

子供の Web アクセスにおける保護者監視支援システム

上田 達巳[†] 高井 昌彰[‡]

[†]北海道大学大学院情報科学研究科 〒060-0814 札幌市北区北 14 条西 9 丁目

[‡]北海道大学情報基盤センター 〒060-0811 札幌市北区北 11 条西 5 丁目

E-mail: [†]ueda@iic.hokudai.ac.jp, [‡]takai@iic.hokudai.ac.jp

あらまし 近年、インターネットの利用は極めて一般的なものとなってきており、子どもたちもまたインターネットから様々な情報を取得する機会が増えてきている。インターネットでは有益な情報を得ることができる反面、リンクを辿っていく事での意図しない有害な情報への接触、操作ミスによる個人情報の誤発信、掲示板などで名誉毀損的な発言をしてしまい収拾がつかなくなるといった問題に遭遇してしまう危険性がある。本稿では有害な Web コンテンツから子供たちを保護し、子供への理解を深めることを支援する保護者監視支援システムについて述べる。また、Windows での実装と動作実験について述べる。

キーワード 監視支援、子供、有害情報、Web ドキュメント

Guardian watch support system for the children's Web access

Tatsumi UEDA[†] and Yoshiaki TAKAI[‡]

[†] Graduate school of Information Science and Technology, Hokkaido University Kita 14, Nishi 9, Kita-ku, Sapporo, 060-0814 Japan

[‡] Information Initiative Center, Hokkaido University Kita 11, Nishi 5, Kita-ku, Sapporo, 060-0811 Japan

E-mail: [†]ueda@iic.hokudai.ac.jp, [‡]takai@iic.hokudai.ac.jp

Abstract The Internet is very widespread in late years, and children's acquiring information from the Internet has increased. There is danger of encountering the problem that a defamation remark is done on the mis-sending individual information by contact to not the intention by tracing the link but harmful information and the operation mistake and bulletin boards, etc. and control doesn't attach though profitable information can be obtained in the Internet. It explains the guardian watch support system that supports the protection of children in this paper from harmful Web contents, and deepens of understanding to guardian's children in addition, and mounting and the operation experiment with Windows are described.

Keyword Watch support, Children, Harmful information, Web document

1. はじめに

近年、インターネットの利用は極めて一般的なものとなっており、子供たちもまたインターネットから様々な情報を取得する機会が増えている[1].

インターネットでは、有益な情報を得る事ができる反面、リンクを辿っていく事による意図しない有害な情報との接触や、操作ミスによる個人情報の誤発信、掲示板などで名誉毀損的な発言をしてしまい収拾がつかなくなるといった問題に遭遇してしまう危険性がある。これらの問題から子供たちを保護するために、インターネット上に存在する有害なコンテンツを排除しようという動きがあるが、現段階では法的に情報発信を規制するには限界がある。

このため、受信者側で受け取る情報を選択または制限する方法としてフィルタリングが用いられている。しかし、フィルタリングには過度に情報を制限してしまい、本来有用となる情報を得られない場合があると

いう問題がある。

また、子供がインターネットに接続する場合は、操作をしている間、保護者がそばにいて監視する事が推奨されているが、常に監視し続けることは難しく子供がインターネットで何をしているかを十分に把握しきれていない場合もある。

本稿では、有害なインターネットの利用から子供たちを保護するための保護者監視支援システムについて述べる。また、Windows での実装と動作実験について述べる。

2. 背景

2.1. 子供を取り巻くインターネット環境

現在、小中高校でのインターネットへの接続率はほぼ 100% となっており、インターネットを用いた教育も盛んに行われている[2]. また、インターネットに接続している家庭において、インターネットを利用して

いる子供は3/4の割合にのぼる[3].

