

# Isolation Forest を用いた IoT 向け異常検知手法に関する考察

菅田 大輔<sup>1,a)</sup> 安全 花子<sup>1,2,†1</sup>

**概要：**本論文では、Isolation Forest を用いた IoT 向け異常検知手法の改善を行なった。この研究の背景に、IoT (Internet of Things) の普及がある。実際、2030 年には約 300 億台ものデバイスが利用されると予測されており、すべての IoT 機器のセキュリティ確保が重要な課題となっている。特に 2016 年の Mirai 型マルウェアによる DDoS 攻撃はその必要性を浮き彫りにした。今回は、専門的なセキュリティ対策が難しい家庭内ネットワークなどの小規模環境に着目し、軽量で高速な異常検知システムの提案を目指した。Isolation Forest は軽量で高速に動作する異常検知手法として注目されているが、これを用いた従来の方法には二つの問題が存在する。閾値を手動で設定する必要がある点と、異常判定のアルゴリズムの精度に限界がある点である。これらの課題を解決するために、本研究ではロジスティック回帰を用いた判定アルゴリズムを提案する。提案手法を二つのデータセットで実験した結果、精度がそれぞれ 84.4%から 91.1%、84.6%から 94.5%に改善したことを確認した。

**キーワード：**Isolation Forest, IoT, IDS, 異常検知

## A Study on Anomaly Detection Method for IoT using Isolation Forest

DAISUKE SUGATA<sup>1,a)</sup> HANAKO ANZEN<sup>1,2,†1</sup>

**Abstract:** abstract

**Keywords:** Isolation Forest, IoT, IDS, Anomaly Detection

### 1. はじめに

### 2. 研究方法

#### 2.1 Isolation Forest の説明

Isolation Forest (以下、iForest) は、外れ値検出のためのアルゴリズムである。iForest は異常データが少数であり、

離れているという前提に基づいている。ランダムにデータを分割していくと、異常データは相対的に早く分離される。iForest は以下のステップで実行される。

#### 1. データの分割

ランダムに選んだ特徴量から、ランダムに選んだ値をもとにデータを分割する。これを一定回数繰り返し、複数のツリーを作成する。

#### 2. 異常スコアの算出

データがツリーの枝に到達するまでの平均パスをもとに、異常スコアを算出する。具体的な計算式は以下の通りである。

ここで、 $E(h(x))$  はデータ点  $x$  の平均パス長、 $c(n)$  は

<sup>1</sup> 東京工業大学 情報理工学院 数理・計算科学系  
Department of Mathematical and Computing Sciences,  
School of Computing, Tokyo Institute of Technology

<sup>2</sup> 株式会社 YY セキュリティ研究所  
Security Laboratories, YY Corporation

<sup>†1</sup> 現在、国立研究開発法人 ZZ 研究所  
Presently with National Institute of ZZ

<sup>a)</sup> sugata.d.aa@m.titech.ac.jp

データセットのサイズ  $n$  に依存する定数である。

### 3. 異常判定

先ほど計算した異常スコアをもとに異常検知を行う。通常、トレーニングデータの異常スコアの上位 10% を閾値として設定し、それを越えたデータを異常と判定する。

#### 2.2 IDS の概要

本研究では、小規模な IoT 環境に適した Intrusion Detection System (IDS) の設計について検討する。具体的には、Isolation Forest を用いた異常検知手法が提案され、その有効性を評価する。

#### 2.3 全体の設計

IDS の設計は以下の 3 つのセクションに分けられる：

##### (1) データの前処理：

- 入力データを適切な形式に変換する。
- 不要な特徴量の削除やデータの標準化、ラベルエンコーディングを行う。

##### (2) 特徴量選択：

- 判定に重要な特徴量を選択し、過学習を防ぎ、検知精度を向上させる。
- Random Forest を用いて特徴量の重要度を算出し、重要な特徴量を選択する。

##### (3) 攻撃の判定：

- iforest を用いて通信が攻撃通信であるかを判定する。
- 特徴量の選択後、Isolation Forest で異常検知を行う。

#### 2.4 実装

実装は以下のステップで行う：

##### (1) データの前処理：

- はじめに、不要な特徴量の削除を行う。(データの標準化を行う必要なし?) Iforest に入力できるのは数値データだけなので、カテゴリカルデータを数値データに変換する。one-hot エンコーディングを用いてラベルのエンコーディングを行う。

