神奈川工科大学

セキュリティ研究センター

研究報告

第 10 巻

2021 年度

工学教育研究推進機構

目次

・分散機械学習モデルに基づいた安全な IoT サービスを実現するための
統合セキュリティ対策技術に関する研究

情報ネットワーク・コミュニケーション学科　岡崎美蘭
情報工学科　納富一宏
情報ネットワーク・コミュニケーション学科　岡本学

・IoT マルウェア解析のためのツールチェインの特定

情報ネットワーク・コミュニケーション学科　岡本剛

分散機械学習モデルに基づいた安全な IoT サービスを実現するための統合セキュリティ対策技術に関する研究

研究者名：　情報ネットワーク・コミュニケーション学科　岡崎　美蘭
情報工学科　納富　一宏
情報ネットワーク・コミュニケーション学科　岡本　学

1. 研究の目的

本研究の目的は，誰もが安心して暮らせる超スマート社会を実現するためにフィジカル空間とサイバー空間を融合したすべての部分で安全性と信頼性を確保することができる高度かつ堅牢な IoT セキュリティ統合対策技術を確立することである．具体的には，超スマート社会実現におけるサイバー攻撃の脅威を明確に分析し，フィジカル空間からサイバー空間まで一貫して安全性を提供可能となる高度な IoT セキュリティシステム構築技術の確立について検討する．そして，不正デバイスと不正アプリケーションを見分けることができるデバイス認証技術とユーザ認証技術を開発するとともに，不正データ・異常データをより早く検知することができる信頼性の高い分散機械学習モデル構築手法の研究開発を行うことである．

今年度は，昨年度検討・提案したスマートフォンとスマートウォッチのようなウェアラブル端末を用いた個人認証方式の有効性について検討することを目的とする．さらに，スマートフォンの行動的特徴量による個人識別に関する検討と文章作成中の打鍵情報による継続的な本人認証についても検討を行う．

2. 研究の必要性及び従来の研究

現在 IoT 機器の急増に伴い，IoT 機器の不十分なセキュリティ設定や脆弱性を突いたマルウェア感染，DDoS 攻撃などサイバー攻撃が増加する中で，世界的に IoT セキュリティ対策に関する研究活動も活発に行われている．例えば，フィジカル空間の不正な IoT 機器を検出するために，ネットワークスキャンなどで脆弱な機器を見つける機器推定機能などが提供されている．その IoT 機器の中に，電子的に個人の認証を行い家のドア等の鍵の開閉を行うスマートロックと呼ばれる製品がある．しかし，スマートロックにおける個人認証方式として，従来のパスワード方式ではパスワードが推測されやすくユーザに記憶負荷がかかるという問題点や，毎回記憶したパスワードを端末に入力する必要があり日常的に使う鍵の開閉を行うための個人認証方式としては煩わしさが残る．また，IC チップの埋め込まれたカードや事前に認証を済ませた端末などを持ち歩くことで認証を行う所持認証と呼ばれる方式があるが，スマートロックにおける所持認証方式では，従来の鍵と同様に紛失や盗難等の危険性が課題として挙げられる．これらの課題点を解消する方式として人間の身

体的特徴を用いて認証を行う生体認証方式がある．しかし，スマートロックにおける生体認証方式として，例えば指紋認証方式では，家の前で数秒間立ち止まらなくてはならず，手袋をつけていればそれを外す必要がある．また，顔認証方式においてもドアの前で数秒間立ち止まる必要があり，マスクや帽子等を身に付けていればそれらを外さなくてはならないという煩わしさがある．

