は $\{2^{-10}, 2^{-9}, \ldots, 2^5\}$ の範囲で調整した．LOF の近傍パラメータ k は $\{2^0, 2^1, \ldots, 2^5\}$ の範囲で調整した．

4.2.2.2　提案手法

トレードオフパラメータ λ と，random Fourier feature 変換に用いられるガウスカーネルパラメータ γ をそれぞれ，$\{2^0, 2^1, \ldots, 2^7\}$ と $\{0.125, 0.25, 0.5, 1.0, 2.0, 3.0, 4.0, 5.0\}$ の範囲で調整した．Random feature 数 L は 200 に固定した．

両手法が共通で含むハイパーパラメータを以下のように設定した．特徴量次元のハイパーパラメータ d は 64 に固定した．勾配降下法における最適化パラメータは以下のように設定した：エポック数=30，最適化手法=SGD (学習率=0.01, モメンタム=0.9)，バッチサイズ=128.

4.3　実験結果

表 2 に実験結果を記す．総合的に提案手法が従来手法を上回る結果を得た．特に PRC においては大きな改善が確認できる．week2 データセットにおいては 5%の偽陽性で攻撃者の IP アドレスを 100%検知した．week3, week4 データセットにおいても，提案手法は AUC，PRC 両基準において，従来手法を大きく改善する結果を得た．

4.4　考察

week2 データセットの結果を参照し実験結果を考察する．

4.4.1　特徴量の質

提案手法によって抽出された IP アドレス特徴量が，従来手法の IP2Vec で抽出された特徴量のような質を保存しているかを検証するために，全 IP アドレス間のコサイン類似度を計算した．結果，PortScan に用いられた 2 つの IP アドレスは互いに最も類似度が高い結果となった．同様に BruteForce 攻撃に用いられた 4 つの IP アドレスに関しても，自身以外の BruteForce 攻撃 IP アドレスの 1 つに最も似ていることが判明した．このことから，提案手法は IP アドレスの特徴量を振る舞いベースで抽出すると同時に，検知に適した形に成型していることが推察される．この現象は他のデータセットの結果からも確認された．

4.4.2　偽陽性

week2 データセットにおいて，提案手法は 3 つの偽陽性 IP アドレスを検知した：14105_53, 17626_160，20464_25 (匿名化されている)．これらは真の攻撃者と同じように，SSH BruteForce 攻撃のようなログを記録していた（宛先ポート 22 のログが 90 回以上記録された）．真の攻撃者と偽陽性 IP アドレス間のログを分析したところ，Duration カラムには差があることが判明した．真の攻撃の Duration の中央値は 2.26 に対し，偽陽性の中央値はそれぞれ 19.83 (14105_53), 2.38 (17626_160), 4.59 (20464_25) であった．この差は真の攻撃者の判定に役立つ情報であるが，本実験では Duration の 1 から 10 の数値は，同じ値に量子化された．粗い量子化は特徴量抽出に有効に働くものの，攻撃者検知に役立つ情報を欠損してしまう恐れがあることを示した．

4.4.3　パラメータ分析

提案手法のハイパーパラメータへの感度を検証するために，λ と γ の異なるパラメータ設定で実験を行った．結果を図 3 に示す．従来手法を上回ったパラメータ領域は連続的に広がっていることがわかる．week2 データセットの結果ではその領域は幅広くはなかったが，既に述べたように適切なパラメータ設計により精度が大幅に改善されることが検証された．

5.　結論

本稿では SOC によるログ分析業務効率化に向け，NLP の分散表現技術と ML に基づくアノマリ IP アドレス検知

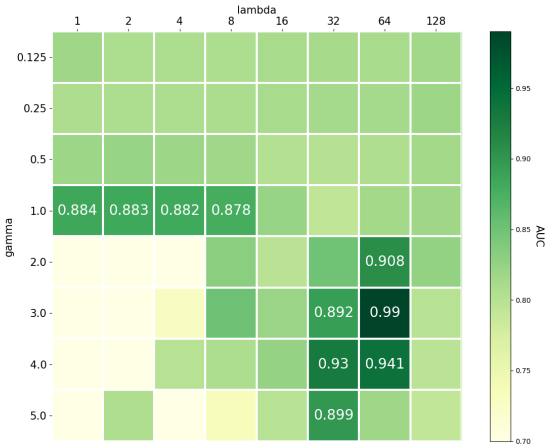


図 3: 提案手法の異なるパラメータ設定における AUC の変化．横軸・縦軸はそれぞれ λ と γ の値を記す．AUC 値が書かれた領域は，提案手法が従来手法の AUC を上回った領域を示す．

手法を提案した．数値実験では，仮想的な組織ネットワークにおけるフローベース通信ログの中から，攻撃者の IP アドレスを検知する実験を実行した．結果，AUC 基準において，提案手法は従来手法と比較して AUC を 0.876 から 0.990 に改善した．本技術により，通信ログに潜むアノマリ IP アドレスを従来より高精度に検知可能となり，SOC のログ分析業務負担軽減，インシデント早期発見に寄与することが期待される．

参考文献

[1] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, “Distributed representations of words and phrases and their compositionality,” in *Proc. of the 26th International Conference on Neural Information Processing Systems*, pp. 3111–3119, 2013.

