建前のあとに：空白を利用したウェブコンテンツのステガノグラフィ

吉 永 珠 里[†]　宮 下 芳 明[†]

本稿では，他人に知られたくない本音を目に見えない空白文字で表現し，記述者本人あるいは許可されたメンバのみが閲覧できるシステムを開発した．本音を記したテキストや画像・動画は，ローカルのファイルに暗号化されて保存され，その相対パスを全角・半角 Space の羅列に変換したものがクリップボードに送られる．ユーザはこれを SNS や Twitter 等の Web サービスで「建前のあとに」貼りつける．ブラウザの拡張機能によってこれを復元し，選ばれたユーザのみが本音を閲覧できる．

Real Intention that Comes After Polite Fiction： Steganography of Web Contents Using Space

Juri Yoshinaga[†] and Homei Miyashita[†]

We developed the system, by which the user can express the real intention not to be known to others by the null characters and enable only permitted members to read it. The texts, images and movies assigned as intention are coded and saved in the local file, and the relative path of them are coded in enumeration of full and half size space, and are sent to the clipboard. The user paste it “ after polite fiction ” in web services such as SNS and Twitter. The extension of the browser decodes this, and only the permitted members can read the real intention.

1. はじめに

情報ハイディングとは，ディジタルコンテンツの中に他の情報を埋め込む技術全般を指す．その中でも，埋め込まれた情報や通信の存在自体を隠す技術，研究分野はステガノグラフィ(steganography) と呼ばれている．これまでに文章中の言葉を同じ意味の別の言葉へ置き換える手法[1] をはじめ，空白文字の色属性や，編集履歴に情報を埋め込む手法[2)3)] が提案されてきた．本稿では，空白文字列を用いたステガノグラフィによって，ウェブコンテンツに「建前」を記述するとともに「本音」を隠す手法を提案する．

例えば，他人から薦められた映画の感想を，不特定多数の人が目にするブログに書く際，建前として「面白い」と書くしかない場合がある．このような時，非公開用に別の独立した記事で「つまらない」と本音だけを書いても，何がつまらなかったかわからない．つまり，本音と建前は一緒に書いてあることで本来の意味を持つと言える．ならば「面白い」と書いておきながら，自分が見た時にだけ本音が閲覧できれば，心置きなく本音を残せるのではないだろうか．また，気心の知れた友人にのみ本音を見せたいというケースもあるだろう．

西村ら[4] は研究室の Web サイトが外部向けに公開する情報と，研究室内部の人が共有したい情報が異なるために管理の手間がかかることを，Web 管理の問題の 1 つとして挙げている．

著者らはこれまで，他人に見せたくない本音を記述したファイルのパスを空白に変換して Web 上に貼り付け，記述者のみが閲覧できるシステムを提案してきた[5]．本稿ではさらに，このシステムに Arcfour(RC4 互換) アルゴリズム[6] による暗号化と，許可されたメンバによる本音共有の機能を追加した．

2. 関連研究

Web ページ中に情報を隠していることから，提案システムは情報ハイディングの領域に位置づけれられる．また，空白を利用してその存在を秘匿しているので，Web ページを対象としたステガノグラフィと言える．

空白を利用したステガノグラフィとして SNOW[7] が挙げられる．これは，テキスト文書の各行末に Space や Tab を挿入することで，メッセージを埋め込むプロ

† 明治大学 理工学部 情報科学科
Department of Computer Science, Meiji University

グラムである．メッセージを埋め込まれた文書は，見た目上の変化はなく，そのままの意味を保つことができる．SNOW ではメッセージを埋め込む際にハフマン符号化を利用して情報量を圧縮している．しかし，埋め込むメッセージが長い場合，変換されたメッセージの一部に長大な空白が付与されてしまう．本稿の提案システムでは，隠したいメッセージを暗号化してローカルに保存し，そのファイルへのパスを短い空白文字列として埋め込んでいる．

空白を利用した研究では，XML 文書のタグ中への空白付加と，無意味な要素の付加によって情報を埋め込む手法を井上ら[8] が提案している．空白を使わない手法としても，違和感のない自然なカバー文書の生成と，自然言語処理により秘密テキストを抽出する手法が提案されている[9]．本稿では，隠す本音の情報量や形式に依らず，カバー文書をそのまま活用している．