本研究ではスマートロックにおける認証の煩わしさを排除するため，ユーザの行動から得られる行動的特徴を用いた歩行認証に着目した．歩行認証とは，歩行時の加速度データや角速度データなどから行動的特徴を抽出し認証を行う生体認証方式の一つであり，日常的に行われる歩行を用いることでユーザが意識することなく認証を行うことができる利便性があると考えられる．しかし，行動的特徴を用いる認証方式では，本人の行動を常に再現することが難しく認証精度に大きな課題がある．これまでも歩行認証は認証精度向上のため多くの研究が行われている．これらの研究では，機械学習や複数のセンサを用いることで認証精度の向上のための研究を行っている．しかし，認証の際に端末の向きを同じにしなくてはならないという煩わしさや，複数の同一センサを身に着ける必要がある等の課題点がある．また，行動的特徴を用いる認証方式には行動を他人に真似され不正なユーザに認証されてしまうという問題点がある．この問題点に対しては，スマートフォンから取得した加速度データを基に歩行認証を行い，なりすまし攻撃に対して有効か確認を行っている研究も実施されているのが現状である．

3. 期待される効果

従来のような端末の向きを考慮せず，さらに複数のセンサを用いることもなく，スマートフォンとスマートウォッチの二つの端末の合成加速度を用いることでユーザが意識することなくスマートロックの個人認証が可能となる．また，事前に登録したウェアラブル端末の ID による所持認証を行うことで歩行認証の認証精度の課題を解消することができると期待される．

4. 研究の経過及び結果・評価

本研究では，加速度センサを搭載した二つの端末を用いた歩行認証システムモデルの提案と評価を行う．提案システムモデルでは，事前に登録した端末の ID による所持認証を行うことで歩行認証の認証精度の課題を解消することを目的とする．加速度センサの計測では，iBeacon による近接検知を行うことによって加速度データの取得開始・終了の判定を行う．また，二つの端末から取得した加速度データから極大値間隔や二つの端末の類似度等の特徴量を抽出することで，学習器を作成しスマートロックの認証を行う．学習器には，機械学習の異常検知モデルを用い，本人のデータのみを学習させ，正常データの場合は鍵を開錠し，異常データの場合は施錠することで認証を行う．

4.1 二つの端末を用いた異常検知による歩行認証システム

提案システムモデルは，図 1 に示すように加速度センサが搭載された二つの端末，加速度データを受信しユーザの認証を行うスマートロックで構成される．

スマートロックと二つの端末の通信方式には Bluetooth を用い，iBeacon によりスマートロックと二つの端末との近接検知を行う．iBeacon では Immediate（約 2cm 未満），Near（約 2cm～1m），Far（約 1m～約 50m）の 3 つの近接検知を行うことができ，本提案システムモデルでは Near と Far の二つの近接検知を用いて加速度データの計測を行う．

二つの端末は，以下の手順で認証を行う．

(1) iBeacon の Far 検知
(2) 事前登録した二つの端末の ID 送信
(3) スマートロック内で二つの端末の ID 確認
(4) 二つの端末での加速度データの計測開始
(5) iBeacon の Near 検知
(6) 計測した加速度データをスマートロックへ送信
(7) スマートロック内で受信した二つの端末の加速度データを処理
(8) 加速度データから特徴量抽出
(9) 機械学習による異常検知を行い正常データなら解錠，異常データなら施錠する

