様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　５ 年　 ６ 月 ２６ 日現在

機関番号：３２８０９
研究種目：研究活動スタート支援
研究期間：2021～2022
課題番号：２１Ｋ２１２９６

研究課題名（和文）深層学習を用いた予測をベースとするDNSへの異常検知手法の構築

研究課題名（英文）Anomaly Detection Method for DNS based on Prediction Using Deep Learning

研究代表者
木村　知史（Kimura, Satoshi）

東京医療保健大学・医療保健学部・助教

研究者番号：９０９０９１１０

交付決定額（研究期間全体）：（直接経費）　　1,100,000円

研究成果の概要（和文）：本研究では，DNS amplification攻撃を最短時間で検知するため，DNSパケット数に深層学習を適用し予測値を算出することで，予測をベースとした異常検知手法の構築を目的とした．本手法を構築するために，DNSパケット数をLSTMを用いて学習し，長期的な予測値を算出する一括予測手法と，短期的な予測値を算出する繰返し予測手法の二つを提案した．さらに，学習中に発生した誤差を予測値に加えてしきい値として設定し，異常を自動的に判定する動的しきい値法を提案した．
本手法を複数の期間わたって適用し評価を行った．その結果，全ての期間で繰返し提案手法が一括予測手法に対し優れた手法であることを明らかにした．

研究成果の学術的意義や社会的意義
本研究で構築したDNSパケット数の予測をベースとする異常検知手法は，高い精度で予測が可能であることを明らかにした．また，誤検知は発生しつつも，自動的にDNS amplification攻撃に関連する行為を検知できることを明らかにした本研究の成果は学術的に意義をもつ．また，本研究で使用したDNSパケットは24時間ごとに周期傾向を有しており，対象とするシステムを他の周期傾向を有するプロトコルへの拡張などの新たな展開に期待できると考えている．
今後の方向性として，異なるネットワーク構造を有する拠点においても提案手法が実用可能であるかを確認する．

研究成果の概要（英文）：In this study, our goal is to construct a prediction-based anomaly detection method by using deep learning techniques. Specifically, we aim to calculate the number of Domain Name System (DNS) packets for predicting and automatically detecting DNS amplification attacks. To develop this method, we trained the system on the number of DNS packets using Long Short-Term Memory (LSTM) models, proposing two prediction methods: a batch prediction method that calculates long-term predictions, and an iterative prediction method that calculates short-term predictions. Furthermore, we introduced a dynamic thresholding method that automatically determines anomalies by setting a threshold value in addition to the prediction value based on the errors that occur during training.
We evaluated the proposed methods over several periods. The results showed that the proposed method was superior to the batch prediction method repeatedly in all periods.

研究分野：情報セキュリティ

キーワード：ネットワークセキュリティ　DNS セキュリティ　異常検知　予測　深層学習　時系列データ

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９（共通）

１．研究開始当初の背景

オープンリゾルバの探索行為やそれを利用した DDoS 攻撃の一種である DNS amplification 攻撃は DNS を狙った攻撃として広く認識されており，政府が目指す未来社会 Society 5.0 を構築するうえで，その対策が強く求められている．DNS amplification 攻撃を抑制するための機能として知られている Response Rate Limiting (RRL) は，単位時間あたりのレスポンス回数を制限する DNS の拡張仕様である．しかし，RRL は一般に権威 DNS に対して実装され，DNS amplification 攻撃を検知する根本的な解決手段ではない．

また，所定の時間あたりに静的なしきい値を設定し，DNS へのアクセス回数がしきい値を上回った場合に異常を検知する手法が考えられる．しかし，DNS に対するアクセス回数は曜日や時間帯によって時々刻々変化する性質を持っているため，静的なしきい値を多様なネットワーク環境において管理者が適切に設定することは難しい．

２．研究の目的

本研究の目的は，DNS amplification 攻撃を最短時間で検知するために，DNS の時系列データに対して深層学習を適用し，時間的な予測値を算出することで，予測をベースとした異常検知手法を構築することである．具体的には，過去に観測した DNS パケット数を Long Short-Term Memory (LSTM) を用いて学習し，48 点先のパケット数を一度に予測する一括予測手法と，パケット数が得られるたびに 48 点の予測値を繰返し算出する繰返し予測手法の二つを採用し，その予測精度を RMSE (Root Mean Square Error) を用いて確認する．さらに，学習中に発生した誤差をそれぞれの手法で算出された予測値に加えてしきい値として設定することで，異常を自動的に判定する動的しきい値法を提案し，その検知精度を確認する．

提案手法を複数の期間に適用し，その再現性や信頼性に関する評価を行う．

３．研究の方法

提案手法の予測値の予測精度と攻撃の検知精度を確認するために，複数の期間に適用し評価を行う．具体的には，2020 年 2 月 15 日 から 2020 年 3 月 12 日 の間に京都工芸繊維大学の学内と学外の間で観測された DNS パケットを 3 つの期間に分割して評価した．その結果，全ての期間において繰返し予測手法の RMSE の値（誤差）が一括予測手法よりも小さく，高い精度であることを示した．

