

O-007

プッシュ型配信を活用したサイバーパトロール活動促進システムの検討 Push-Based Information Distribution System for Encouraging Cyber Patrol Activities

横山 未有[†] 角田 裕[‡]
Miyu Yokoyama Hiroshi Tsunoda

1. はじめに

スマートフォン等の普及に伴い、Social Networking Service (以下、SNS) に起因する児童被害が社会問題化している。対策として、各都道府県警察に登録されたサイバー防犯ボランティアにより、SNS 上の有害投稿を SNS 事業者や警察に通報するサイバーパトロール活動が行われており、筆者らもこの活動に参加している。ただし、ボランティア参加者は限られた活動時間で多くの有害投稿を発見する必要があり、活動の効率化が必要とされている。サイバーパトロール活動の効率化を目指す先行研究としては、池辺らによる活動支援システムの開発[1][2]などがある。これらの研究では、情報収集の効率的な枠組みを構築し、Web システムとして活動参加者に提供することで活動支援を行なっている。我々は、煩雑な作業の自動化による効率化と共に、参加者のスキマ時間での活動を促すプッシュ型配信の活用に着目した。本論文では、SNS 上で自動でキーワード検索を行い、有害と思われる投稿をプッシュ配信して情報の違法性・有害性の最終的な判別を参加者に依頼するサイバーパトロール活動促進システムの設計と試作結果について報告する。

2. システムの設計

2.1 システムの概要

図 1 に従来の活動の流れを示す。従来の活動では、参加者それぞれが SNS にアクセスして検索を行い、有害な投稿を探して通報を行っていた。この方法では、投稿の検索や通報の操作が煩雑で、まとまった活動時間が取れないと活動が難しいという問題点がある。また、同じ投稿を重複して通報してしまう、個人で活動するためモチベーションを保ちにくいなどの課題もある。

図 2 に本システムの概要を示す。本システムはまず、自動で投稿検索・収集を行い、収集した投稿を参加者それぞれの分に分けてプッシュ配信する。そして、プッシュ配信された投稿に対し参加者が有害か無害かの判別を行うと、有害と判定された投稿が自動で通報される仕組みになっている。本システムにより、煩雑な検索や通報の作業が自動化され、参加者はプッシュ配信された情報を見て判別を行うだけで良いので、まとまった活動時間が取れない場合にもスキマ時間で活動することが可能になる。また、投稿を各参加者に分配することで通報の重複を避け、プッシュ配信で依頼する形を取ることで参加者に活動の動機づけを与えることを狙う。

[†] 東北工業大学大学院工学研究科 Graduate School of Engineering, Tohoku Institute of Technology

[‡] 東北工業大学工学部情報通信工学科 Department of Information and Communication Engineering, Tohoku Institute of Technology

