**Individual In-depth Report**

**Member name:** [Anuranjan Dubey](mailto:adubey37@asu.edu)

**Evaluated by:** [Sangeeth Santhosh](mailto:ssantho9@asu.edu)

**Date:** 09/16/23

**Tasks Assigned:**

* Literature review for Machine learning techniques applied in: “Weakly Supervised Extraction of Computer Security Events from Twitter”
* Write In-depth Report while comparing it to methods seen before.
* Review and Evaluate 1 other person's report.
* Prepare Gantt Chart.

**Summary:**

* Baseline algorithms included One-Class SVMs and Expectation-Maximization (EM)
* The paper compared several algorithms, including One-Class SVMs and Expectation Regularization, with the latter showing better precision and recall.
* Specific algorithms beyond the baselines were also used for tasks like NLP and entity extraction. Expectation regularization was a proposed approach that outperformed the baseline methods.
* A bag of tweets mentioning each seed event is collected.
* Seed instances are historical examples (Entity, Date) of events used to train a system to detect new, similar events in real-time Twitter data.
* Extracting candidate events involves Providing historical seed examples (e.g., (Spamhaus, 3/18/2013)), collecting tweets mentioning seed events, tracking relevant keywords (e.g., "hacked" for account hijacking), extracting named entities and tweet dates, filtering candidate events based on seed examples.
* The paper's results showed that the proposed weakly supervised approach outperformed previous methods in terms of precision and recall for security-related events.
* Label regularization and other techniques address challenges in learning from limited positive seeds and unlabeled events.
* Conclusions highlighted Twitter's value as a resource for security event information and the effectiveness of the weakly supervised seed-based approach.
* Specific accuracy percentages are not mentioned in the provided text.
* In summary, the paper used machine learning to enhance security content classification on Twitter, with the Expectation Regularization approach showing promise. Seed instances were used for training, and the proposed method outperformed previous approaches for specific security events.

**Outcome:**

The proposed weakly supervised approach outperforms previous methods, including heuristic negatives, semi-supervised EM, and one-class SVMs. Significant improvements in precision and recall are seen for security-related events like DoS attacks, data breaches, and account hijacking. The weakly supervised seed-based approach allows rapid training of event extractors with minimal supervision.

**References** *(with citation)*

[1] A. Ritter, E. Wright, W. Casey, and T. Mitchell, “Weakly supervised extraction of computer security events from twitter,” Florence, Italy: International World Wide Web Conferences Steering Committee, 2015, pp. 896–905. doi: https://doi.org/10.1145/2736277.2741083.

**Evaluation of Report**

**Evaluation summary with justification.**

This research paper successfully introduced the Expectation Regularization Approach by conduction of comparison between different machine learning algorithms on Twitter. A clear extraction procedure was highlighted while showcasing the benefits of employing historical seed cases for training. Precision and recall were emphasized, and the usefulness of the data obtained from Twitter was acknowledged, thus emphasizing the importance of the study.

**The quality of the major result(s) with justification.**

The paper stresses on the importance of weakly supervised learning with the Expectation Regularization approach.

**The usefulness of the paper to the overall project.**

The paper is very much useful to the overall project, because the algorithms discussed could be used for the process of detecting suspicious activity on social media platforms.

**Other comments**

None

**Evaluation Approval  
  
Evaluation by:** [Sangeeth Santhosh](mailto:ssantho9@asu.edu) **Date:** 09/17/23

**Is the written report of the in-depth study complete with all the major result(s) of the paper(s)? If not, provide as many examples of the major result(s) missing in the written report as possible. (in bullet form). [Normally within 100 words]**

Yes, the written report of the in-depth study is complete with all major results of the paper emphasized.

**Is each section of the guidelines sufficiently completed? If not, point out what is missing. [Normally within 40 words].**

Yes, each section of the guidelines is sufficiently completed.

**Is the quality of this version of the written report satisfactory? If not, then why not? [Normally within 40 words]**

Yes.

**Approval.  
  
Approved by:** [Krupaben Kothadia](mailto:kkothadi@asu.edu) **Date: 09/18/2023  
  
Is the quality of this written in-depth study report and Evaluation report satisfactory? If not, then why not? (limit: 40 words)**

Yes the quality of the in-depth study report is satisfactory, as the approach is discussed clearly. Also, the evaluation report is satisfactory as the content mentioned, denotes that the evaluator has properly understood the study report.