

見守りネットパトロールシステムの構築

ソリューション本部ソリューションサービス部

太田 桂吾

1. はじめに

「秋葉原通り魔事件」(2008年6月8日)をきっかけに、“ネット上の若者の書き込みの看視による、いじめ等の早期発見”を目的として、京都府立大学の吉富教授らのグループ(情報環境学グループ)が開発した「学校裏サイト等の看視支援システム」。そのシステムをベースとし、さらなる機能強化のため、京都府立大学情報環境学グループと当社との共同研究(2012年4月～2013年3月)の成果として開発された「見守りネットパトロールシステム」に関して以下に述べる。

2. システムの構成と概要

2.1 構成要素

以下の2つのモジュールから構成される。

- 1)「サイトクローラー」
- 2)「サイト看視」

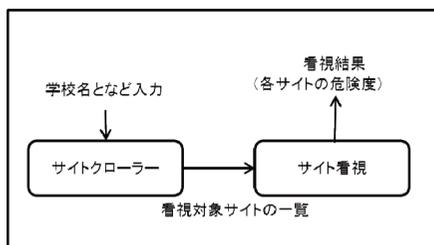


図1 システム構成

2.2 「サイトクローラー」の概要

サイトクローラーとは、いわゆる”プロフィールサイト”、“リアル”をインターネット上から検出し、サイトの一覧を出力する機能である。”プロフィールサイト”(以下プロフ)とは、インターネット上に構築できる自己紹介サイトであり、“リアル”とは携帯電話からプロフと同様に気軽に書き込めるブログサイトのようなものである。これらは若者を中心に利用されており、セットで利用されることが多い。

文部科学省の「平成21年 青少年が利用するコミュニティサイトに関する実態調査」では、「注意を要する、問題のある投稿のサイト種別」として、プロフサイトが53%、ゲームSNS系(ゲーム系SNSおよび一般的SNS)が25%、掲示板が9%、リアル7%という結果となっている。なお、ゲームSNS系は、本システムの今後の課題とし取り込む予定である。

2.3 「サイト看視」の概要

「サイトクローラー」で発見したサイトの看視を一括で行う。

看視するサイトのURLリスト取り込み、各サイトを確認し、システム内に持っている辞書とサイトの書き込み内容を比較し、各サイトの危険度を判定している。

3. 「サイトクローラー」

3.1 機能

ユーザーは以下の入力を行い、システムは、それらキーワードをもとにインターネット上のサイトの検索を行う。

- ・正式学校名(必須)
- ・学校名検索キーワード(必須)
- ・その他検索キーワード(任意)
- ・検索対象とするサイト提供会社(1つ以上のチェック 必須)

結果として、その学校の生徒が関係するプロフィールもしくはリアルを発見し、「サイト閲覧」モジュールで読み込み可能なリストを作成する。

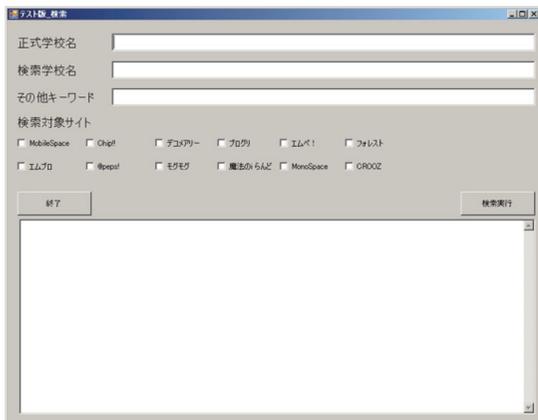


図2 サイトクローラー画面

3.2 仕組み/流れ

検索は、プログラム上で Google 検索オプションを使用した URL を作成し、Google 検索サイトに問合せることによって実施している。

ただし、検索結果そのままでは、プロフィール内の特典ページであったりし、そのまま使用できるものではない。そのため、以下のようなサイト提供会社別の URL の特徴を利用し、ユーザーID を取得