このように、非常に多くの子供たちがインターネットを利用していることがわかる。

2.2. インターネットが抱える問題点

インターネットは誰でも情報発信をすることが可能であるという特徴から、非常に多様で有益な情報を得ることが可能である。

しかしその反面、リンクを辿っていくことで意図せずに違法あるいは有害な情報に接触してしまう、誤操作により個人情報を発信してしまい被害を受ける、掲示板などで名誉毀損的な発言をしてしまい収拾がつかなくなってしまう、気がつかないうちに事件の被害者や加害者になってしまうなどといった問題を抱えている。また、近年インターネットから情報を得て行われる犯罪が増加しているとされ、インターネット上の違法・有害情報への対策を求める声が大きくなってきている。

家庭において子供がインターネットを閲覧する際には、保護者がそばについて監視することが推奨されているが、実際に常にそばに居続けることは難しく、子供がインターネットで何をしているのかを把握しきれていない場合もあり、保護者の多くは子供がインターネット上の有害な情報に接触していないかどうかを不安に感じている[4]。低学年の子供たちが使用するインターネット環境は違法・有害情報から隔離され、安全が確保されている状態であることが望ましいが、高学年の子供たちについては情報やメディアの扱い方を学習していく必要がある。このとき、保護者が子供の受け取っている情報を把握して情報モラルの教育に役立てることが望ましい[5]。

2.3. 違法・有害情報への対策

2.3.1. 違法な情報の規制

インターネット上の違法な情報の発信を規制するためにインターネット関連法規が施行されている。現在施行されている法規のなかで性・アダルトに関連する違法情報を取り締まるものについて表1に示す。

例えば改正風俗営業適正化法は、ISPのサーバから違法な映像が発信されないようにするために、ISPに対し努力義務を求める。さらに、インターネットを配信経路としてアダルト映像の発信を行う業者に対して都道府県公安委員会への届出を求めるという法規である。しかし、違法ではないけれども有害な映像に対してのISPの努力義務は無く、無料で映像配信を行っている業者についても本法規では適用外となる。

このように、法規により違法な情報の発信については取り締まることができるが、法規の適用外となってしまうような有害な情報の発信に対処することはできないため、他の手法を用いることが必要となる。

表1 違法情報に対する法規制の例

Table 1 Regulations system to illegal information.

改正風俗営業適正化法

(風俗営業等の規制および業務の適正化等に関する法律)

ISPが自らのサーバ上にコンテンツプロバイダが猥褻な映像または児童ポルノ映像を記録したことを知ったときは、当該映像の送信防止のための必要な措置を講ずるよう努めなければならない。インターネット上でアダルト映像を有料で送信する業者は都道府県公安委員会に届け出なければならない。

出会い系サイト規制法

(インターネット異性紹介事業を利用して児童を誘引する行為の規制等に関する法律)

インターネット異性紹介事業者に対して、利用者への18歳未満の利用禁止の伝達、利用者が児童でないことの確認を義務付ける。

児童買春・ポルノ禁止法

(児童買春、児童ポルノに係る行為等の処罰及び児童の保護等に関する法律)

児童ポルノを提供したり、公然陳列したり、これらの目的で製造・所持・運搬・輸出入したり、製造したりしてはならない。

表2 既存の有害情報定義

Table 2 Definition of harmful information.

	性	暴力	犯罪	ドラッグ	人権侵害
kids.us	○	○	○	○	
SafeSurf	○	○		○	○
CyberLIST	○	○	○	○	
SafetyOnline2	○	○	○		
ICRA	○	○		○	○

表3 本稿における有害情報の範囲

Table 3 Range of harmful information in this paper.

カテゴリ	内容
性	ポルノ、ヌード、性行為、異性愛、同性愛、個人広告、デートサービス
暴力	テロ、武器、兵器、軍事、過激活動
犯罪	違法行為、犯罪行為
人権侵害	差別(人種、性)、中傷
薬物	違法薬物、麻薬
その他	不快な描写、個人情報、アルコール、タバコ

2.3.2. 有害な情報の定義

情報が有害であるかどうかという判断は、その情報の受信者の年齢や文化により異なっている。例えば、過激な性表現は子供にとって有害な情報となるが、そのような情報に対して興味を持つ成人にとっては有害な情報とはならない。

どのような情報が有害なものであるのか、という問題については様々な意見があり、一概に何が有害な情報であるかということ定義することは難しいといえる。

既存の有害情報の定義を表2に示す[6-10]。既存の定義の多くが性、暴力、犯罪、薬物、人権侵害が有害な情報であるとしている。

そこで、本稿における有害な情報は子供にとって有害な情報であるとし、既存の定義を参考にして有害情報の範囲を表3のように定義する。

2.3.3. 有害情報の自主規制

情報の発信者側から有害情報を発信しないようにするための取り組みがなされている。

インターネット接続サービスプロバイダの多くは、会員からの違法・有害情報の発信を制限するために利用規約を設けており、このような情報の発信があった場合には情報の発信停止、情報の削除、サービスの利用停止や契約解除を行うことができる。