[2] M. Ring, A. Dallmann, D. Landes, and A. Hotho, “IP2VEC: Learning similarities between ip addresses,” in *Proc. of 2017 IEEE International Conference on Data Mining Workshops*, pp. 657–666, 2017.

[3] R. S. M. Carrasco and M.-A. Sicilia, “Unsupervised intrusion detection through skip-gram models of network behavior,” *Computers and Security*, vol. 78, pp. 187 – 197, 2018.

[4] J. Cui, J. Long, E. Min, and Y. Mao, “WEDL-NIDS: improving network intrusion detection using word embedding-based deep learning method,” in *International Conference on Modeling Decisions for Artificial Intelligence*, pp. 283–295, 2018.

[5] C. Bertero, M. Roy, C. Sauvanaud, and G. Tredan, “Experience report: Log mining using natural language processing and application to anomaly detection,” in *Proc. of 2017 IEEE 28th International Symposium on Software Reliability Engineering*, pp. 351–360, 2017.

[6] M. Mimura and H. Tanaka, “Heavy log reader: learning the context of cyber attacks automatically with paragraph vector,” in *International Conference on Information Systems Security*, pp. 146–163, 2017.

[7] A. Pande and V. Ahuja, “WEAC: Word embeddings for anomaly classification from event logs,” in *2017 IEEE International Conference on Big Data*, pp. 1095–1100, 2017.

[8] X. Zhuo, J. Zhang, and S. W. Son, “Network intrusion detection using word embeddings,” in *2017 IEEE International Conference on Big Data (Big Data)*, pp. 4686–4695, 2017.

[9] G. Yuan, B. Li, Y. Yao, and S. Zhang, “A deep learning enabled subspace spectral ensemble clustering approach for web anomaly detection,” in *2017 International Joint Conference on Neural Networks*, pp. 3896–3903, 2017.

[10] E. Min, J. Long, Q. Liu, J. Cui, and W. Chen, “TR-IDS: Anomaly-based intrusion detection through text-convolutional neural network and random forest,” *Security and Communication Networks*, vol. 2018, pp. 1–9, 2018.

[11] B. Li, G. Yuan, L. Shen, R. Zhang, and Y. Yao, “Incorporating url embedding into ensemble clustering to detect web anomalies,” *Future Generation Computer Systems*, vol. 96, pp. 176–184, 2019.

[12] 江田智尊，及川孝徳，古川和快，村上雅彦 “分散表現を用いたアラートログにおけるアノマリ検知,” コンピュータセキュリティシンポジウム 2019 論文集, pp. 443–450, 2019.

[13] A. Rahimi and B. Recht, “Random features for large-scale kernel machines,” in *Advances in neural information processing systems*, pp. 1177–1184, 2008.

[14] D. M. Tax and R. P. Duin, “Support vector data description,” *Machine Learning*, vol. 54, no. 1, pp. 45–66, 2004.

[15] L. Ruff, R. Vandermeulen, N. Goernitz, L. Deecke, S. A. Siddiqui, A. Binder, E. Müller, and M. Kloft, “Deep one-class classification,” in *Proc. of the 35th International Conference on Machine Learning*, vol. 80, pp. 4393–4402, 2018.

[16] M. Ring, S. Wunderlich, D. Grudl, D. Landes, and A. Hotho, “Flow-based benchmark data sets for intrusion detection,” in *Proc. of the 16th European Conference on Cyber Warfare and Security*, pp. 361–369, 2017.

付　　録

A.1 Toy 実験で用いた Toy ログ

Toy 実験で用いた Toy ログ．太字で書かれた送信元 IP アドレスはアノマリ IP アドレスを示す．そのアノマリ要因となる情報も同様に太字で示す．

クラスタ 1 のログ

#	src. IP address	dst. IP address	dst. port	signature
1	1	201	22	A
2	2	201	22	A
3	3	201	22	A
4	4	201	139	A
5	5	201	139	A
6	6	201	139	A
7	7	201	445	A
8	8	201	445	A
9	9	201	22	A
10	9	201	139	A
11	10	201	22	A
12	10	201	445	A
13	11	201	22	A
14	11	201	139	A
15	11	201	445	A
16	12	201	22	A
17	12	201	139	A
18	12	201	445	A
19	**13**	**203**	22	A
20	**13**	**203**	139	A
21	**14**	201	**3389**	A

クラスタ 2 のログ

#	src. IP address	dst. IP address	dst. port	signature
22	101	202	22	B
23	102	202	22	B
24	103	202	22	B
25	104	202	139	B
26	105	202	139	B
27	106	202	139	B
28	107	202	445	B
29	108	202	445	B
30	109	202	22	B
31	109	202	139	B
32	110	202	22	B
33	110	202	445	B
34	111	202	22	B
35	111	202	139	B
36	111	202	445	B
37	112	202	22	B
38	112	202	139	B
39	112	202	445	B
40	**113**	**203**	22	B
41	**113**	**203**	139	B
42	**114**	202	**1**	B
43	**114**	202	**2**	B
44	**114**	202	**3**	B
45	**114**	202	**4**	B
46	**114**	202	**5**	B
47	**114**	202	**6**	B
48	**114**	202	**7**	B
49	**114**	202	**8**	B
50	**114**	202	**9**	B
51	**114**	202	**10**	B