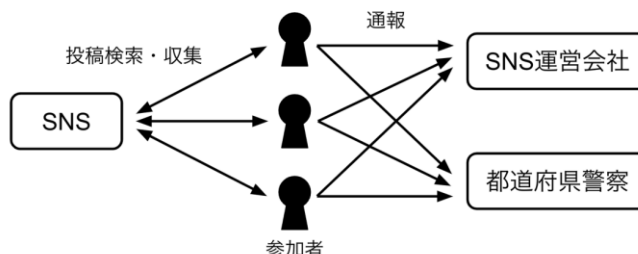


図 1 従来の活動の流れ

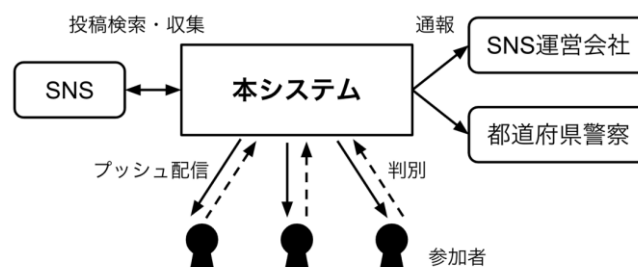


図 2 活動促進システムの概要

2.2 各機能の設計

図 3 に本システムのブロック図を示す。本システムのメインとなる機能は、(1)投稿収集機能・(2)プッシュ配信機能・(3)自動通報機能である。

(1)投稿収集機能では、キーワード検索により SNS 上から有害と思われる投稿を収集する。ただし、有害な投稿では隠語として一般的な単語を用いることが多いため、収集される投稿には実際には有害ではない投稿も多く含まれている。例えば、援助交際を意味する隠語として「円」があるが、これをキーワードとして検索すると為替相場につい

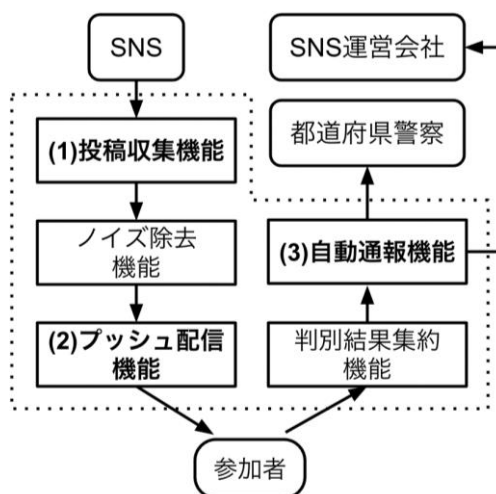


図 3 活動促進システムのブロック図

での投稿なども多く収集されてしまう。そのため、ノイズ除去機能により明らかに有害ではない投稿を取り除く。

(2)プッシュ配信機能では、投稿データを見やすい形に成形し、各参加者へプッシュ配信する。参加者が判別結果を返すと、判別結果集約機能により結果が集約される。

(3)自動通報機能では、集約された判別結果から通報すべき投稿を抽出し、SNS 運営会社と警察へ自動で通報する。

3. システムの運用試験

3.1 試験実装

運用試験を通じたシステムの効果測定と課題抽出のために、基本機能である投稿収集機能・プッシュ配信機能・自動通報機能と、判別結果集約機能を試験実装した。なお、運用試験は Twitter を対象に行うこととした。

(1)投稿収集機能では、Twitter の公式 API を利用してキーワード検索を行い、有害投稿の収集を行う。検索に用いるキーワードには、児童被害に関わる有害投稿内によく見られる単語・ハッシュタグ 10 個を利用する。

(2)プッシュ配信機能では、Google Apps Script[3] (以下、GAS) により、Google フォームを利用した投稿判別フォームを作成し、Gmail を利用して判別フォームの URL を記載した判別依頼メールを作成して参加者へ送信する。判別フォームには投稿の内容と URL が 10 件ずつ記載されており、判別結果を選択肢から 1 つ選択する形になっている。選択肢は、有害(県内)・有害(県外)・無害・削除済み・わからないの 5 つとした。有害の項目を県内と県外に分けたのは、県内情報のみを都道府県警察に通報する必要があるためである。また、有害投稿は SNS 運営会社により削除される場合もあるので、すでに削除された投稿を通報しないように削除済みの選択肢を設けた。フォームの回答が送信されると、判別結果集約機能により Google スプレッドシートに結果が集約される。プッシュ配信は、毎日午後 5 時に行う。

通報の方法について、都道府県警察への通報はメールで行うこととなっており、SNS 運営会社への通報は Web 上の通報フォームから行う必要がある。そのため、(3)自動通報機能では、警察へのメール通報には GAS と Gmail を利用し、SNS 運営会社の通報フォームへの通報には筆者らが以前開発した自動通報システム[4]を利用する。ただし、試験実装ではメールによる通報のみを実装し、宛先は研究室で利用しているメールアドレスとした。