し、そこからプロフィール、リアルの URL を取得している。

プロフィール別の特徴(例)	
プロフィール提供会社MのURLパターン	
TOPページ	http://nn.****.jp/ID / (nnは任意の数字、****は本来は固定文字)
プロフィールページ	http://nn.****.jp/ID -nnnnnn-pr.php?guide=on/ (nnは任意の数字、****は本来は固定文字)
プロフィール提供会社NのURLパターン	
TOPページ	http://**.****.co.jp/i.asp?l=ID / (nnは任意の数字、**.****は本来は固定文字)
プロフィールページ	http://dp.****.co.jp/action=profile_view_main&ocd=user&oid=ID (nnは任意の数字、****は本来は固定文字)

図3 サイト提供会社別特徴
(参考文献3より引用)

トップページから、プロフィール、リアル取得は表-1 のように、取りこぼしなく収集できている。

表-1 (参考文献3より引用)

サイト提供会社	トップページ数	プロフィールが正しく取れている数
A社	73	73
M社	17	17
D社	47	47
Z社	3	3
P社	4	4

また効率的にサイトを収集するため、収集したサイトのリンクページ(友人のトップページ等をまとめたページ)から、多くのサイトを収集する。

なお、収集したサイトが、対象の学校の関係者(主に生徒)の作成したものかどうかは、以下の内容で判定する。

(1) プロフィールの内容

(以下はサンプルです)

HN(ハンドルネーム):○○○

誕生日:○○年○○月○○日

学校:○○学校○○年○○組

- (2) リンク時の紹介文
(以下はサンプルです)

Myリンク

OOちゃん OO学校でーす

4. 「サイト看視」

4.1 機能

(1) 看視機能

ユーザーが「サイト看視」モジュールを実行すると、看視対象サイトの URL リスト(「サイトクローラー」モジュールで収集、もしくはユーザー自身でも設定可能)を読み込み、サイトの書込内容を判定する。

判定の結果、書込内容の危険度が高い順にサイトをソートし、その危険度とともに表示する(初期では、最も危険度が高いサイトが表示される)。危険度は、「レッドカード(非常に危険、この中でさらに3段階あり)」「イエローカード(危険)」「ブルーカード(危険なし)」に分かれる。

(2) ハイライト機能

危険と判定した、いじめ語等をハイライト表示する。

(3) 履歴管理

一度看視したサイトは、以降の看視で新たな書込(更新分)のみを確認する。

4.2 仕組み/流れ

サイトの内容を MeCab(オープンソースの形態素解析エンジン)に入力して形態素解析を行い、あらかじめ登録した「いじめ語」辞書を使用し、そのサイトの危険度を判定する。また「いじめ語」の付

近に人名があるかどうかも判定(「松茸用人名テキスト」記載の 80,000 語を使用)し、通報に役立てる。なお、「いじめ語」辞書には、「いじめ語」指数が登録されており、その指数により深刻さを図る。また、「いじめ語」には、「死にたい」などの言葉も含まれており、いじめの事案以外も考慮されている。

5. 効果

「サイト看視」モジュールは、2010 年9月より、京都府立大学 情報環境学グループと京都市教育委員会が連携し、モデル校として選定された学校の非公式サイト(学校裏サイト)の看視を実施し、2011 年 8 月から 2012 年 3 月の期間は、20 校のモデル校で月平均 5 件、の「問題のある書き込み」を、京都市教育委員会に通報した。通報内容は、友人、教職員、親などに対する誹謗中傷、および、自殺願望であった。

上記のことから以下の効果を認めることができた。

- (1) 看視時間削減(2%~4%へ 96~98%削減)
- (2) チェック漏れの防止
- (3) 人件費等の抑制

「サイト看視」モジュールの導入で、すべてのサイトのすべての書き込みを目視確認する作業(「昨日はどこまで確認したかな…?’等の労力)に比較し、問題がありそうなサイトの、未チェックの書き込みを優先し確認していくことで、大きな作業効率化ができたためである。

