戦略的情報通信研究開発推進事業(SCOPE)＿ICT基礎・育成型研究開発

ポスター番号

P-03-14

原因に基づく悪性DNSクエリ分類技術の研究開発

佐藤彰洋[1], 中村豊[1], 福田豊[1]　[1]九州工業大学

1. はじめに

マルウェアはインターネットにおける重大な脅威のひとつである．ネットワーク内の感染端末を検出するためには，ブラックリストを利用した通信の監視が一般的である．しかしながら，ブラックリストに基づく検出は，(1)ブラックリストは必ず幾つかの誤りを含むこと，(2)検出結果の正誤の判断が困難であることが問題となる．故に，単純にブラックリストに合致するか否かでは済まず，管理者による通信の調査と原因の特定が必須となる．

本研究開発では，ブラックリストに基づいて検出された悪性DNSクエリを原因ごとに分類する技術の実現可能性を検証する．この原因に基づく分類の実現により，分類結果の代表的なDNSクエリのみに調査範囲を限定できるため，分析を要する悪性DNSクエリの数を大幅に削減することが可能となる．

2. 原因に基づく悪性DNSクエリ分類技術

本研究開発では，ブラックリストによる検出結果の効率的な分析のため，悪性DNSクエリ分類技術を実現する．その独自性は，(1)ブラックリストと合致したクエリの前後には，その原因の推定を助けるクエリ群が存在することに着目したこと，(2)クエリの数値表現のため，クエリ間の共起関係を利用したこと，(3)クエリに対する重みの付与のため，一般的なマルウェアの性質，すなわち同一マルウェアファミリに感染した端末は，共通の悪性ドメイン群と繰り返し通信する性質を考慮したことの3点である．これにより，従来のドメイン文字列による表層的な類似性に基づく分類とは異なり，悪性クエリとそれに付随するクエリ群が潜在的に示す原因に基づく分類を実現する．

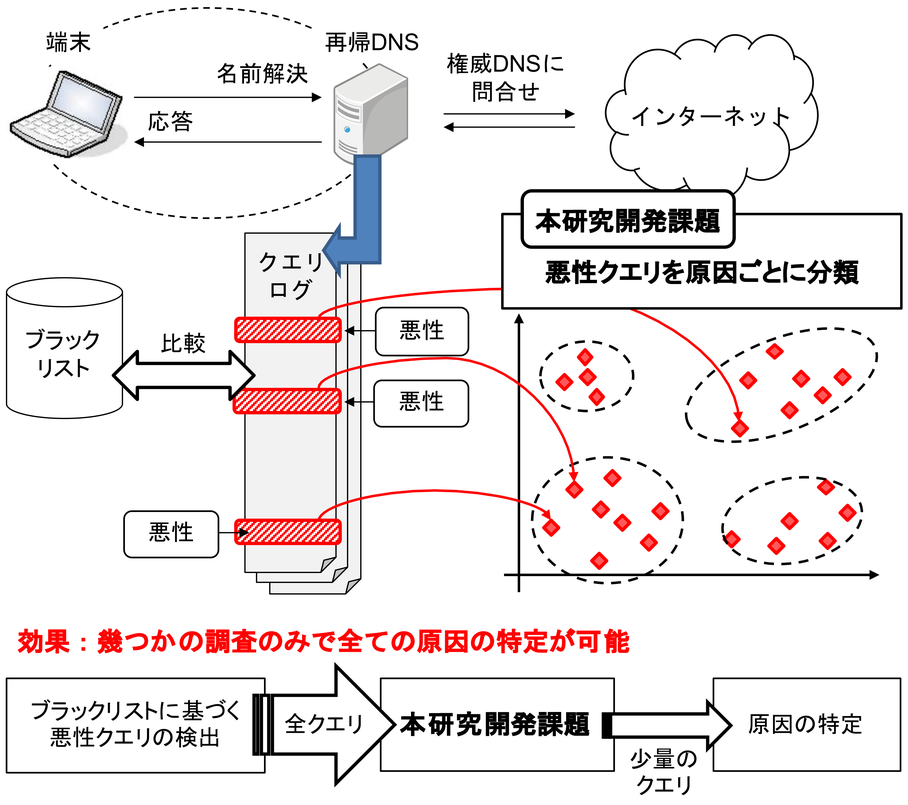

図1 研究課題の概要

クエリ部分ログ選択機能：(a)マルウェアの通信は幾つかのタスクにより構成されること，(b)ブラックリストと合致した悪性クエリはマルウェアのタスクの一部によるものであり，その悪性クエリの前後にはマルウェアの別のタスクによるクエリが生じることに着目した．