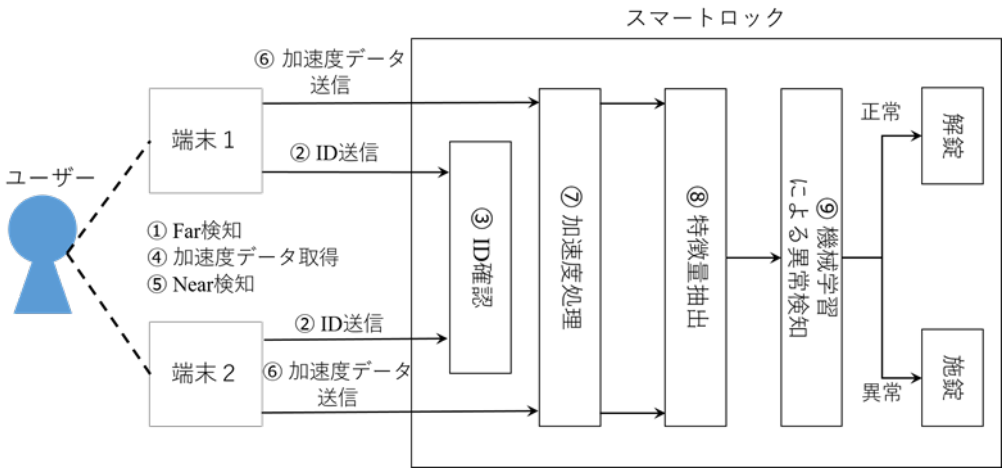

図１．二つの端末を用いた歩行認証システム

4.2 実装と実験

提案手法の有用性を確かめるために，端末 1 にスマートフォン Sony Xperia XZs を用い，端末 2 にウェアラブル端末 Sony SmartWatch3 を用い実装を行った．また，スマートロックの代わりに Linux OS をインストールしたパソコンをサーバとして用意し，各端末から加速度データを取得するため加速度の開始及び終了命令を送信するための

プログラムを実装した.実装したプログラムでは,まずより良い特徴量選択を行うため,全ての特徴量の組み合わせに対して認証率を計算し,それぞれの機械学習モデルに適した特徴量の選択を行う.次に,選択された特徴量を用いて,機械学習モデルを作成しそれぞれの機械学習のパラメータの調整や異常度を算出しその異常度の閾値を調整することで認証率を計算し,最適なパラメータ及び閾値を決定する.機械学習には Python モジュールの scikit-learn を用い,異常検知アルゴリズムとして,Elliptic Envelope, GMM, Isolation Forest, KDE, LOF, One Class SVM を用いそれぞれの認証率を計算し評価を行った.

次に提案方式の識別率と認証精度を評価するための実験を行った.実験方法は,次のように 3 つの実験に分けて実施した.

① 予備実験

機械学習の教師ありアルゴリズムを用いて一つの端末を用いた場合より二つの端末を用いた方が歩行認証の識別率が向上するかを確認した.ここでの識別率は,全ての試行回数中の正しく分類された回数の割合とする.

② 普段の速度での歩行実験

次に提案方式が日常の歩行に適用できるかどうか確認するため,機械学習の異常検知を用いた認証精度の確認実験を行った.ここでは,日にちを空けずに普段通りの速度で歩行してもらった.

③ 異なる速さの歩行実験

異常検知を用いた歩行認証の継続的な認証及び異なる速度での歩行認証の認証精度を確認するため,日にちを空け早歩き,普通歩き,遅歩きの三種類の速さでの歩行実験を行った.

実験では,日にちを空けた場合の普段の速度での歩行認証の有効性を確認するため,学習データには本人のデータのみを学習させ,評価を行った.結果として,最も良かった識別率では機械学習に KDE を用いた場合の平均 FAR が 18.6%, 平均 FRR が 13.8%であった.ここでは,大きく悪くなった機械学習のアルゴリズムもあればあまり変化のないアルゴリズムもあり,例えば KDE では, FAR が+3.4%, FRR が-3.9%とそれほど認証精度が変わらないという結果となった.次に,異なる速度での歩行認証の有効性を確認するため,学習データには異なる速度の歩行データを学習させ,評価を行った.結果として,最も良かった識別率は機械学習に GMM を用いた場合の FAR が 18.8%であり,FRR が 15.0%であった.こちらも,認証精度は悪くなるという結果となったが,大きく FAR や FRR が悪くなってしまった機械学習アルゴリズムもあれば変化の少ないアルゴリズムもあり,特徴量の抽出方法や加速度処理の方法によっては,認証精度の向上ができると考えられる.

