

Distant Supervision によるサイバーセキュリティ文書のマルチラベル分類 Multi-label Classification of Cybersecurity Text with Distant Supervision

石井 将大 * 森 健人 * 桑名 亮一 * 松浦 知史 *
Masahiro Ishii Kento Mori Ryoichi Kuwana Satoshi Matsuura

キーワード Distant Supervision, テキスト分類, マルチラベル分類, セキュリティインテリジェンス

あらまし

近年の高度かつ複雑に進化するサイバーセキュリティ攻撃・脅威に対抗するために、セキュリティインテリジェンスを含む多様な形式の大規模なデータの詳細な分析が必須である。特に、自然言語で記述された非構造化データに対するサイバー攻撃・脅威の分類、あるいはセキュリティインテリジェンスの抽出を行う高精度なモデルが必要である。

本研究では、インシデント対応のコスト低減を目指した、サイバーセキュリティ文書のマルチラベル分類を行う。セキュリティインシデントの詳細な分析にはセキュリティインテリジェンスの抽出とそれらの関連性を活用したイベント抽出タスクなどを行う統合的なモデルが必要である。本研究はその前段階として、MITRE ATT&CK¹, Common Attack Pattern Enumeration and Classification (CAPEC)²に関する標準的なサイバーセキュリティの攻撃・脅威モデルに対する文書レベルのマルチラベル分類を行う。さらに、モデルの高精度化のための大規模教師データの作成コストの低減は本質的な課題であるため、distant supervision [1] による教師ラベル付与の自動化を行う。教師ラベルの付与については、定められた分類カテゴリに関する文書から得られるキーワードの抽出方法や複数のラベル付与規則の比較検討を行う。データセットとしてソーシャルニュースサイトやセキュリティベンダによって投稿される脅威レポートやブログ記事などによるサイバーセキュリティ文書を利用する。これらの文書に対し、BERT を始めとした事前学

習済みモデルによって得られる汎用的な文書分散表現と与えられた教師データを用いてマルチラベル分類モデルを学習し、分類精度と各カテゴリに対する分類結果を報告する。特に、distant supervision による教師ラベル付与の各手法において、分類モデルの高精度化とインシデント対応のコスト低減に寄与する特徴について議論する。

参考文献

- [1] Mintz, M., Bills, S., Snow, R. and Jurafsky, D.: Distant supervision for relation extraction without labeled data, *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, Suntec, Singapore, Association for Computational Linguistics, pp. 1003–1011 (2009).

* 東京工業大学, 〒 152-8550 東京都目黒区大岡山 2-12-1, Tokyo
Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo,
152-8550 Japan

¹ <https://attack.mitre.org>

² <https://capec.mitre.org>