限られた人にだけ伝わる書き方として，伏字が挙げられる．「インタラクシ◯ン」ならば，「インタラクション」という言葉を知っている人は伏字部分を補間して読める．しかし，文章中の名詞を伏字に変換するサービス[10] や，伏字で隠された文字を解読するサービス[11] の存在から，伏せる箇所や言葉によっては，記述者の意図とは違った解釈をされることがわかる．

読者の知識に依存しない伏字では，増井が伏字を登録するサービス fuseji.com[12] を構築している．登録された伏字は，特定の順番にクリックすることで，中身が見える仕組みになっている．ブログにこのサービスで登録した伏字を導入すると，クリックの順番を知っている人だけが，中身を見ることが可能になる．

また増井は「界面潮流」[13] で，秘密の情報をパソコンで扱う際の問題として，ファイルの置き場所を忘れてしまうことを挙げている．提案システムでは，ファイルの作成から利用までを自動で行うため，本音ファイルをユーザが意識する必要はなくなっている．

mixi に代表される SNS サイトには，人と人とのつながりを支援するために日記やブログ機能を持つものが多い．ブログはコミュニケーション的役割を持ち，既存の人間関係を意識して構築されることが多いと山本ら[14] は述べている．したがって，人間関係を意識した結果，建前のために本音が排除されてしまうことはブログを書く際に直面しやすい問題と言える．

永田らは手軽に情報の公開範囲を操作できるツール Enzin[15] を開発した．公開範囲の操作により，個人的なメモ，メール，メーリングリストなど様々な形態としての機能を実現している．しかし，メッセージの一部だけに公開制限をかけることを想定していないので，

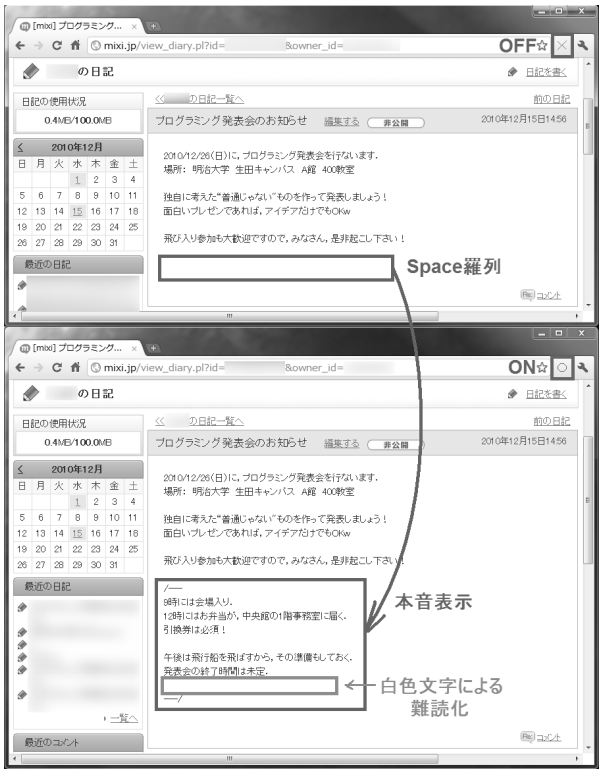

図 1 mixi の閲覧：記述者以外 (上)，記述者と共有者 (下)
Fig. 1 Viewing mixi : Public(Upper), the describer and the part owners(Lower)

その場合は変更を加えたメッセージを新規に作成する必要がある．また操作対象はメッセージのみで，画像や動画を扱うことができない．

中村のネタバレ防止ブラウザ[16] は他者が発信した情報のうち，ユーザが受信したくない情報をフィルタリングする．これに対し，提案システムは自らが発信した情報の公開範囲を管理し，他者に知らせないツールであると言える．

3. システム

本稿では他人に見せたくない記述を「本音」，誰にでも見える記述を「建前」と呼ぶ．記述者は本音と建前を一緒に書いて発信できる．本音は実質的に Space のみの空白として表示され，ブラウザの拡張機能により記述者には本音が見える．我々は本音を保存して暗号化するシステムと，復号化システムを構築した．図 1 に mixi の日記における動作例を示す．

3.1 本音の保存と暗号化

本音を保存・暗号化するシステムでは，Arcfour(RC4 互換) アルゴリズム[6] により本音を暗号化する．その後，ローカルにテキストファイルとして保存し，ファイルの相対パスを全角・半角 Space の羅列に変換したものをクリップボードへ送る． Space で構成され