また、有害な情報が検索結果に表示されず、子供にとって安全と思われるページだけを表示するような子供向けの検索エンジンを提供しているポータルサイトもあり、このようなサイトをブラウザのホームページとして登録しておくことで、子供を有害な情報からある程度遠ざけることができると考えられる。

2.3.4. 有害情報のフィルタリング

法規や情報発信側の自主規制のみでは有害な情報の発信を制限することには限界があり、受信者が望まない有害情報への接触をしてしまう場合がある。

このため、インターネット上での情報発信を制限するのではなく、受信者側で受け取る情報が有害かどうかを判定し取捨選択する手法としてフィルタリングが用いられている。

主に用いられるフィルタリングの情報判定方式としてリスト方式、コンテンツチェック方式、レーティング方式がある。

以下にそれぞれの特徴を述べる。

- リスト方式 (図 1a)

コンテンツが有害であるかどうかをあらかじめチェックしておき、その結果と URI をデータベースにまとめておく。ユーザが Web からコンテンツを取得する際には、コンテンツの URI をデータベースと照会し、有害であった場合には、アクセスを拒否する方法である。

規制対象を登録したブラックリスト方式、登録サイトのみにアクセス可能なホワイトリスト方式がある。

Web にアクセスする際に、コンテンツの URI をデータベースと照会するだけなので、処理速度が速く、CPU 負荷が低いという特長がある。また、Web ページを目視によって確認し、有害であるかどうかの判断を行っている場合、判定の精度も高い。

短所はデータベースに含まれていない新規サイトの判定は行うことができない、判定精度がデータベースの供給に依存すること、データ

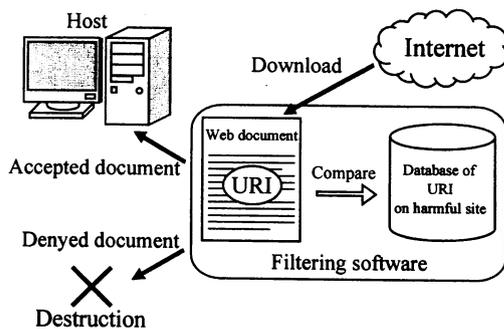


図 1a URI リスト方式

Fig.1a URI list method.

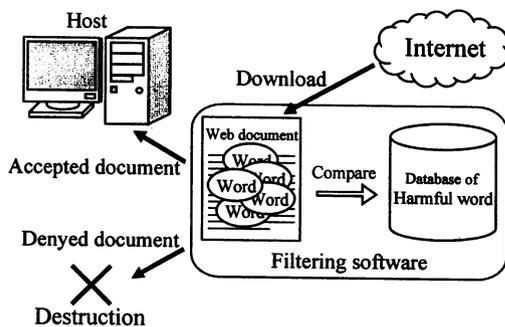


図 1b コンテンツチェック方式

Fig.1b Contents check method.

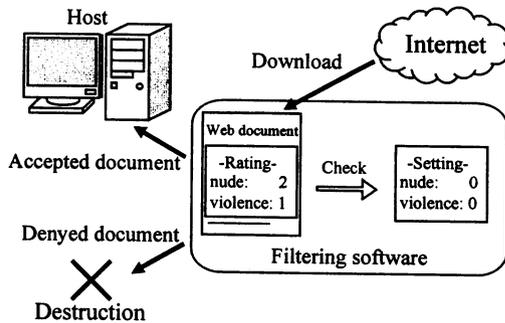


図 1c レーティング方式

Fig.1c Rating method.

