

問題投稿に付帯されるハッシュタグの特徴 —売買春を事例として—

吉見憲二†

上田祥二‡

針尾大嗣††

佛教大学†

(株) セールスフォース・ドットコム‡

摂南大学††

1. はじめに

スマートフォンの普及を背景に、ソーシャルメディアの利用が一般化、低年齢化していくに伴い、子どもたちが直面するリスクも増大してきている。警察庁の調査では、出会い系サイト規制法施行以降、出会い系サイトに起因する事犯の被害児童数が減少し、過去最少を記録している一方で、SNS 等における被害児童数は増加傾向が継続し、過去最多を迎えていることを指摘している[1]。また、被害児童の SNS へのアクセス手段は約 9 割がスマートフォン経由であること、Twitter における被害児童数は平成 27 年から平成 28 年にかけてほぼ倍増しており、平成 29 年には全被害児童数の 1,813 人のうち Twitter による被害が 3 分の 1 以上の 695 人に達していることも調査から示されている[1]。このまま有効な対策が打たれなければ被害数はさらに増大していくことが予想されるため、事業者や行政の連携によって有効な対策が行われる必要がある。

本研究では、青少年を取り巻くリスクの中でも近年最も深刻な社会病理問題となっている「売買春」に焦点をあて、これらに関連する Twitter の問題投稿とそれに利用された一連のハッシュタグについて調査を行った。

2. 先行研究

サイバー補導とは「インターネット上で援助交際を求めるなどの不適切な書き込みをした未成年者に対し、注意や指導をすること（出典：デジタル大辞泉）」であり、2013 年 10 月より全国の警察で実施されている。「サイバー補導の現状と今後の取組について」（2015 年 3 月 12 日広報資料）では試行期間である 2013 年 4 月から 2014 年末までに 18 歳未満の児童 597 人が補導され、その 6 割以上に非行・補導歴がなかったことを報告している[2]。

全国 12 か所の都道府県警察本部少年課を対象とした聞き取り調査からは、サイバー補導の課題として「効率よく検索できる仕組み」が挙げられている[3]。そもそもサイバー補導における交信作業はさまざまな制約から「不適切な書き込みを発見し、現場での接触にこぎつけられるのは、ほんの数%程度である」という難しさがある。その中で、「ネット上にあふれる交流サイトの掲示板の中から、不適切な書き込みを発見する作業は、容易ではない。変化を続ける掲示板等のサイトの情報を、取りまとめてくれる人材と、情報共有できる仕組みが欲しい」というのが現場からの声となっている。警察に代わって対象者との交信作業を行うことは現実的ではないが、サイバー補導の対象となる書き込みを「効率よく検索できる仕組み」についてはシステムの解決が期待できるものである。

上記のような社会背景や先行研究を踏まえて、本研究では Twitter 上の投稿内容の分析から、より売買春に関する問題投稿の蓋然性が高いハッシュタグの構成を明らかにすることを目的とする。

3. 調査・分析方法

2018 年 5 月 2 日～31 日の 30 日間の Twitter 投稿を対象に、明確に「売買春」を示唆していると考えられる 4 つのハッシュタグ（「援助」「援」「円」「サポ」）を含んだ投稿を収集し、同時に用いられている（共起関係にある）ハッシュタグの特定を試みた。

こうしたハッシュタグ間の関係性に着目した研究として、東日本大震災後の Twitter における震災関連ハッシュタグの計量的分析を行い、話題のピークや中心的な話題の変化について検証しているものがある[4]。当該の研究では、ハッシュタグの共起関係について、共起ネットワーク分析を用いて可視化しているため、同様の手法を採用した。

4. 分析結果

期間中に取得した総投稿数は 120,020 件とな

Characteristics of Hashtags Incidental to Problem Posting - Case of Prostitution -

† Kenji Yoshimi, Bukkyo University

‡ Shoji Ueda, Salesforce.com, Ltd

†† Daiji Hario, Setsunan University

った。毎日 3500 件以上もの問題の疑いのある投稿が観察され、そのうち RT（リツイート）は 1000 件から 1500 件程度であった。登場したハッシュタグは 5,045 種類となった。出現数上位のハッシュタグは、「援」「サポ」「円」「オフパコ」「援助」「裏垢女子」「家出」となり、いずれも 1 割以上の投稿に当該ハッシュタグが含まれていた。「パパ活」や「神待ち」といった通常のコミュニケーションでは用いないようなハッシュタグの利用も見られた。