図 2 Twitter の閲覧：記述者以外 (左)，記述者 (右)
Fig. 2 Viewing Twitter : Public(Upper), the describer and the part owners(Lower)

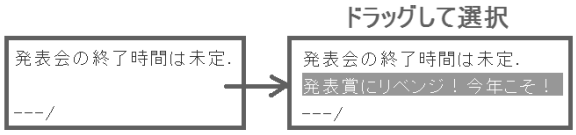

図 3 白色文字による難読化
Fig. 3 Obfuscation by white character

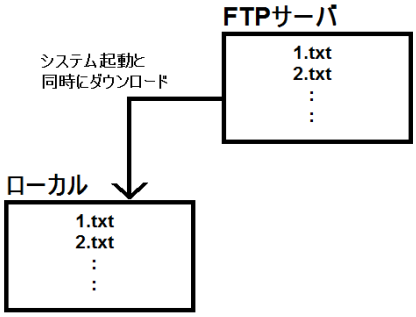

図 4 共有のために FTP サーバ上に置かれた本音
Fig. 4 Real intention on FTP server for sharing

たファイル名はそのまま Web 上に公開される．通常では，記述者を含めたすべての閲覧者にファイル名が空白で表示される．第三者がそれを解読したとしてもローカルファイルの相対パスしかわからないため，外部から本音を表示することは原理的に不可能である．

3.2 復号化とブラウザ表示

本音の復号化システムは，Web ページ中にある空白の羅列をファイル名に復元する．さらに同名ファイル中の本音を復号化し，空白の羅列と置き換える．これはブラウザ Google Chrome の拡張機能として実装した．復号されたテキストは，HTML ファイルの変換部分を書き換えることで記述者へ表示される．

復号化システムでは，HTML ファイルの書き換えを行っている．そのため，本音に HTML ソースを記述し，リンクや画像・動画を貼り付けることが可能である．これにより，テキスト情報としての本音にとどまらず，他人に見せたくない画像や動画をも埋め込んで表示できる．図 2 は Twitter において，画像と動画を表示した例である．このように，本システムを用いれば，HTML タグの使用が禁止されている Web サイトでも自由な表現が可能である．

また，図 3 は「自分の PC を他人が使う場合でも本音が見えないようにする」使い方である．本音を記す時に，タグを使用して白色の文字として表現する．Chrome 拡張を ON にした上で，さらにその領域を選択し，文字色を反転させないと読むことはできない．

4. 指定メンバとの共有

これまでに述べてきた本音情報は，個人の中に留めておきたいものを対象としていた．よって中身を誰かに見せる際は，本音の書かれたテキストファイルを渡し，相手もブラウザの拡張機能を導入する必要があった．しかし，実生活では研究室やサークルといった集団や限られた人と常に情報を共有したいこともある．著者らが WISS2010 にて「建前のあとに」[5] のデモ発表を行なった際，限られた人と本音を共有したいとの要望を多数得ている．そこで，集団で本音を共有可能な集団用システムを試作した．

試作した集団用システムでは，本音の書かれたテキストファイルを FTP サーバ上に置いて共有する．基本的な操作は，個人用システムと同様であり，3 つの追加点がある．1 つ目は，本音をローカルに保存した後，FTP サーバに転送する点．2 つ目は，本音の保存・暗号化するシステムの起動時に FTP サーバへアクセスし，これまでに蓄積された集団の本音をローカルにダウンロードする点（図 4）．3 つ目は，集団としての本音は秘密度が高いと予想されるため，パスワード認証を設けた点である（図 5）．これは拡張機能を有効に切り替えようとする度に表示され，正しいパスワードを入力して初めて有効となる．

本音のダウンロードに必要なパスワードは，暗号化システム中に記述してある．そのため，ユーザは拡張機能を有効にするパスワードのみを知っていれば良い．集団用システムで空白部分を復元されたとしても，本音ファイルはローカルと FTP サーバのみに存在している．空白文字列には，FTP サーバの URL と ID，パスワードといった情報がないため，外部から本音を表示することは原理的に不可能である．