3.2 運用試験

運用試験として、試験実装したシステムを 10 日間に渡って筆者らの所属する研究室の学部 4 年生 8 名、学部 3 年生 1 名の合計 9 名に利用してもらった。運用試験終了後に事後アンケートを実施したところ、プッシュ配信の方法と時間帯に対する回答は図 4 のようになった。配信方法については、「気付きやすかった」との回答が 89%であるため適切であったと考えられる。ただし、自由記述欄では「LINE の方が気付きやすい」との意見も挙がった。配信時間帯については、「どちらとも言えない」との回答が半数以上を占め、自由記述欄では「自分で配信時間を設定できると良い」、「回答がない場合は翌朝にも通知が来ると良い」などの意見が挙げられた。また、回答できない日があった参加者に対して、その理由について確認した結果、「メールは確認したが回答を忘れていた」などの理由が多

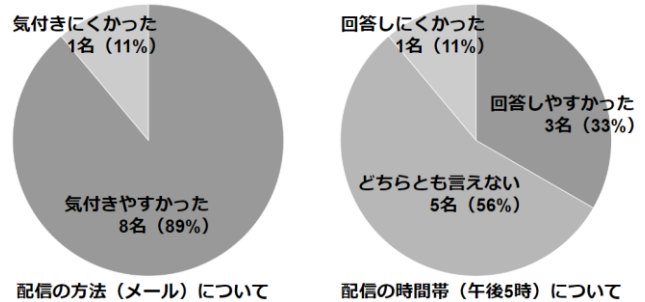


図 4 運用試験事後アンケートの結果

く、ほとんどの場合でメールの確認までは行われていることがわかった。よって、本システムにより活動のきっかけ作りができたといえる。

ただし、運用する中で、判別フォーム生成に用いる投稿データ数が増えると GAS の利用制限によりフォームの生成に失敗し、メールには前日のフォームの URL が記載されてしまうという不具合が生じることがあった。

3.3 実運用に向けた課題

実運用に向けて、試験実装した機能の他にノイズ除去機能と SNS 運営会社への自動通報機能の実装を進める。また、運用試験の事後アンケートから得られた要望に基づき、配信方法に LINE を追加する、配信時間帯を選べるようにする、一定時間回答がない場合にリマインドするなどの機能追加も課題となる。さらに、GAS の利用制限による不具合に対し、運用試験ではフォーム生成に用いる投稿データ数を少なく制限することで対処したが、実運用にあたって参加人数が増えた場合、対応が難しくなる可能性がある。そのため、利用制限のある GAS を使わず、別途専用のサーバを用意してシステムを実装・運用することも検討する。

4. おわりに

本論文では、煩雑な作業の自動化による効率化と共に、参加者のスキマ時間での活動を促すプッシュ型配信の活用に着目したサイバーパトロール活動促進システムの設計と試作結果について報告した。今後の課題としては、サイバー防犯ボランティアで扱っている児童被害関係以外の有害情報についても対応すること、Twitter 以外の SNS への対応などが挙げられる。また、活動へのモチベーションをより高めるためには、参加者へのフィードバックを行うなどの工夫も必要になる。

謝辞

本研究は、東北工業大学学内公募研究 (2022-14) の援助により行われたものである。

参考文献

- [1] 池辺 正典, “サイバー防犯ボランティアのためのサイバーパトロール支援システムについて”, 警察学論集, Vol.71, No.5, (2018).
- [2] 池辺 正典, 川合 康央, 櫻井 淳, “SNS を対象とした都道府県警察のサイバーパトロール支援システムについて”, 情報教育シンポジウム論文集, Vol.2021, (2021).
- [3] Google LLC, “Google Apps Script”, <https://workspace.google.co.jp/intl/ja/products/apps-script/>, (accessed Jun. 22, 2022).
- [4] 横山 未有, 角田 裕, “サイバー防犯ボランティアの通報活動を効率化するシンプルな GUI アプリケーション”, 2021 年度電気関係学会東北支部連合大会, 3C04, (2021).