ベースの維持管理に多大な労力を要することが挙げられる。

- コンテンツチェック方式 (図 1b)

コンテンツに含まれる単語からコンテンツが有害であるかどうかを判定する方法である。コンテンツを取得する際に、コンテンツの内容

をリアルタイムに判定するため、新規サイトの有害コンテンツに対しても対応可能である。しかし、処理速度が遅く、CPU 負荷が高い。また、日本語環境においてはキーワード解析が難しいため、規制ミス、規制抜けが多いといった問題がある。

- レイティング方式 (図 1c)

コンテンツに対して情報発信者がレイティングした内容を検索し、受信者が望むレベルで規制を行う方法である。主に PICS ルールが採用されているが、第三者機関がレイティングしたデータベースを提供するサードパーティレイティングもある。

Microsoft® の Internet Explorer にも採用されており、無料で利用することができるというメリットはあるが、対応サイトが少なく、セルフレイティングの場合は偽装することが可能であるため事実上はあまり機能していない。

また、設置方式でネットワーク型とホスト型に分けることができる。

- ネットワーク型

ネットワーク型はネットワークのフィルタリングソフトウェアをゲートウェイに設置する方法で、配下のネットワークからのリクエストの処理を行う。配下のネットワークに大量のホストが接続されている場合は、ホスト型よりも導入コスト及び管理コストが低く済む。

- ホスト型

ホスト型はフィルタリングソフトウェアを各ホストに対してインストールする。ホストで稼働中のプロセスを監視することや、ホストの資源に対して直接アクセスすることが可能のため、多くの機能を持たせることが可能である。

フィルタリングソフトウェアは Web 上に存在する有害な情報が子供に伝わらないようにするのに有用であるが、フィルタリングソフトウェアの検出能力は完全ではなく、有害な情報が受信されてしまう場合もあるため、保護者がそばにいて適切な対応をとることが望ましい。

2.4. 本研究の目的

本研究は、子供が有害な情報に接触することを防止し、保護者に対して子供が有害な情報に接触しているかどうかを知らせることを目的としている。また、子供が興味を持っている分野を保護者に伝え、子供に対する理解を深める手助けをすることを目指すものである。

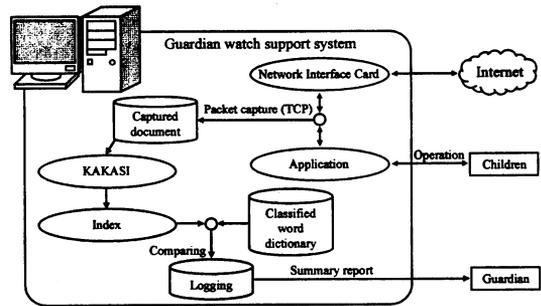


図 2 監視支援システムの概要
Fig.2 Outline of watch support system.

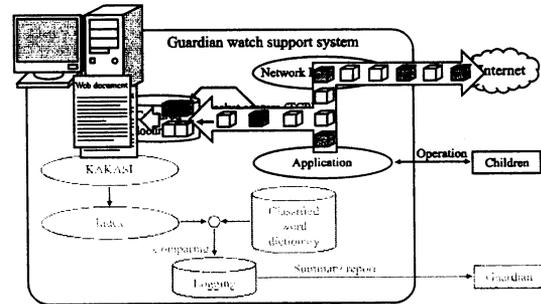


図 2a Ethernet フレームの取得
Fig.2a Acquisition of Ethernet frame.

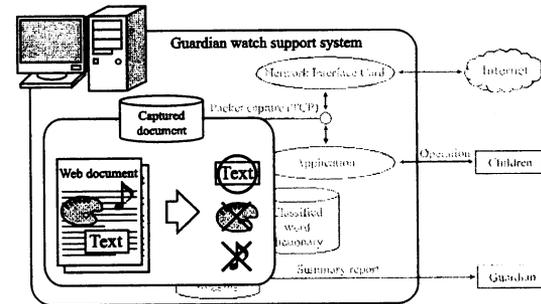


図 2b テキストデータとバイナリデータの分離
Fig.2b Separation of text data and binary data.

3. システムの概要

3.1. 全体構成

本システムの概略図を図 2 に示す。本システムは子供が直接利用する PC にインストールして使用するホスト型の設置方式を取っており、インターネットと P

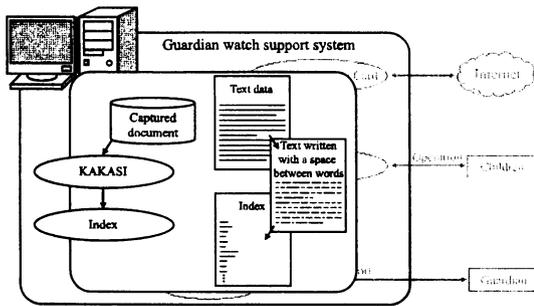


図 2c インデックスの作成
Fig.2c Making of index.