図 5 パスワード認証
Fig. 5 Input password

5. まとめと課題

本稿では，ローカルに保存した他人に知られたくない本音を記述者本人あるいは許可されたメンバのみが閲覧できるシステムを提案した．これにより本音の存在を他人に知られず，建前と一緒に記述できるようになる．また，本音を FTP サーバ上に置き，限られたメンバで共有・閲覧するシステムも構築した．

現段階では，空白を埋め込んだ Web ページと本音の関連付けがされていない．そこで，空白文字列が書かれた Web ページを定期的に巡回してその存在を確認し，本音と同期させることを検討している．

今後，本システムの評価実験を行う予定である．評価対象は個人用システムと集団用システムの 2 つとする．一定期間本システムを使って情報発信を行ってもらう．個人用システムの評価実験では，ブログの記事作成や Twitter でのツイートを行ってもらう．集団用システムの評価実験では，著者らの所属する研究室内で本音を共有する．研究室が行う外部向けの情報発信として Twitter や研究室の HP があり，ここでのシステム使用を予定している．

評価アンケートでは SNS や Twitter 等の Web サイト上での情報発信において，自分の置かれている環境に影響されずに本音を記述できるか，空白を利用した手法の妥当性，本音と建前を併せた記述・表示の有用性を検証する．本システムを通じて，今までユーザの置かれた環境により躊躇われた，Web サイト上での率直な本音の記述が可能になることが期待される．

また，システム使用者が発信した情報に対する調査を行う予定である．提案システムの機能を知っている閲覧者が，システム使用者が発信した情報に対して，印象が変化するかを調査項目に含める．変化がなければ，記述者の意図通りに記述できたと言えるであろう．

参 考 文 献

1) 中川 裕志，三瓶 光司，松本 勉，柏木 健志，川口 修司，牧野 京子，村瀬 一郎：意味保存型の情報ハイディング—日本語文書への適用，情報処理学会論文誌, Vol.42, No.9, pp.2339–2350 (2001).
2) 北野 宗之, 増田 英孝, 中川 裕志:Word2003XML 文書への情報ハイディングシステム，電子情報通信学会研究報告，Vol.105，No.193，pp. 205–212 (2005).
3) 井上 照久，遠山 毅，松本 勉：編集履歴付きディジタル文書を媒体とするステガノグラフィの一方式，情報処理学会, Vol.2009, No.20, pp.313–318 (2009).
4) 西村 美咲，横井 茂樹，安田 孝美：研究室の情報共有・公開を支援する CMS を基盤とした Web システムの構築，情報処理学会, Vol.2008, No.16, pp.23–28 (2008).
5) 吉永 珠里，宮下 芳明：建前のあとに，第 18 回インタラクティブシステムとソフトウェアに関するワークショップ (WISS2010)，日本ソフトウェア科学会, pp.165–167 (2010).
6) Kalle Kaukonen，Rodney Thayer：A Stream Cipher Encryption Algorithm ”Arcfour”，Internet draft (draft-kaukonen-cipher-arcfour-03.txt)，Network Working Group (1999).
7) SNOW Home Page. http://www.darkside.com.au/snow/
8) 井上 信吾，村瀬 一郎，滝澤 修，松本 勉，中川 裕志：XML におけるステガノグラフィ手法の提案，暗号と情報セキュリティシンポジウム (SCIS2002)，pp.301–306 (2002).
9) 滝澤 修，山村 明弘：自然言語テキストを用いた秘密分散法，情報処理学会論文誌, Vol.45, No.1, pp.320–323 (2004).
10) 伏字変換サービス. http://fuseji.info/
11) 伏字検索○○○. http://fuseji.net/
12) fuseji.com - なぞなぞ伏字サービス. http://fuseji.com/
13) 増井俊之の「界面潮流」. http://wiredvision.jp/blog/masui/
14) 山本 仁志，諏訪 博彦，岡田 勇，山本 浩一：ブログ空間上のコミュニケーションスタイル，経営情報学会 2007 年春季全国研究発表大会予稿集, pp. 112–115 (2007).
15) 永田 周一，安村 通晃：Enzin:情報の公開範囲を手軽に変更できるコミュニケーションツール，情報処理学会論文誌, Vol.48, No.3, pp.1134–1143 (2007).
16) 中村 聡史：ネタバレ防止ブラウザの実現，第 18 回インタラクティブシステムとソフトウェアに関するワークショップ (WISS2010)，日本ソフトウェア科学会, pp.41–46 (2010).