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

プロキシログから抽出した通信パターンによる異常検知
Anomaly detection focused on communication pattern from Proxy log

名倉 悠 * 青木 茂樹 * 宮本 貴朗 *
Yu Nagura Shigeki Aoki Takao Miyamoto

キーワード 侵入検知, プロキシログ, GMM クラスタリング, マルコフモデル

あらまし

はじめに

端末に不正侵入する攻撃は事前に防ぐことが難しいため, 感染後の迅速な異常の検出が重要視されている. このような攻撃の検知手法の 1 つとして, 解析対象をクラスタリングし, クラスタ間の遷移確率を用いる手法が提案されている. 本稿では, クラスタ間の遷移にはユーザごとに固有の特徴が現れると考え, 遷移パターンを確率モデルで学習することにより高精度に異常を検知する手法を提案する.

提案手法

まず, Proxy サーバのログの各行から受信バイト数, ドメインの長さなどの 7 次元の特徴量を抽出する. これらの特徴量の尺度を合わせるため, それぞれ標準化する.

次に特徴量を GMM を用いてクラスタリングし, 得られたクラスタ番号を各行に付与する. ログを端末ごとに, 時間情報を基に分割することでセッションを定義する. セッション内のクラスタ番号の列をクラスタシーケンスとし, ユーザの普段の通信パターンとして抽出する.

端末ごとに, マルコフモデルを基にして全シーケンスにおけるクラスタ間の遷移の発生確率を全て算出し, それらをまとめた確率モデルを作成する.

検知を行う場合は, 検証用データについて, 学習データ同様にクラスタシーケンスを作成する. 作成したクラスタシーケンス内のクラスタ間の遷移確率を確率モデルから算出し, 尤度とする. そして, 尤度が閾値より小さいものを異常なセッションとして検出する.

* 大阪府立大学大学院人間社会システム科学研究科, 599-8531 大阪府堺市中区学園町 1-1, Graduate School of Humanities and Sustainable System Sciences, Osaka Prefecture University, 1-1 Gakuen-cho, Naka-ku, Sakai, Osaka 599-8531, Japan. sba00203@edu.osakafu-u.ac.jp, aoki@kis.osakafu-u.ac.jp

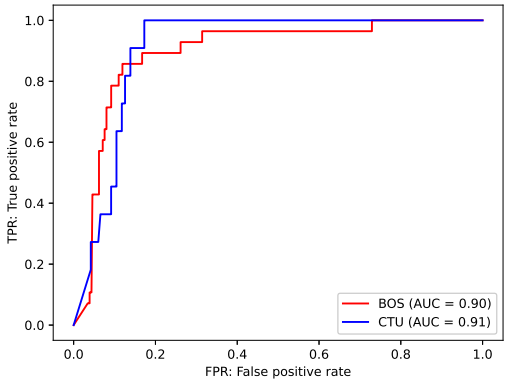

図 1: 検知結果の ROC 曲線

実験

本手法の有効性を確認するため, 2 種類のデータセットを用いて実験を行った. 1 つは本学の端末の Web アクセスログに, BOS2018 データセットから抽出した悪性ログを挿入したデータセット, もう 1 つはチェコ工科大学 (CTU) の Stratosphere IPS Team が公開しているマルウェアデータセットから抽出した悪性ログを挿入したデータセットである. それぞれ C&C サーバ, 攻撃者端末と通信しているログを悪性ログとし, これらが含まれるセッションを異常, それ以外のセッションを正常として実験を行った. 結果を図 1 に示す. AUC はそれぞれ 0.90, 0.91 となり, 有効性を確認できた.

おわりに

本稿では, プロキシログに対してクラスタリングを行い, セッションごとの各ログのクラスタ遷移を確率モデルに適用することで異常を検知する手法を提案した. 今後の課題としては, 他の特徴量の有効性の確認や抽出方法の検討, 他の確率モデルを適用した場合との有効性の比較検証などが挙げられる.