5. 今後の計画

歩行認証という人間の行動的特徴を用いた認証方式は，同一人物でも常に同じ歩き方をするというのはなかなか難しく，認証率の精度をいかに上げるかが課題になっている．そこで，異なる速度での歩行認証の有効性を確認するため，学習データには異なる速度の歩行データを学習させ，評価を行った．認証率は悪くなるという結果となったが，大きくFARやFRRが悪くなってしまった機械学習アルゴリズムもあれば変化の少ないアルゴリズムもあり，特徴量の抽出方法や加速度処理の方法によっては，認証精度の向上ができると考えられる．

最後に，今回の実験で使用したデータの数は十分とは言えない．データが少ないことによって生じうる影響として，学習データにいつもと違う歩行データが混ざってしまった場合にそのデータに影響を受けてしまいFRRが高くなってしまう可能性が考えられる．他にも，データが少ないことにより他人との境界をつけづらくFARも高くなってしまうと考えられる．今後の課題としては，認証精度向上のためより幅広い年代の十分なデータでの分析，より良い特徴量の抽出や，より良い特徴量選択の方法などが挙げられる．今後の課題としては，眼鏡型のウェアラブル端末など様々なウェアラブル端末から歩行状態の加速度データを取得した場合でも高い精度で認証を行うことができるか確認することが挙げられる．

6. 研究成果の発表

(1) 論文

[1] 朴 美娘，渡辺 一樹，油田 健太郎，岡崎 直宣 “スマートロックにおける2端末による機械学習を用いた歩行認証に関する研究” IPSJ, Vol.62, No.12, pp.2011-2023 (2021-12-15) http://id.nii.ac.jp/1001/00214243/

[2] M. Park, K. Aburada, N. Okazaki, ″Proposal and Evaluation of a Gesture Authentication Method with Peep Resistance for Smartwatches”, Proceedings of the 2021 9th International Symposium on Computing and Networking, CANDAR 2021, pp.359-364, 2021/11/23

[3] S. Usuzaki, K. Aburada, H. Yamaba, T. Katayama, M. Mukunoki, M. Park, N. Okazaki, ″Proposal and Evaluation for Color Constancy CAPTCHA”, Journal of Artificial Life and Robotics, 2021/09

(2) 学会発表

[1] 伊藤 憂香，朴美娘 “視線入力デバイスを用いた画像認証方式に関する研究”電子情報通信学会　総合大会2022, A-17-1 (2022/03/16)

[2] 渡辺 一樹，油田 健太郎，岡崎 直宣，朴 美娘 “継続的な認証及び異なる歩行速度での二つの端末を用いた異常検知による歩行認証の研究,″ ICCS/SPT 研究会 Vol.2021

SPT-41 No.16, 2021/3/1.

[3] 嘉藤 鴻介, 渡辺一樹, 油田健太郎, 岡崎直宣, 朴美娘, “ウェアラブル端末を用いた覗き見耐性を持つジェスチャー認証手法の提案と評価,” ICCS/SPT 研究会 Vol.2021 SPT-41 No.17, 2021/3/1

(3) 特許

[1] 特願 2018-036579 「加速度センサを搭載したモバイル端末を正規の通信相手として認証する方法および認証装置」

発明者：岡崎　美蘭

出願者：学校法人　幾徳学園

[2] 特願 2019-189314 「システム、方法、情報処理装置およびプログラム」

発明者：岡崎　美蘭

出願者：学校法人　幾徳学園

(4) 助成金採択

[1] 2020 年～2022 年度　文部科学省科学研究費補助金　基盤 (C)