４．研究成果

(1) 提案手法の概要

予測をベースとする DNS への異常検知手法を構築するために，学習を行う学習区間と，予測を行う予測区間を設ける．学習区間では，パケットの取得間隔を 30 分とし，観測された 48 点分のパケット数を入力値として LSTM に入力し，その 1 点先のパケット数と出力値の誤差が最小になるように学習を行う．予測区間は，48 点先のパケット数を一度に予測する一括予測手法と，パケット数が得られるたびに 48 点の予測値を繰返し算出する繰返し予測手法の二つを採用し，その予測精度を RMSE を用いて比較する．なお，RMSE は値が小さいほど誤差が小さいことを示す．

さらに，学習区間で得られた観測値と予測値の誤差を RMSE によって算出し，その誤差を予測区間で算出した予測値の上限と下限に設定することで，動的にしきい値を算出する動的しきい値法を適用する．

(2) 予測精度の比較

実験で用いるデータは，2020 年 2 月 15 日から 2020 年 3 月 12 日の間に京都工芸繊維大学の学内と学外の間で観測された DNS パケットを 3 つの期間に分割して使用した（表 1）．それぞれの期間に対して一括予測手法と繰返し予測手法を適用した結果，全ての期間において繰返し予測手法の RMSE の値（誤差）が一括予測手法よりも小さく，高い精度であることが示された（表 2）．

表 1: 学習区間と予測区間

期間番号	学習区間	予測区間
第 1 期間	2020/2/15 - 2020/2/22	2020/2/23
第 2 期間	2020/2/24 - 2020/3/2	2020/3/3
第 3 期間	2020/3/4 - 2020/3/11	2020/3/12

表 2: RMSE の比較結果

期間番号	RMSE（一括予測手法）	RMSE（繰返し予測手法）
第 1 期間	3.54×10^4	$\mathbf{1.77 \times 10^4}$
第 2 期間	4.29×10^4	$\mathbf{1.82 \times 10^4}$
第 3 期間	6.29×10^4	$\mathbf{2.23 \times 10^4}$

(3) 動的しきい値法の検知精度
表 2 の中で，RMSE が最も小さくなった第 1 期間に対して予測精度が優れた繰返し予測手法を用いた動的しきい値法を適用し、攻撃の検知精度を確認した（図 1）．その結果，48 点のパケット数のうち 10 点が異常と判定され，このうち 3 点は，オープンリゾルバを探索する突発的なスキャン行為だと思われるもので，その他は誤検知だった．

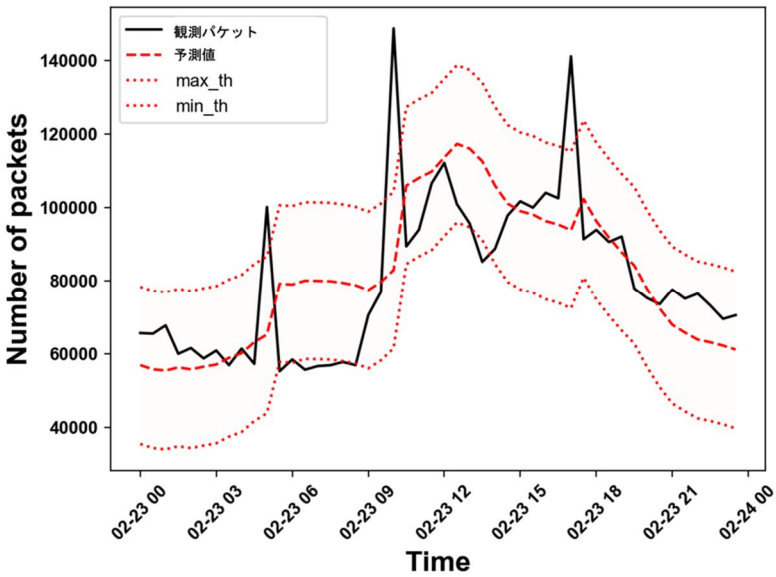

図 1: 予測区間 2020/2/23 における動的しきい値法

(4) 今後の方向性
本研究により，提案手法が複数の期間にわたって適用可能であり，繰返し予測手法が一括予測手法に対し高い予測精度を有することを明らかにした．また，動的しきい値法により，誤検知が発生しつつも，オープンリゾルバを探索する突発的なスキャン行為の検知が可能であることを明らかにした．

今後の研究の方向性として，所属している東京医療保健大学の DNS パケットを半年から 1 年の単位で長期間にわたり収集する予定である．得られた DNS パケットに対して提案手法を適用することで，施設における提案手法の予測値の予測精度や攻撃の検知精度を確認する．さらに，施設により予測精度や攻撃の検知精度が異なる場合，LSTM のパラメータや入力する特徴量の変更等を改良し提案手法の一般化可能性を高める予定である．

５．主な発表論文等

〔雑誌論文〕　計0件

〔学会発表〕　計1件（うち招待講演　0件／うち国際学会　0件）

１．発表者名 木村知史，稲葉宏幸
２．発表標題 深層学習を用いた予測に基づくDNSへの異常検知手法の評価
３．学会等名 2022年電子情報通信学会総合大会
４．発表年 2022年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

―

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関