##### (2) 特徴量選択：

- Random Forest を用いて特徴量の重要度を算出する。その後、重要度をもとに上位 1 割の特徴量を使用する。

##### (3) 攻撃の判定：

- 攻撃通信、正常通信のそれぞれでトレーニングされたサブシステムが、Isolation Forest によって異常検知を行う。
- それぞれのサブシステムの結果を 2 通りで組み合わせ、最終的な判定を行う。

### 3. iForest の問題点の整理

Isolation Forest を異常検知手法として使用する際、どのような問題があるのかを明らかにするため、事前実験を

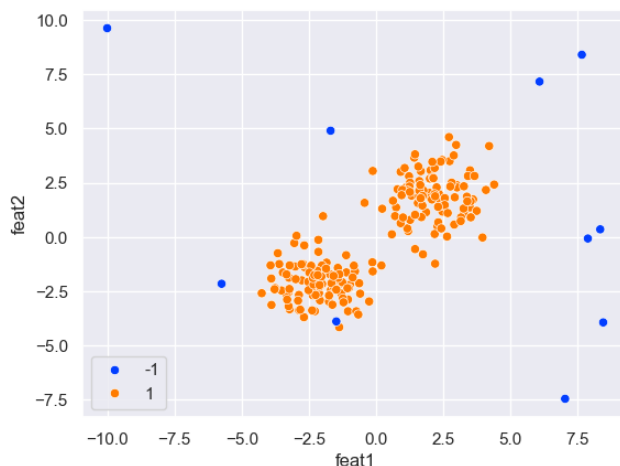


図 1 デモデータに関する図の説明 (和文)

Fig. 1 Description of the dim.vs.accu figure (English).

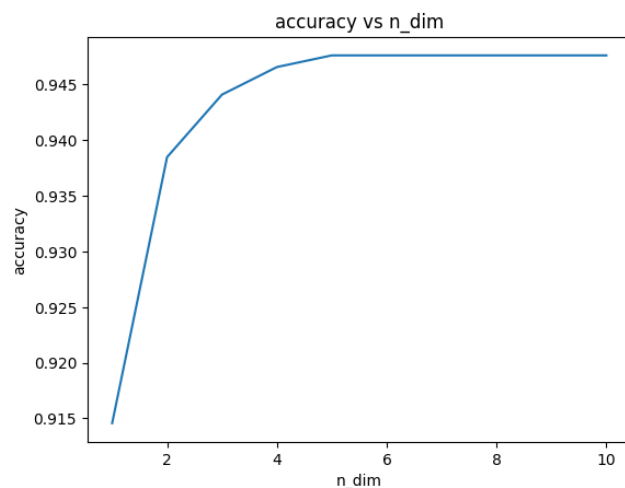


図 2 dim.vs.accu に関する図の説明 (和文)

Fig. 2 Description of the dim.vs.accu figure (English).

行った。はじめにデモデータを使用して、iForest の挙動を確認した。その考察をもとに、特徴量選択手法の提案を行った

使用したデモデータは、(2.5, 2.5) と (-2.5, -2.5) を中心とした正常データ群と、10 から -10 の範囲に一樣に分布した異常データ群からなる。二次元の場合のデモデータを図 1 に示す。

はじめに、デモデータと特徴量数の関係を調査した。図 2 に示すように、特徴量数が増えるにつれて、異常検知の精度が単調に向上することがわかった。また、精度の伸びは増加に反比例して緩やかになっていることもわかる。ゆえに、iForest は目標とする精度に対して十分な特徴量数が存在すると言える。

パケット通信を監視して得られたデータセットの全ての特徴が、異常検知に有効であるわけではない。iForest は特徴量同士の重みづけを行わないため、判定に有効でない特徴量が混ざると精度が低下すると考えられる。そこで、図

