頻度を考慮したAPIコールグラフによるマルウェア分類手法の提案

Malware Classification Method using API Call Graph considering Frequency

菊地 寿史・ネットワーク分科会・情報セキュリティ大学院大学

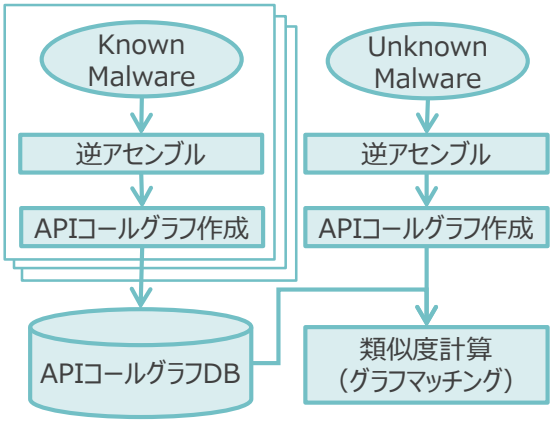
We focus on the fact that malware makes various API calls to realize its purpose and function, and defined API Call Graph that graphed the order and appearance frequency of API calls from assembly code. Therefore, we proposed a method using API Call Graph to calculate similarity between malware.

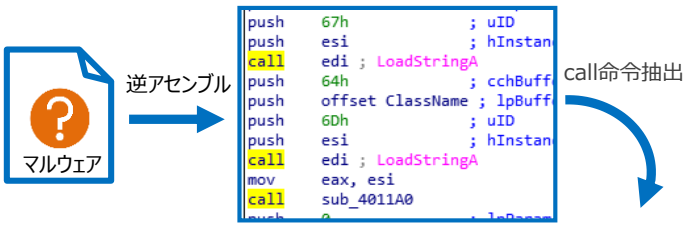
研究背景と提案手法

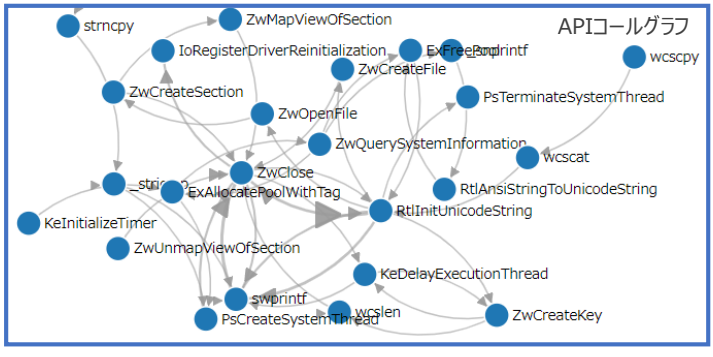
- 未知マルウェアがどのファミリーに分類できるか分かれば、感染時の被害状況の早期特定や解析の効率化を期待できる
- 本研究では、マルウェアのアセンブリコードから抽出できるAPIコールを特徴としたマルウェア分類手法を提案する

【提案手法】

- マルウェアが自身の目的や機能を実現するために様々なAPIを利用することに着目
- アセンブリコードから**APIコールの順序**と**出現頻度**を抽出してグラフ化（=APIコールグラフ）
- 未知マルウェアと既知マルウェアのグラフを比較し、マルウェア間の類似度を0～1で算出。類似度計算にはJaccard係数を利用







提案手法の評価

- ①～⑩のマルウェアファミリーの亜種を各10個ずつ使用し、提案手法の評価を実施

①Backdoor.Afcore	⑥Trojan-PSW.Coced
②Backdoor.Swz	⑦Trojan-Ransom.Hexzone
③Rootkit.Podnuha	⑧Virus.Delf
④Trojan-Downloader.Hmir	⑨Virus.Xorer
⑤Trojan-Dropper.BAT	⑩Worm.FlyStudio

【実験1】同一ファミリー内の類似度

- 同一ファミリー内の亜種同士で類似度を算出

【実験2】異なるファミリーとの類似度

- ①と他ファミリー②～⑩の亜種の類似度を算出

実験1、2の結果（類似度平均）

	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩
実験1	.556	.498	.960	.384	.559	.544	.565	.379	.550	.801
実験2	-	.020	.026	.003	.007	.008	.007	.008	.006	.010

【実験3】正規ソフトウェアとマルウェアの類似度

- 著名な正規ソフトウェアとマルウェアの類似度を算出

正規ソフト	概要	マルウェアとの類似度平均
FFFTP	FTPクライアント	.0094
Chrome	Webブラウザ	.0067
Tera Term	ターミナルエミュレータ	.0072
+Lhaca	ファイル圧縮・解凍	.0104
サクラエディタ	テキストエディタ	.0021

まとめ

実験の結果、本提案手法の有効性を確認できた。

	対象	類似度の傾向	
実験1	同じファミリー内の亜種	高	0.37～0.96
実験2	異なるファミリーの亜種	低	0.003～0.2
実験3	正規ソフトウェアとマルウェア	低	0.01以下

本提案手法により、マルウェアの亜種分類が可能！

- ✓ 高い類似度になった場合、同じファミリーの可能性が大きい
- ✓ 異なるファミリーや正規ソフトの誤判定の可能性は小さい