「分散機械学習モデルに基づいた高度かつ堅牢な IoT セキュリティ対策技術に関する研究」

研究代表者：岡崎　美蘭

IoT マルウェア解析のためのツールチェインの特定

研究者名：情報学部　情報ネットワーク・コミュニケーション学科　岡本　剛

１．研究の目的

IoT 機器の増加とともに，IoT 機器に感染するマルウェアが増加している．多くの IoT マルウェアにおいて，ライブラリ関数が静的結合され，関数などのシンボル情報が消去されているため，関数レベルでのマルウェア解析が困難であることがわかっている．我々の先行研究で提案した手法を使えば Intel 80386 の IoT マルウェアに静的結合されたすべてのライブラリ関数とツールチェインを特定できることがわかったが，Intel 80386 以外の IoT マルウェアに対して提案手法が有効であるかが明らかでなかった．そこで，提案手法を Intel 80386 以外の IoT マルウェアに適用する方法を明らかにするとともにハニーポットで収集したマルウェア検体のツールチェインを特定することを本研究の目的とする．なお、ツールチェインを特定すればマルウェアが使用するライブラリ関数のほとんどを特定できることもわかっている。

２．研究の必要性及び従来の研究

多くの IoT マルウェアにおいて，関数レベルでのマルウェア解析が困難であることがわかっている．具体的にはマルウェアの動的解析において関数の実行履歴を解析できない。静的解析においてはマルウェア定義関数を解析する必要があるが、これらが呼び出すライブラリ関数が明らかではないため、マルウェア定義関数の解析に多くの時間を必要とし、これが原因で静的解析が敬遠されやすい。ツールチェインの特定はこれらの問題を解決するための有効の手段となる。

関数の特定にはこれまで様々な方法が提案されている．パターンマッチングによって関数を特定する方法は，コンパイラやライブラリの組み合わせ（ツールチェイン）ごとにマシンコードが変化し，組み合わせが多様であるため，関数を特定することが難しい．このような理由からパターンマッチングによる方法はほとんど行われていなかった。マシンコードや制御フローグラフの類似度によって関数を特定する方法は，ツールチェインの変化に対して柔軟に対応できるため、最近の研究のトレンドであるが、オペランドのレジスタや即値が異なる場合でもその周辺のマシンコードが一致する場合，関数を誤検知する．特にライブラリ関数にはオペランドの即値のみが異なる関数が多く存在するため，誤検知が発生しやすい．

３．期待される効果

研究成果は高精度の関数特定（ツールチェイン特定）であり、本成果の期待される効果は、関数レベルの解析である。動的解析ではライブラリ関数の実行履歴を追跡できるようになる。そのおかげで関数レベルでマルウェアの機能を解析できる。静的解析ではライブラリ関数の特定を省略できるおかげで、マルウェア定義関数の解析に注力する時間を増やせる、言い換えると解析に要する時間を短縮できる。この他に、マルウェアが利用するライブラリ関数のリストを比較すればマルウェアファミリを特定できる可能性がある。

４．研究の経過及び結果

研究は予定通り完了して次の結果が得られた。これらの成果をマルウェア対策のコミュニティと共有するため、すべての検体についてすべての関数のアドレスと関数名を GitHub に公開した.

- 提案手法は Intel 80386 を含むすべての検体のツールチェインを特定できることを示した.
- 特定したツールチェインは合計 14 種類あり，すべてのツールチェインは カスタマイズされたものではなく，Web サイト上で公開されているものであった.
- Intel 80386 以外のアーキテクチャの検体の 91.4%が Firmware Linux 0.9.6 を使用してビルドされており，Intel 80386 と同様に Firmware Linux 0.9.6 の検体数が最も多かった.
- C ライブラリは 99.0%が uClibc，1.0%が musl を使用していた.

上記の成果はコンピュータセキュリティ研究会（CSEC）にて発表した。

５． 今後の計画

ツールチェインを手がかりにしてマルウェアが利用するライブラリ関数の完全な特定を今後の課題とする。

６． 研究成果の発表

1. 赤羽秀，岡本剛，”ライブラリ関数が静的結合された IoT マルウェアのビルドに使用されたツールチェインの特定”，情報処理学会研究報告，2021-CSEC-93，pp. 1 - 8，2021. 査読なし