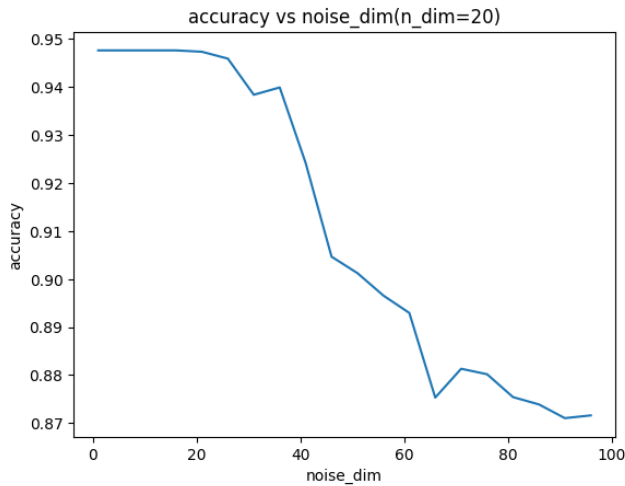


図 3 noise\_accu に関する図の説明 (和文)

Fig. 3 Description of the noise\_accu figure (English).

3に示すように、ノイズとなる特徴量を混ぜた時の精度を調査した。この結果から、ノイズとなる特徴量が混ざると精度が低下することがわかった。また、今回の実験の場合だと、ノイズとなる特徴量が判定に有効な特徴量数の2倍以上になると、精度が急激に低下することがわかった。

前の実験から、データセットからノイズとなる特徴量を取り除くことが重要であることがわかった。ところで、iForestはツリーベースの異常検知手法である。そこで、同じくツリーベースのRandom Forestから特徴の重要度を算出すれば、ノイズとなる特徴量を取り除けるのではないかと考えた。図4は、Random Forestで算出した特徴量の重要度を表している。このグラフは、ノイズ特徴量を判別できていることがわかる。そして、実際にノイズ特徴量を取り除いた場合の精度を調査したところ、精度は???%から0.945%まで向上した。この結果から、Feature Importanceによる特徴量選択手法は精度の向上に有効ではないかと考えた。

### 3.1 より効果的な判定の組み合わせ方について

Iforestを用いた異常検知手法では、攻撃データと正常データのそれぞれでトレーニングされたサブシステムが、Isolation Forestによって異常検知を行う。最終的な判定は、これらの結果を組み合わせることで行う。AbuAlghamらの研究では、表1のようにそれぞれの判定器の結果を組み合わせることで、最終的な判定を行なっている。しかし、表1のようにマッピングを行うことは2つの問題を抱えている。一つ目は、1. それぞれの判定器の異常スコアの閾値を適切に設定する必要があること、2. 異常判定が、2つの異常スコアから直線的に決定されるため、異常スコアの分布によっては適切な判定ができないことである。

そこで、本研究では、異常スコアの分布によって適切な判定ができるように、異常スコアをロジスティック回帰を

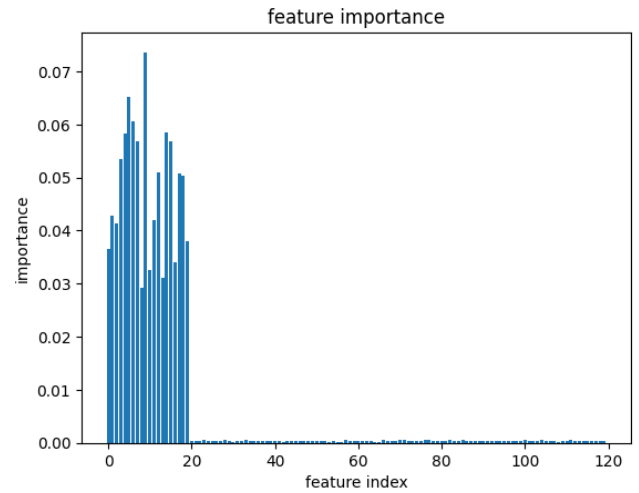


図 4 select\_noise に関する図の説明 (和文)

Fig. 4 Description of the select\_noise figure (English).

表 1 組み合わせアルゴリズム (和文)

Table 1 combination algorithm (English).

