G-11

単純マルコフ連鎖と再帰型ニューラルネットワークを用いた訓練用標的型メール文章自動生成システムの開発

A Development of an Automatic Generation System for Training-Purpose Targeted Emails Using Simple Markov Chains and Recurrent Neural Networks.

○山口智誠[1]，五味悠一郎[2]
*Tomonari Yamaguchi[1], Yuichiro Gomi[2]

Abstract : The purpose of this study is to demonstrate that automatically generated training emails can support organizations in conducting effective targeted email training. In this paper, we developed automated email generation systems based on a simple Markov chain and a recurrent neural network. Functional tests confirmed that both systems successfully generated training-targeted email texts.

１．はじめに

サイバー攻撃の 1 つとして，標的型攻撃がある．組織に対する脅威として,「機密情報等を狙った標的型攻撃」の項目は, IPA の情報セキュリティ 10 大脅威に 2016 年から 10 年間連続で選出されており, 優先的に対策を行うべき攻撃の 1 つである[1]．標的型攻撃において，標的組織のネットワークへの初期侵入時にメールを用いて行われる攻撃手法を標的型メール攻撃という．

標的型メールは，情報窃取を目的として，ごく少数または多数ながら特定された範囲のみに対して送られ，利用者の PC をマルウェアに感染させることや，フィッシングサイトを介して情報を入力させることを目的としている．特徴としては，送信者の詐称や URL クリックまたは添付ファイルダウンロードを促すメール内容などが見られる．標的型メール攻撃による被害は，システムによるセキュリティのみでは完全に回避することが難しく，人的セキュリティも併せて行うことが重要である．

標的型メール攻撃への対策の 1 つとして，標的型メール訓練が挙げられる．しかし，標的型メール訓練を行うサービス費用の問題から，企業が標的型メール訓練を積極的に行えない状況にある．また，既存の訓練サービスに頼らず，独自に標的型メール訓練環境を構築する方法も考えられるが，技術的な知識やノウハウが求められるため困難である場合が多い．そこで本研究では，自動生成した訓練用標的型メールおよび訓練用標的型メール配信システムを開発し，企業が簡単に標的型メール訓練を行える状況を作ることで問題解決を図る．

２．目標

本稿では,「単純マルコフ連鎖を用いた訓練用標的型メール自動生成システム」と「再帰型ニューラルネットワークを用いた訓練用標的型メール自動生成システム」を開発し，正常に訓練用標的型メールを生成できることを明らかにする．

３．単純マルコフ連鎖を用いた訓練用標的型メール自動生成システム

自動生成プログラムは Python3.6 で開発し, 形態素解析エンジンは Janome を用いた．Janome は固有名詞の判定とマルコフ連鎖で自動生成する際に，学習データであるメール文章を分析するために使用した．マルコフ連鎖を実装するためのライブラリには markovify を用いた．自動生成プログラムではマルコフ連鎖を使用し，形態素解析された自作メールから文章を組み立てて文章自動生成した．開発したシステムを「訓練用標的型メール自動生成システム Ver.1.0」と命名した．

４．再帰型ニューラルネットワークを用いた訓練用標的型メール自動生成システム

Googlecolaboratory を用いて, RNN/LSTM を使った訓練用標的型メールの自動生成プログラムを開発した．生成の元となる学習データは，GoogleDrive に設置し，ファイルのパスを指定することで参照した．本プログラムは，janome tokenizer を使って文章を単語ごとに分割し，LSTM モデルの入力データとして使えるように処理した後，ニューラルネットワークライブラリの Keras を用いて LSTM モデルを構築し，LSTM モデルの学習によって文章を生成している．構築した LSTM モデルのパラメータを Table 1 に示す．開発したシステムを「訓練用標的型メール自動生成システム Ver.2.0」と命名した．

1：日大理工・学部・情報　2：日大理工・教員・情報

Table. 1 LSTM model parameters

パラメータ	設定値	説明
入力シーケンス長 [maxlen]	4	モデルが考慮する過去の単語数
語彙サイズ [len(unique_words)]	可変	学習データの中で登場するユニークな単語の数
LSTMユニット数 [units]	128	LSTMの隠れ層のニューロン数(記憶容量に影響)
出力層の活性化関数	softmax	次の単語の確率を出力するための関数
最適化アルゴリズム [optimizer]	RMSprop(lr=0.01)	重みを更新するアルゴリズム
損失関数 [loss]	Categorical_crossentropy	多クラス分類の誤差を計算する関数
学習回数 [epochs]	1以上の整数で設定	モデルの学習を繰り返す関数
バッチサイズ [batch_size]	128	一度に処理するデータ数

５．機能テスト

先行研究で作成した自作標的型メール文章を学習データとして，単純マルコフ連鎖を用いたシステムと，再帰型ニューラルネットワークを用いたシステムの機能テストを行った[2]．各生成手法における機能テストでは，学習データである自作標的型メール文章 280 通を，テキストファイル形式で指定して，プログラムを実行した．なお，「再帰型ニューラルネットワークを用いた訓練用標的型メール自動生成システム」では，生成する訓練用標的型メールの初めの 4 単語は，自作標的型メールで使われた「いつも」「お世話」「に」「なって」とし，学習回数を 1000 回として生成を行った．機能テストでは，学習データを認識し，正常に訓練用標的型メールを生成できることを確認した．

６．結果と考察

Figure 1 に単純マルコフ連鎖を用いて生成したメール文章の例, Figure 2 に再帰型ニューラルネットワークを用いて生成したメール文章の例を示す．

それぞれの訓練用標的型メール自動生成システムにおいて，正常に訓練用標的型メールを生成することができた．これらの結果から，「単純マルコフ連鎖を用いた訓練用標的型メール自動生成システム」および「再帰型ニューラルネットワークを用いた訓練用標的型メール自動生成システム」は実用に耐えうるといえる．

拝啓。
配達員が注文番号数値の商品をお預かりしていただいた電話登号に誤りがあったことが分かりましたが連絡がつかなかったため、持ち帰りました。
このメールに添付されている委託運送状を印刷して、配送確認をお願い致します。
従ってご注文時に入力して、最寄りの組織名取り扱い郵便局までお問い合わせください。
敬具

Figure 1. Email text generated using a simple Markov chain.

この度は×××へご注文をいただき誠にありがとうございます。
商品を発送いたしましたのでお知らせいたします。
出荷した商品は配送伝票番号と出荷日が受注当日以内にお揃えが出来ない場合、その理由と出荷予定日を案内しています。
必ず添付の「注文商品内容のご案内」でご確認ください。
この度は×××へご注文をいただきありがとうございます。

Figure 2. Email text generated using a recurrent neural network.

７．まとめと今後の課題

「単純マルコフ連鎖を用いた訓練用標的型メール自動生成システム」および「再帰型ニューラルネットワークを用いた訓練用標的型メール自動生成システム」を開発し，それぞれのシステムが実用に耐えうることを確認した．

今後は，生成される訓練用標的型メール文章の品質を評価し，標的型メール訓練として効果的な文章の自動生成を目指す．

謝辞

本研究にご協力いただいた丹野隆裕様および釜田諒大様に感謝申し上げます．

８．参考文献

[1] 独立行政法人情報処理推進機構(IPA): 情報セキュリティ 10 大脅威 2025. https://www.ipa.go.jp/security/10threats/10threats2025.html, 参照 2025-09-23.

[2] 釜田諒大，五味悠一郎，“N 階マルコフ連鎖を用いて自動生成した訓練用標的型メールの比較と評価”，第 67 回日本大学理工学部学術講演会，2023.