Copyright ©2022 The Institute of Electronics,
Information and Communication Engineers

SCIS 2022 2022 Symposium on
Cryptography and Information Security
Osaka, Japan & Online, Jan. 18 – 21, 2022
The Institute of Electronics,
Information and Communication Engineers

関数呼び出しグラフと関数埋め込みに基づくマルウェア分類手法
Malware Classification based on Function Call Graph and Function Embedding

林　実奈美 *‡　　大坪　雄平 * ††　　大塚　玲 ‡
Minami Hayashi　　Yuhei Otsubo　　Akira Otsuka

キーワード　機械学習，マルウェア，ファミリ分類，グラフニューラルネットワーク

あらまし

年々，高度化及び多様化するマルウェアに対応するためには，機械学習の活用が不可欠である．マルウェアの種類は膨大であるが，その大半は既存のマルウェアを少しだけ改変した亜種である．したがって，マルウェアのファミリ分類を適切にできれば，挙動の推測が容易となり，解析の効率化に役立てることができる．機械学習を活用したマルウェア分類手法は様々なアプローチで行われているが，最近では，グラフニューラルネットワーク（GNN）[1] を応用した研究が注目されている．GNN をマルウェア分類に使用することで，プログラムの条件分岐や関数呼び出しの関係を考慮した分類が可能となる．

本研究では，GNN を用いてマルウェアのファミリ分類器を実装し，プログラムの関数呼び出しグラフを入力として，分類器の学習及び評価を行った．関数の特徴抽出には，Massarelli らによって提案された SAFE(Self-Attentive Function Embeddings for Binary Similarity)[2] を用いた．SAFE では，関数の特徴抽出に自然言語処理技術を応用することで，関数の特徴を捉えたベクトル表現の獲得に成功している．我々は，SAFE を使用して抽出した関数の特徴量を，関数呼び出しグラフのノードに割り当て，GNN での学習・評価を行った．特徴抽出に自然言語処理技術を用いたことで，従来手法と比較して高精度での分類に成功した．本稿では，本提案手法の概要及び実験結果について説明した後，GNN を用いたマルウェア分類手法の有効性について述べる．

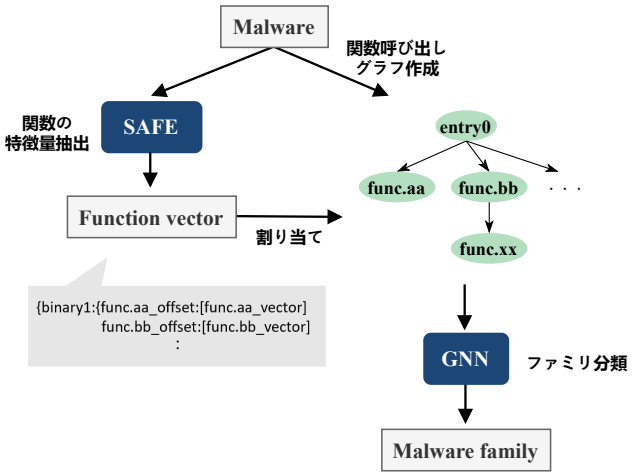

図 1: 提案手法

参考文献

[1] Scarselli, Franco, et al. 2009. “The Graph Neural Network Model.” IEEE Transactions on Neural Networks / a Publication of the IEEE Neural Networks Council 20 (1): 61–80.

[2] Massarelli, Luca, et al. 2021. “Function Representations for Binary Similarity.” IEEE Transactions on Dependable and Secure Computing, 1–1.

* 警察大学校, 東京都府中市朝日町 3-12-1, National Police Academy, 3-12-1 Asahi-cho, Fuchu, Tokyo
† 警察庁, 東京都千代田区霞が関 2-1-2, National Police Agency, 2-1-2, Kasumigaseki, Chiyoda, Tokyo
‡ 情報セキュリティ大学院大学, 神奈川県横浜市神奈川区鶴屋町 2-14-1, Institute of Information Security, 2-14-1 Tsuruya-cho, Yokohama, Kanagawa