Normal subsystem	Attack subsystem	combination result
Normal	Anomaly	Normal
Anomaly	Normal	Anomaly
Normal	Normal	Anomaly
Anomaly	Anomaly	unknown

用いて判定する手法を提案する。この手法では、それぞれのサブシステムの異常スコアを入力とし、ロジスティック回帰によって異常判定を行う。この手法を用いることで、1.iForestの閾値を設定する必要がない 2. 閾値による直線的な分割よりも、より正確な分割が行えるという利点がある。

## 4. 結果と考察

### 4.1 比較するアルゴリズム

本研究では、以下の2つの比較を行った。特徴量エンジニアリングの効果の比較を行うため、特徴量選択を行わなかった場合と、Random Forestを用いた特徴量選択手法を用いた場合の結果を比較する。また、判定の組み合わせアルゴリズムの比較には、2つのサブシステムの結果をマッピングする手法と、Logistic Regressionを用いて判定を行う手法の結果を比較した。

#### (1) 特徴量エンジニアリングの比較：

- 特徴量選択なし
- Random Forestを用いた特徴量選択手法

#### (2) 判定の組み合わせアルゴリズムの比較：

- マッピングして判定
- Logistic Regressionを用いて判定

#### 4.1.1 実験環境

実験は Mac Book Pro 2017 2.3GHz Intel Core i5,

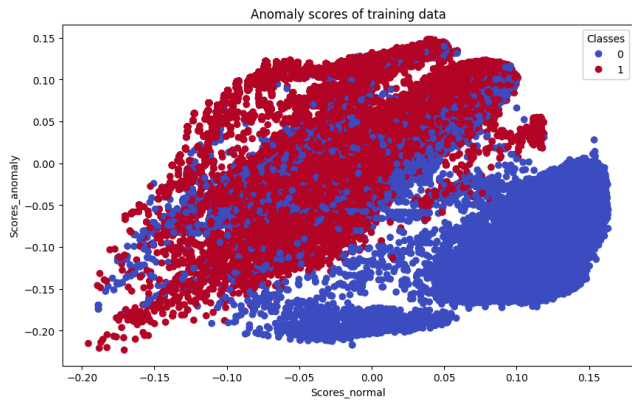


図 5 UNSW1 (和文)

Fig. 5 Description of the collectness\_UNSW figure (English).

表 2 UNSW でのモデルの性能評価結果 (和文)

Table 2 Performance evaluation results of models (English).

model	accuracy	precision	recall	f1
特徴量選択なし	0.7506	0.7949	0.7506	0.7526
特徴量選択あり	0.8440	0.9321	0.8440	0.8817
ロジスティック回帰で判定	0.9112	0.9146	0.9112	0.9082

8GB RAM で行った。また、実験に用いたプログラムは Python3.10.4 で実装した。Isolation Forest や Random Forest の実装には、scikit-learn のライブラリを用いた。

#### 4.1.2 データセット

本実験では、以下の 2 つのデータセットを用いた。

- (1) **NSL-KDD**: KDDCUP99 の問題点を解決するために提案されたデータセットであり、データの冗長性や攻撃データの割合を調整したもの。
- (2) **UNSW-NB15**: 既存のデータセットの問題点を解決し、現代のネットワークトラフィックと低フットプリント攻撃を包括的に反映するために作成されたデータセット。

#### 4.1.3 評価指標

本研究では、以下の 4 つの評価指標を用いてモデルの性能を評価した。

- (1) **Accuracy**: 正確度を示し、予測が実際のクラスと一致する割合を示す。
- (2) **Precision**: 予測が正常と判定されたデータのうち、実際に正常であるデータの割合を示す。
- (3) **Recall**: 実際に正常であるデータのうち、正常と判定されたデータの割合を示す。
- (4) **F1-score**: Precision と Recall の調和平均を示す。