クエリ数値化機能：全クエリ部分ログに含まれる膨大な数のクエリの特徴を効率的に表現するため，クエリ間の共起関係に基づいたクエリの数値化を試みる．これにより，各クラスタには，共起関係が類似したクエリ，すなわち機器間の通信において担うタスクが類似したクエリの集約が可能となる．

悪性クエリ類似性導出機能：(a)共起関係が類似したクエリはマルウェアの通信において担うタスクが類似したクエリであること，(b)クエリ部分ログにおいて，そのクエリ群が担うタスクの類似性はその原因の類似性に強く依存することに着目した．

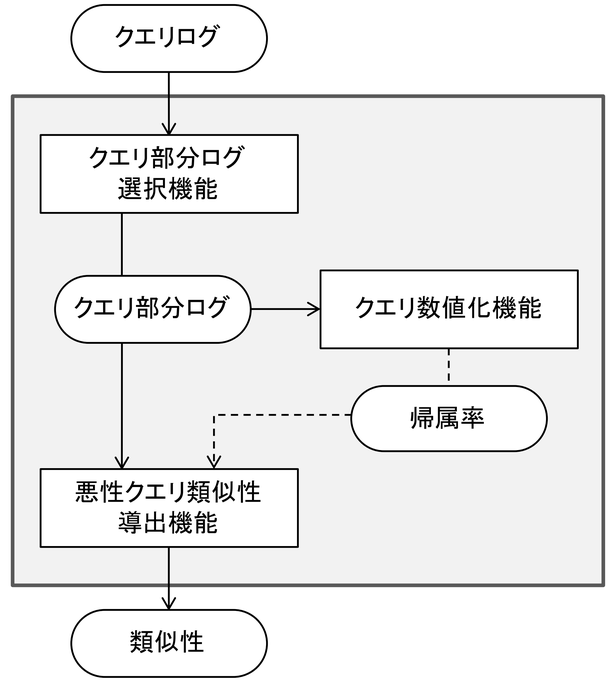

図2 悪性DNSクエリ分類技術

3. 評価

既存手法のDoc2Vecの結果は，各シンボルが散乱しているため，これらから悪性クエリの類似性を判断することが困難である．それに対して，本システムはブラックリストにより検出された388の悪性クエリを3つのクラスタに分類できること，各クラスタが共通の原因のクエリのみで構成されることを確認した．この結果は，管理者が各クラスタの代表的なクエリのみを調査することで全ての原因を追跡できるため，ネットワーク内の感染端末を迅速に排除できることを示している．

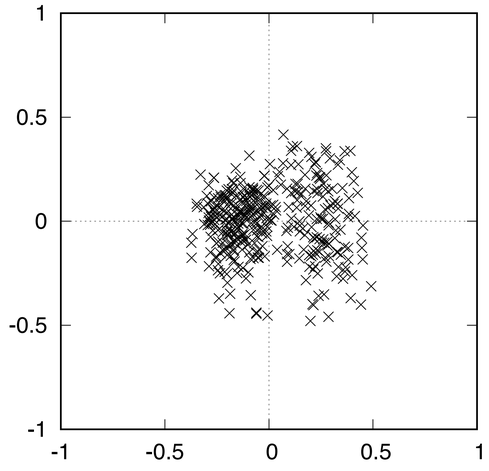

図3(a) Doc2Vec

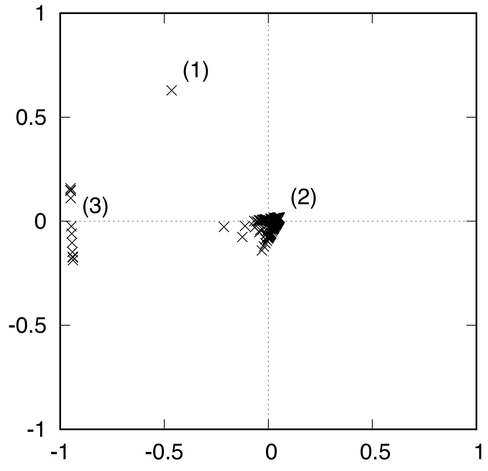

図3(b) Our Work

ICTイノベーションフォーラム2020
戦略的情報通信研究開発推進事業（SCOPE）