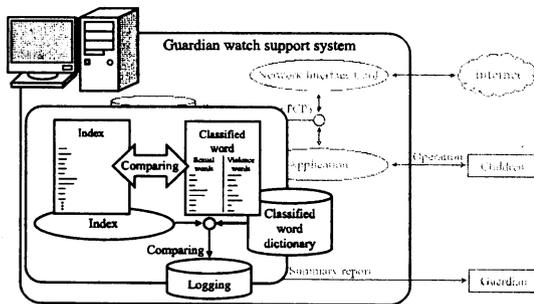


図 2d カテゴリー辞書との比較
Fig.2d Comparison with classified word dictionary.

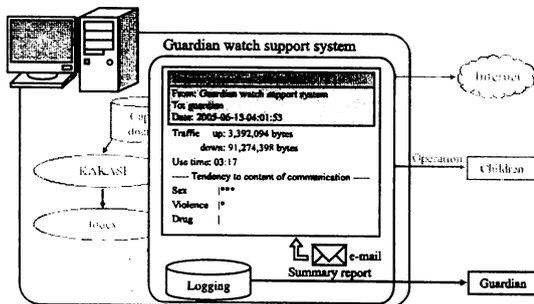


図 2e 保護者への通知
Fig.2e Notification to guardian.

表 4 辞書に含まれる単語の一例

Table 4 Example of word included in dictionary.

カテゴリー	単語例
性	アダルトビデオ、イメクラ、ポルノ、痴漢
ドラッグ	コカイン、シンナー、ヘロイン、大麻
人権侵害	(放送禁止用語などから抜粋)
2ちゃんねる	きぼんぬ、ぬるぽ、もうだめぽ、インスコ、オマエモナー
IT	イーサネット、エンコード、スパイウェア、乗積回路

C間で送受信されるパケットに含まれる Web ドキュメントを監視対象とする。検出はコンテンツチェック方式を採用し、Web ドキュメント内に有害な情報が含まれていないかどうか調べ、電子メールを用いて分析結果の用紙を保護者へと通知する。

なお、システムの運用にあたっては、子供と保護者の間での事前了解を取ることが望ましい。

3.2. 端末内でのパケットの取得

インターネットを介して使用されるアプリケーションの多様性を考慮し、汎用的に通信の監視を行うことのできるパケットキャプチャによって通信内容の取得を行うために、パケットドライバ WinPcap を用いたツールを作成した[11]。

キャプチャされたパケットは、データリンク層のフレームとして得られ、ヘッダを取り除いて IP データグラムを取り出す。同様にしてネットワーク層、トランスポート層のヘッダを参照して、TCP パケットのみを取得して再構築し、各ヘッダを取り除いて TCP データを得る。(図 2a)

また、これらのヘッダを用いることで、TCP/IP およびアプリケーション層のプロトコルの挙動を模擬する。インターネットとアプリケーション間の通信を再現することにより、インターネットとの通信に含まれる画像、音声などのバイナリデータを取り除き、テキストデータを得ることができる。(図 2b)

本稿では、これらのうち暗号化されていない通信において送受信される Web ドキュメントに含まれる日本語のテキストデータを取得し使用する。

3.3. カテゴリー別単語の辞書化

表 3 の定義に基づきカテゴリー辞書の作成を行う。本稿で用いる有害情報検出用のカテゴリー辞書は表 3 に示したカテゴリーの中から性、ドラッグ、人権侵害に関するものを選択し作成した。また、有害情報以外のカテゴリーの検出を行なうために、2ちゃんねると IT に関連する単語からなるカテゴリー辞書も作成した。

表 4 に辞書に含まれる単語の一例を挙げる。それぞれの辞書に含まれる単語がそれぞれのカテゴリーに属していることがわかる。

3.4. 単語の検索

Web ドキュメントに含まれるテキストデータに対し KAKASI を用いて分かち書きを行い単語ごとに区切った後、ドキュメントに含まれる単語のインデックスを作成する[12]。(図 2c) インデックスとカテゴリー辞書と比較し、送受信されたパケットにこれらの単語が含まれているかどうかを調べて記録する。(図 2d)