#### 4.2 結果

初めに、UNSW と NSL のデータセットに対して、トレーニングデータの異常スコアの分布を調査した。

結果は以下の通りである。

この結果から、2 つのデータセットにおいて、特徴量選

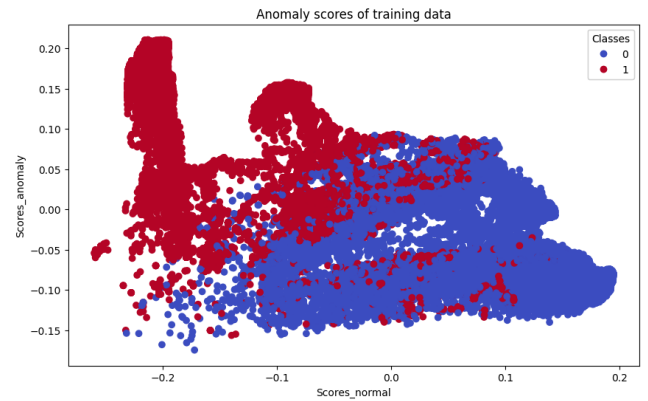


図 6 NSL1 (和文)

Fig. 6 Description of the collectness\_UNSW figure (English).

表 3 NSL でのモデルの性能評価結果 (和文)

Table 3 Performance evaluation results of models (English).

model	accuracy	precision	recall	f1
特徴量選択なし	0.7873	0.8749	0.7873	0.8171
特徴量選択あり	0.8466	0.9174	0.8466	0.8775
ロジスティック回帰で判定	0.9452	0.9453	0.9452	0.9452

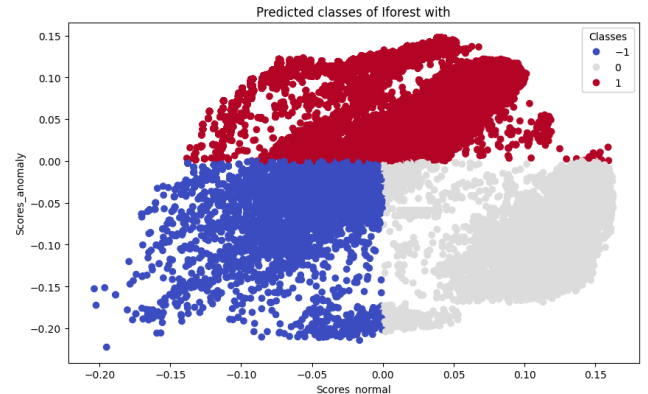


図 7 UNSW2 (和文)

Fig. 7 Description of the collectness\_UNSW figure (English).

択を行うことで精度が向上することがわかった。また、ロジスティック回帰を用いて判定を行うことで、精度が向上することがわかった。

#### 4.3 考察

今回の実験から、特徴量選択を行うことで精度が向上することがわかった。また、ロジスティック回帰を用いて判定を行うことで精度が向上させられることがわかった。これは、今回使用した 2 つのデータセットの異常スコアの分布の境界線が斜めであり、線形な境界線でうまく分割できたためであると考えられる。境界線が曲面であったり、データセットが

混ざってしまっている場合、うまくいかない→なんでこれを iforeset を使っているのか→判定の速さが

plusAlpha でかければいいことモデル構築のコストの比

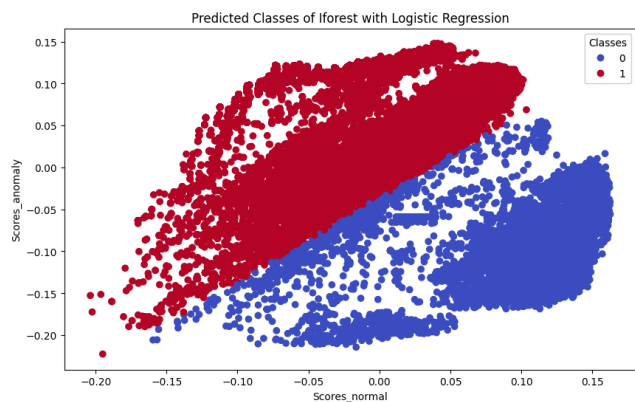


図 8 UNSW3 (和文)

Fig. 8 Description of the collectness\_UNSW figure (English).

