

インターネット上の違法・有害情報検出技術の研究開発

(1) 研究の目的

ISP やコンテンツ監視事業者が共通で利用可能な「監視業務支援システム」を研究開発し、監視業務のコストを削減すること

(2) 研究期間

平成 21 年度から平成 23 年度（3 年間）

(3) 委託先企業

(株) KDD I 研究所

(4) 研究予算（百万円）

平成 21 年度	179
平成 22 年度	164
平成 23 年度	158

(5) 研究開発課題と担当

課題ア：モバイル対応型情報収集・解析結果提供技術

課題イ：違法・有害情報の大規模コーパス自動構築技術

課題ウ：違法・有害情報の高速高精度検出技術

課題エ：ISP やコンテンツ監視事業者等の実運用環境における実証実験

(6) これまでの主な研究成果

特許出願：国内出願	16 件	外国出願	0 件
外部発表：研究論文	5 件	その他研究発表	14 件
報道発表	7 件	展示会	6 件
		標準化提案	0 件

**具体的な成果**

(1) 報道発表関連

(ア) 課題ウ：外形的特徴に基づく有害サイト検出手法について報道発表により研究成果を幅広くアピールした。

- ① 一般誌 4 誌に掲載
  - 1. 読売新聞(9月14日、8面)、毎日新聞(9月14日、27面)、東京新聞(9月14日、8面)、フジサンケイビジネスアイ(9月16日、7面)
- ② NHK 番組「おはよう日本」で紹介(10月17日)
  - (イ) 課題ア:有害サイトクローラ技術について報道発表(2月1日)を行い、Web ニュースを中心に 50 サイトに掲載されるなど、研究成果を幅広くアピールした。
- (2) 表彰関連
  - (ア) 船井ベストペーパー賞を受賞(2010年9月8日)
    - 外形的特徴に基づく有害サイト検出手法について、2010年9月に開催された情報科学技術フォーラム(FIT2010)に投稿した論文で受賞。(700件以上の論文のうちわずか3件のみ)
  - (イ) 大会奨励賞を受賞(2011年3月3日)
    - 係り受け解析を用いた高精度有害サイト検出手法について、情報処理学会全国大会で発表した研究で受賞。
- (3) 課題ア:違法・有害情報に関するPCサイト・携帯サイトのクローリング作業効率化技術に関する課題において、有害サイトらしさをURL等から自動的に推定し、それらを優先的に収集する機能を持つ有害クローラを開発した。20%の有害ページを含む10,000ページを対象とし、このうち2,000ページを収集する条件で有害ページの収集効率を計測する実験を行ったところ、一般のクローラと比較して3.5倍以上の有害ページを収集できることを確認した。
- (4) 課題ア:機械による自動判定結果の信頼性を向上するための、携帯専用サイト識別機能を開発した。評価実験の結果、91.4%の精度で携帯専用サイトを識別できることを確認した。
- (5) 課題イ:違法・有害情報の高精度な監視ツール構築のベースとなる、3億記事の大規模コーパスを構築する技術に関する課題において、機械学習向けの人手によるラベル付きのコーパス3種類を構築し、年度目標を達成した。
  - 1. カテゴリラベル付与コーパス 2665万記事(目標:2000万記事)
  - 2. 品詞ラベル付与コーパス 50万記事(目標:50万記事)
  - 3. 品詞・格構造ラベル付与コーパス 10万記事(目標:10万記事)
- (6) 課題イ:学習・評価用データとして用いる大規模コーパスをウェブページから構築するための能動学習手法を新たに提案した。従来の能動学習手法では、学習データに酷似文書が含まれると、精度向上幅が小さくな

る問題点があった。提案手法では類似文書のクラスタリングを導入し、この問題を解決した。違法有害な情報に関連するウェブ文書 100 万件を抽出して精度評価を行ったところ、精度向上が確認できた。

- (7) 課題ウ：違法・有害文書に特徴的に現れる、違法・有害なキーワードを統計的な指標に基づいて抽出し、それらの出現回数を文書の特性として、高速 SVM (Support Vector Machine) を用いて大量の文書から違法・有害文書の特性を自動的に学習することで、高精度な違法・有害文書検出を実現した。昨年度の方式と比較して、再現率を 5% 向上し、再現率 75%、適合率 90% の H22 年度目標を達成した。
- (8) 課題ウ：Web サイトの背景色や画像数などの外形的な特徴に基づき、高適合率条件下 (90% 以上) において、違法・有害サイトを高速・高再現率で検出する手法を考案した。提案手法は Web サイトの HTML タグ部分などに注目し、違法・有害サイトに特徴的に出現するような外形的特徴を発見し、それらを高速 SVM を用いて組み合わせることで判定に利用することで、単独で高い適合率の判定を可能とした。テキストベースの方式と検出可能な有害サイトの傾向が異なる点ことから、本手法をテキストベース方式の前に適用することで、処理時間を短縮すると共に、システム全体の適合率を高めることが可能であるという知見を得た。
- (9) 課題ウ：係り受け解析を用いることで違法・有害文書を高精度に検出する手法を考案した。違法・有害性の高い表現と係り受け関係にある文節の組を判定に用いることで、従来手法では判定を誤るような表現を正しく判定することが可能となった。課題イで収集したコーパスを用いた実験により、再現率と適合率がともに向上することを確認した。
- (10) 課題エ：課題イ、ウの成果を取り入れた違法・有害情報判定システムを試作した。また、CP (Contents Provider) および監視事業者向け監視ツールを試作した。今後は実際の事業者等の意見を反映した改良を行い、実証実験に活用する予定である。
- (11) 課題エ：平成 23 年度に予定している実証実験のための実験データを集めるため、監視事業者や CP などの 4 つの参加形態に分類した実験計画書を完成した。