これにより、子供が有害な情報に接触していないかどうか、どのような情報に興味を持っているかの傾向

を知ることができる。

3.5. 保護者への通知

パケットに含まれる単語を分析して作成した要旨を、あらかじめ登録された保護者の電子メールアドレスに通知する。この要旨には、通信時間、転送量、パケットに含まれる単語の傾向が示される。(図 2e)

また、この要旨にはインターネットと PC 間で送受信されたテキストデータが一切含まれていないため、子供のプライバシーを守ることにもつながる。単語の辞書を有害なものだけでなく、幅広いカテゴリーについて用意することで、保護者が子供の興味の変遷を知ることが可能である。

4. システムの動作実験

Windows PC 上に実装されたシステムの動作を確認するために、Impress Watch, Yahoo! Japan News, 2ちゃんねる News カテゴリー, アダルトサイトをそれぞれ 1 時間閲覧し、検出された単語の分布を比較する。

各サイトから受信したドキュメントに含まれる単語の種類および総数は表 5 のようになった。また、単語をカテゴリー辞書と比較して得られた結果を図 3 に示す。2 ちゃんねるとアダルトサイトの単語種数が多いが、これはこれらのサイトにおいて口語的表現が多用されており、分かち書きを行うことが難しいことによるものだと考えられる。

Impress Watch と Yahoo! Japan News では IT に関連する単語、アダルトサイトでは性に関連する単語が多く検出された。また 2 ちゃんねるでは複数のカテゴリーに渡って有害単語が少量ずつ検出された。

5. まとめ

子供の Web アクセスにおける保護者監視支援システムの Windows プラットフォーム上への実装とカテゴリーごとの単語検出の動作実験について述べた。今後、カテゴリー辞書を充実させ、さらにドキュメントに含まれる単語の傾向を用いることで検出精度の向上を図る。また、PC の操作にあまり詳しくない保護者に配慮し、簡単に操作することが可能な GUI の実装を目指す。

表 5 各サイトを閲覧した結果

Table 5 Result of inspecting each site.

	単語種数	単語総数
Impress Watch	10,259	60,948
Yahoo! Japan News	11,507	124,205
2ちゃんねる	27,900	124,044
アダルトサイト	17,285	120,888

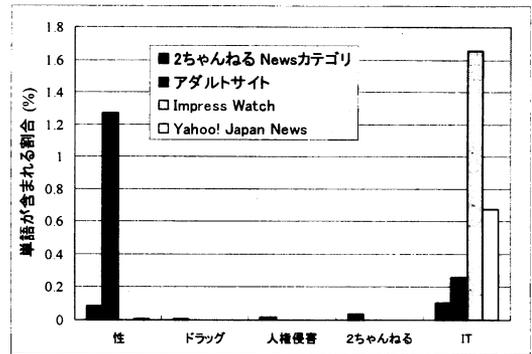


図 3 辞書との比較結果

Fig. 3 Comparison result with dictionary.

文 献

- [1] 平成 16 年版 情報通信白書, 総務省
- [2] “学校における情報教育の実態等に関する調査結果”, 文部科学省, 平成 15 年度
- [3] “有害サイトブロックに関するアンケート”, パナソニックネットワークサービス株式会社, トレンドマイクロ株式会社, <http://www.trendmicro.com/jp/about/news/pr/archive/2004/news040705.htm>
- [4] “家庭教育におけるテレビメディア調査/青少年とインターネット等に関する調査”, (社) 日本 PTA 全国協議会, 平成 16 年度
- [5] “インターネット活用のための情報モラル指導事例集”, 財団法人コンピュータ教育開発センター
- [6] “Kids.us Content Policies”, NeuStar Inc, http://www.kids.us/content_policy/content.html
- [7] “The SafeSurf Internet Rating Standard”, SafeSurf, <http://www.safesurf.com/ssplan.htm>
- [8] “CyberList”, SurfControl plc, <http://www.cyberpatrol.com/default.aspx?id=139&msgid=7>
- [9] “SafetyOnline2”, 財団法人インターネット協会, http://www.nmda.or.jp/enc/rating/rating_standard.html
- [10] “The ICRA vocabulary”, Internet Content Rating Association, <http://www.icra.org/vocabulary/>
- [11] WinPcap, <http://winpcap.org/>
- [12] KAKASI, <http://kakasi.namazu.org/>