Webページ向けフィンガープリント方式の一提案

鈴木貴志[1)]　青木輝勝[1)、2)]　沼澤潤二[1)、2)]

[1)] 東北大学 工学部 情報知能システム総合学科

[2)] 東北大学　電気通信研究所/情報科学研究科

1．　はじめに

近年、電子商取引やBlog、web掲示板等の情報発信など、Webページを介した情報交換や取引が活発となっている。Webページの受信者を操作ミスや不正送信者、第3者から守る技術や法律が用意されているが、Webページの受信者が改ざんやなりすましなどの不正をした場合、正規の送信者を守る技術は不十分である。そのため、本研究では正規の送信者を守るための手法を提案する。

2．課題点と従来研究

Webページは通常容易に書き換えが可能であり、改ざんされる前の原本が手元にあったとしてもその証拠能力は低い。

一般に、不正受信者の検出のためにはコンテンツ配信分野等で広く研究が進められているフィンガープリンティング技術をWebページに適用する方策が考えられるが、同じWebページを複数入手し、差分をとって変化内容を検出する、結託攻撃に対して十分な耐性を持たせることは困難である。上述の通りWebページは書き換え可能であるため、変化箇所を検出し、入れ替えられてしまった場合、受信者の正確な特定ができなくなってしまう。

Webページを保護するための従来研究としてはインターネット・マーク[1]、Webフィンガープリント方式[2]や文書にIPアドレスを挿入するする手法[3]などがある。

しかし、インターネット・マークは、主に受信者の保護を目的とした技術であるため本稿で検討する問題、すなわち、正規送信者の保護（または不正受信者の検出）という問題を解決することには適さない。また、Webフィンガープリント方式は、Webページの不整合問題の解決を目的としているが前述の結託攻撃耐性が必ずしも十分ではない。文献[3]の手法は、ソースコードの"＝"、"＜＝"、"！＝"などの符号の前後にすかしを挿入するという方法でありブラウザには出力されない。つまり、ソースコード保護には使用できるが、アナログドメイン(いったん印刷されたWebページなど)においてその効力はゼロとなってしまう。

本稿では、以下の背景のもと、既存手法と比較して結託攻撃の耐性の向上が期待できるWebページ向けフィンガープリンティング手法として、ブラウザにすかしが挿入された文書を出力する手法、DWPF(Dynamic Web Page Fingerprinting)を提案する。

3．DWPF方式の提案

提案手法では、改ざん元を特定可能にするために、Webページを動的に変化させ、アクセス者を判別できるようにする。IPアドレスはIPv4での2進数表示で32桁なので、Webページに32ビット分の変化をもたせることとする。そして、アクセス者のIPアドレスに応じて変化させることで、Webページにアクセス者のIPアドレスを挿入することにより、必要に応じてこの情報を読み出すことにより不正受信者を特定する。

3．1　CSSの変化

Webページを改ざんし、送信者にクレームをつける場合、改ざん内容に信憑性を持たせるために、改ざん場所以外は可能な限りもとのWebページ

"a proposal on a new fingerprinting method for web pages"

[1)]Takashi Suzuki・Imformation and Intelligent Systems,Engineering,Tohoku Univ

[1),2)]Terumasa Aoki・RIEC/GSIS,Tohoku Univ

[1),2)]Junji Numazawa・RIEC/GSIS,Tohoku Univ

のままにするはずである。よってCSS(Cascading Style Sheet)を変化させ、文章出力の各タグ(<body>,<a>,<h3>など)に対し、次のような箇所を用いて32ビット分の変化をもたせる。

- color
- font-size
- margin もしくは padding
- line-height
- letter-spacing

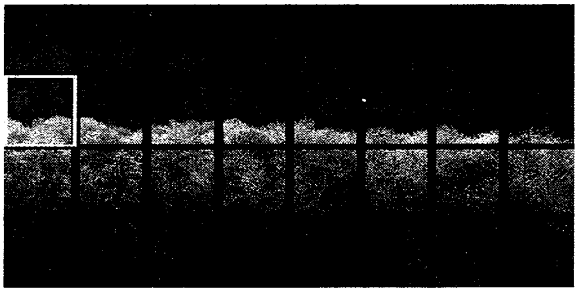
CSSに変化をもたせることで、ブラウザの出力文章のフォントや文字間隔に違いを持たせることができる。

3．2　掲載画像の分割と接合

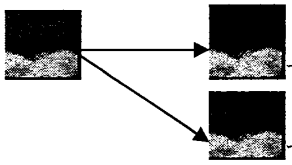
Webページに画像を掲載する場合、画像の1つ1つに電子すかしでIPアドレスを挿入すれば、画像入れ替えという形の結託攻撃にも対応できる。掲載する画像を32個以上に分割し、それぞれに2通りの電子すかしを入れる。すかしを入れた画像をサーバにおき、アクセスごとにアクセス者のIPアドレスに応じて接合する。電子すかし自体はあらかじめ挿入してあり、アクセス毎にその接合を行うことにより、電子すかし挿入に伴う演算時間の大幅な削減が期待できる。これは文献[2]の手法と異なり、掲載する画像が1つの場合でも対応できる利点がある。

また、画像をより細かく分割し、挿入する電子透かしの種類を増やせば、挿入する電子すかしの組み合わせをより複雑なものにでき、差分を検出する方の結託攻撃の耐性も向上できる。

①画像を分割



②電子透かしの挿入



③すかし入りの画像をサーバに保存、受信者のアクセスごとに接合

図1.　画像の分割と接合の手順　例

4．DWPF法の評価

4．1　CSSの変化

データの増減はほとんど皆無であることに加え、すかし挿入もリアルタイム処理で可能であることを実証した。また、主観ではあるが、文章自体に空白スペースの有無やかなと漢字の変換などの変化をもたせる場合と比較しても、CSSに変化をもたせるほうが識別が難しかった。

4．2　掲載画像の分割と接合

IPアドレスのすかし入りコンテンツを用意しようとするとデータ量が膨大になるが、この手法では元の画像のデータ量に対してたかだか2倍で済むためサーバへの負荷は極めて小さい。

5．　まとめ

本研究は、受信者に気づかれにくいようにWebページを動的に変化させることで、Webページを改ざんされた際、変化の内容から改ざん元を特定するとともに、個人などでも使えるような、低コストですむ手法を提案した。

文献

[1]州崎誠一、吉浦裕、永井康彦、豊島久、佐々木良一、手塚悟、"Web サイトの真正性を確認可能とするインターネット・マークの提案"、情報処理学会論文誌、2000.8.

[2]青木輝勝、沼澤潤二、安田浩、"Web ページ真正性保証のための WebCoFIP 方式"、電子情報通信学会 MIH 研究会、2007.10.

[3]Subin Park、Dongsub Cho、"Web Server for Web Page Fingerprinting" Granular Computing, 2008. GrC 2008.IEEE International Conference on、Aug. 2008