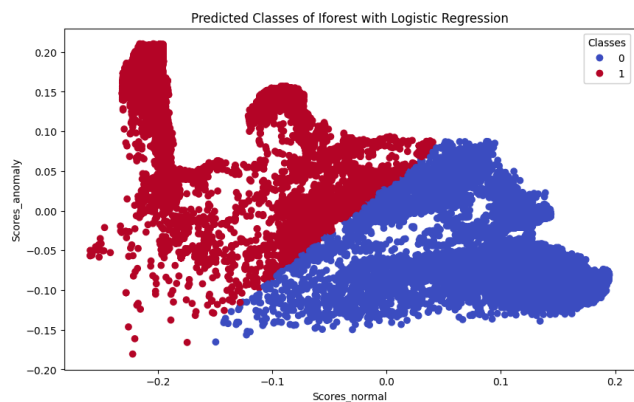


図 11 NSL3 (和文)

Fig. 11 Description of the collectness\_UNSW figure (English).

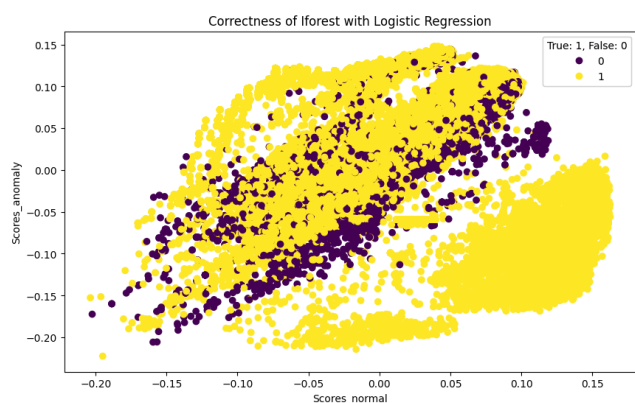


図 9 UNSW4 (和文)

Fig. 9 Description of the collectness\_UNSW figure (English).

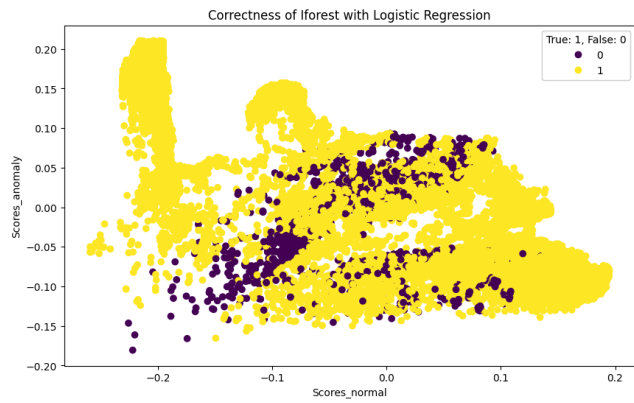


図 12 NSL4 (和文)

Fig. 12 Description of the collectness\_UNSW figure (English).

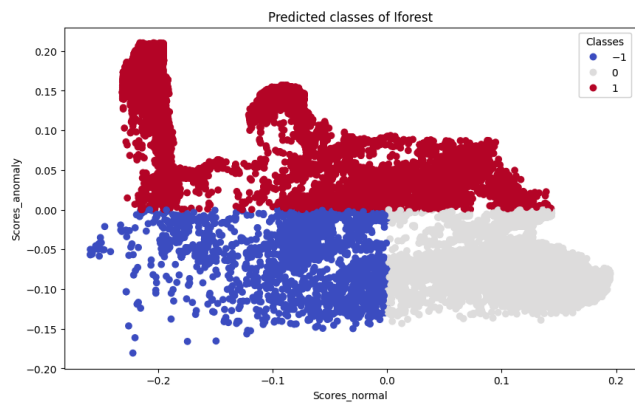


図 10 NSL2 (和文)

Fig. 10 Description of the collectness\_UNSW figure (English).

較・ロジスティック回帰のコスト vs iforest の探索 ( ) の計算量 (時間) を比較する

ロジスティック回帰におけるランダムサンプルしての精度の比較 (調べてもいい) サンプリングの適用範囲 (データ量と精度のトレードオフ)

svcc 判定の決定技を作成して (さまざまな観点から決定技だと redteam と合意できる) first(国際機関の cert) ≠ ≠ ≠

## 5. おわりに

おわりにを書く.

謝辞 謝辞を書く.

## 参考文献