作業の効率化は、看視範囲の拡大を可能とし、いじめ等の看視に大きく貢献できると考えている。

また、「サイトクローラー」に関しては試験段階だが、上記と同等の効果が上がると考えられる。

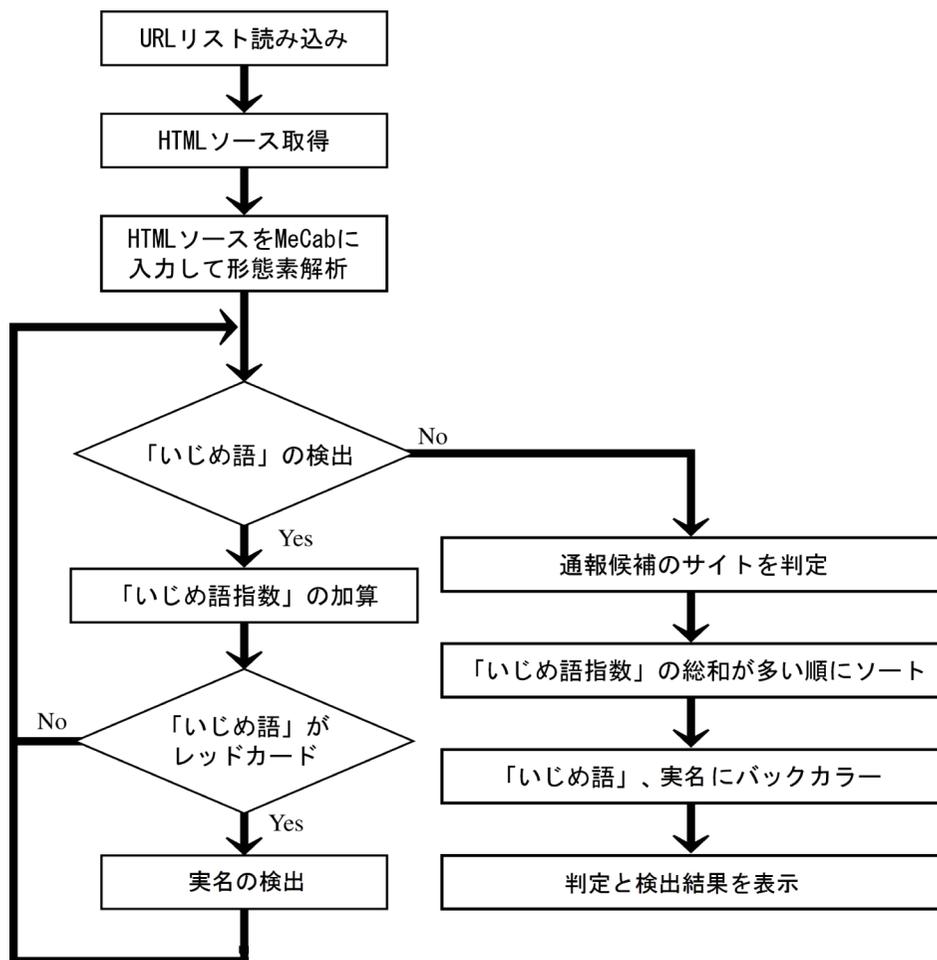


図4 サイト看視のフロー（参考文献1より引用）

6. 今後の課題

6.1 本システムの課題

さらに効率的、かつ広範囲な看視を実施するために、以下の課題へも積極的に取り組みたい。

(1) SNS サイト

Twitter/Facebook に代表されるような SNS サイトの看視。

(2) 携帯電話専用サイト

使用される比率は下がってきたが、効率的な看視方法を見つけていきたい。

(3) パーソナリティ分析

テキストマイニング技術を応用し、プロフィール、リアルに記載された内容から、その人物のパーソナリティを分析し、ユーザーの分析等に役立てたい。

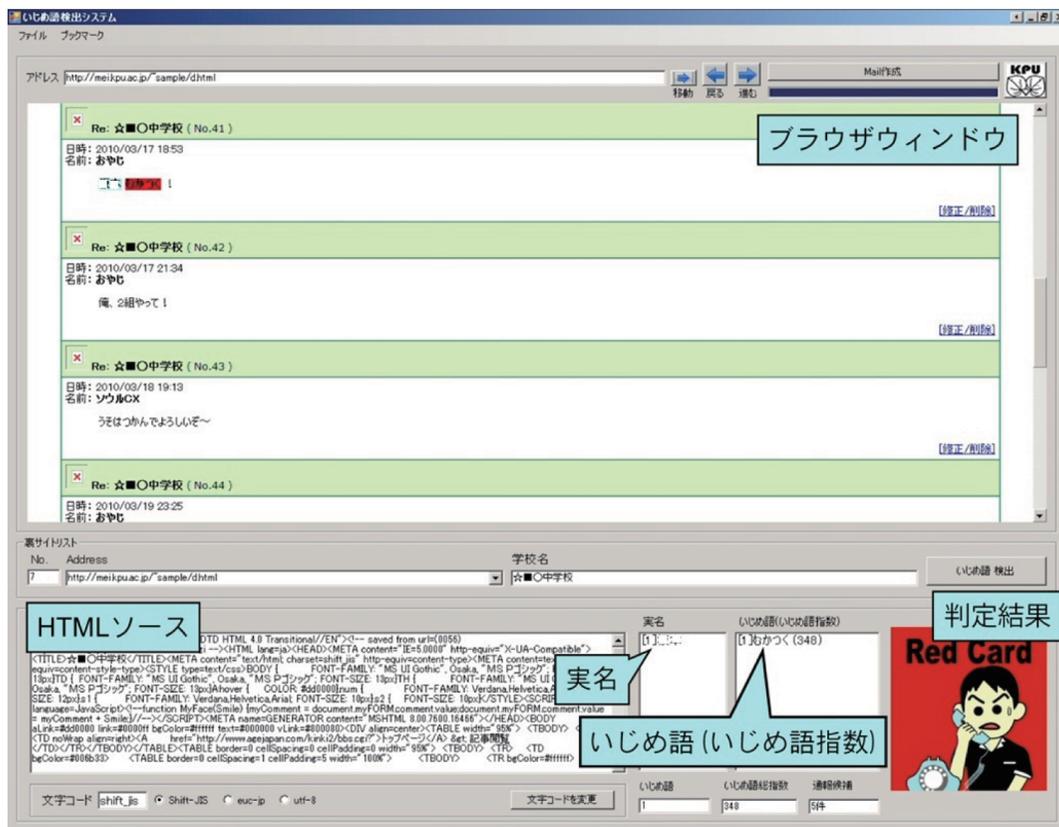


図5 サイト監視結果画面（参考文献2より引用）

6.2 本システムの機能を利用した拡張

本システムのサイトクロウリング、監視機能を利用することで、以下のような課題へも今後取り組んでいきたい。

(1) 産業向けネット監視システム

「いじめ語」辞書を、その業界の評判を表す辞書へ変更することで、比較的容易に、ネット上のその企業/製品の評価を監視することが可能であると考えている。

<参考文献>

- 「「いじめ語指数」と個人名検出を併用した学校裏サイト監視支援システム」(第10回情報科学技術フォーラム講演論文集、Vol. 3、pp. 749-750、2011)
- 「平成23年度京都府立大学地域貢献型特別研究の成果」より「A13 学校裏サイト等の監視支援システムの実用化」(京都府立大学、平成24年5月15日)

<http://www2.kpu.ac.jp/kanrika/kikaku/actr23/A13.pdf>

- 3) 京都府立大学情報環境学グループ作成資料、2012
- 4) “Development of Monitoring Support System for Unofficial School Websites”, Proc. of 1st WSEAS International Conference on Information Technology and Computer Networks, pp.164-168, 2012
- 5) 「平成21年度 青少年を取り巻く有害環境対策の推進 青少年が利用するコミュニティサイトに関する実態調査 報告書」(文部科学省、平成22年3月)

http://www.mext.go.jp/a_menu/seisyounen/houshin/1295040.htm

- 6) 松茸用人名テキスト

<http://www.vector.co.jp/soft/data/writing/se066955.html>