また、「都内」「池袋」「大阪」といった地名が出現数上位に見られている。付帯ハッシュタグ数の平均値は 4.8 個、中央値は 4 個であり、複数のハッシュタグが同時に用いられる傾向にあった。10 個以上のハッシュタグが用いられている投稿も少なくなかったが、こうした極端にハッシュタグが多い投稿の詳細を確認すると、業者によるものと疑われる投稿も散見された。

なお、ハッシュタグではないものの、本文に「DM（ダイレクト・メッセージ）」という単語を含む投稿は 28,122 件で全体の 23.4%に達していた。

共起ネットワーク分析の結果は、図 1 と図 2 の通りである。共起ネットワーク分析は描写の基準に応じて結果が変わるため、Jaccard 係数（以下、J 係数）の異なる結果を示している。図 1 と図 2 からは、援助交際を示唆する「援」「円」「サポ」といったハッシュタグが強い共起関係にあることが明らかとなっている。このことは複数の隠語が同時に用いられていることで、その意味を補強していることを示している。こうした事実は去の新聞報道等でも言及されており、改めてその事実が確認された。その他には、地名同士の関係性の強さや特定のハッシュタグ間の共起頻度の多さも見られている。ただし、後者については業者アカウントが繰り返し同様の表現を用いていることから、結果として共起頻度が高まっているものも含まれている。

5. まとめ

本研究では、「売買春」に関連する Twitter の問題投稿とそれに利用されたハッシュタグについて調査・分析を行った。テキストマイニングを用いた付帯ハッシュタグの分析から、地名や隠語を含めた複数のハッシュタグが同時に利用されている傾向があり、こうしたハッシュタグを特定することで問題投稿の事前検知に有効であることが示された。

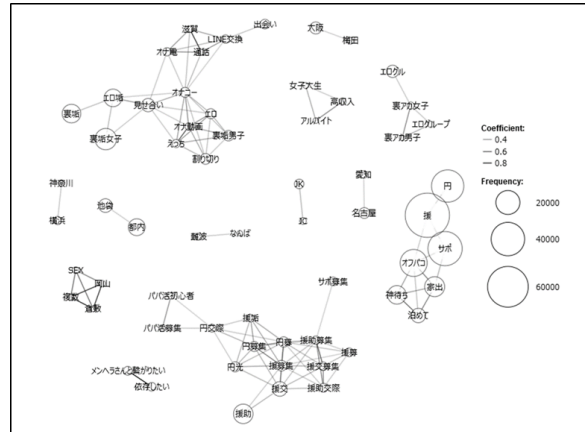


図 1 ハッシュタグ間の共起関係 (J 係数=0.2)

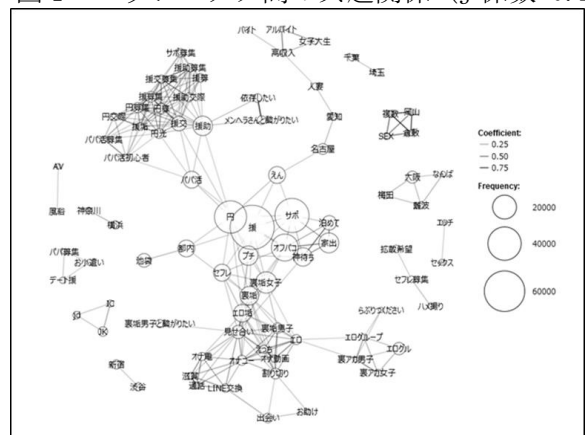


図 2 ハッシュタグ間の共起関係 (J 係数=0.1)

謝辞

本研究は公益財団法人電気通信普及財団の研究助成（「SNS 上の発生リスクとデータリンク構造（代表：針尾大嗣）」）を受けたものである。

注

本稿の内容は 2018 年社会情報学会 (SSI) 学会大会における報告内容「問題投稿に付帯されるハッシュタグの特徴—売買春を事例として—」を加筆・発展させたものである。

参考文献

- [1] 警察庁 (2018) 「平成 29 年における SNS 等に起因する被害児童の現状と対策について」
- [2] 警察庁 (2015) 「サイバー補導の現状と今後の取組について」
- [3] 矢作由美子 (2014) 「サイバー補導の現状と課題：不良行為少年の補導活動の拡大化」『早稲田大学社会安全政策研究所紀要』第 7 号, pp.113-140
- [4] 村井源 (2012) 「東日本大震災後の Twitter 利用傾向 -震災関連ハッシュタグの計量的分析-」『情報知識学会誌』22 巻 2 号, pp.97-106