ユーザフレンドリかつ不正攻撃に強い、
コンピュータと人間の識別能力の違いを活用した判別基盤の開発
— 新しい CAPTCHA の開発 —

1. 背景

現在、インターネット上には、不正なボット（スパムなどを自動送信するためのプログラム）により、スパムが氾濫している。スパムの種類と形態は多種多様であり、このスパムを如何に防ぐかは、現在の IT 社会において非常に厄介な問題である。
これらの問題点を解決するための一つの方法として、インターネット上の簡易なコミュニケーションやインタラクションにおけるスパム防止手段として、CAPTCHA（キャプチャ、"Completely Automated Public Turing test to tell Computers and Humans Apart"; コンピュータと人間を区別する完全に自動化された公開チューリングテスト）が生まれたが、近年、この手法も破られ、危険な状況となりつつある。

2. 目的

本プロジェクトの目的は、利用者が使いやすく不正な攻撃に強い人と機械の判別基盤の構築を通して、世界中のスパムを削減し、IT 資源と人資源の効率的な利用をはかり、持続可能な安全・安心な IT 社会の実現するために、新しい発想や斬新な技術を駆使し、システムの正当な利用者である人間と、スパムを送信する不正なプログラムである機械を、短時間で判別する新しい「CAPTCHA」を開発することとした。

3. 開発の内容

新しい CAPTCHA 技術を実現するために、今回のプロジェクトでは下記４点を開発した。

- CAPTCHA アルゴリズム

  Oblivious CAPTCHA と Reading CAPTCHA に適した CAPTCHA アルゴリズム

- Oblivious CAPTCHA

  CAPTCHA の新しい耐性強化軸として、真の文字列に“ダミー文字”を挿入し、真の文字列やその個数を答えさせる CAPTCHA。 利用者が使えば使うほどダミーの精度が高まり、ダミーの基礎個数が増える仕組みを備える。

- Reading CAPTCHA

  機械が読み取れなかったスキャンされた紙媒体の文字を人間が読む CAPTCHA。利用者が使えば使うほど紙媒体の文字の電子化が進み、また電子化の精度が高まる仕組みを備える。

- CAPTCHA サーバと Web・携帯プラグイン

  CAPTCHA を Web サービスとして提供するための、CAPTCHA サーバと Web プラグイン（サーバが提供する Web サービスを、Web サイト・携帯向け Web サイトが容易に利用・組み込み可能な形にするための Web サイト用の拡張機能）。

3．1 CAPTCHA アルゴリズム

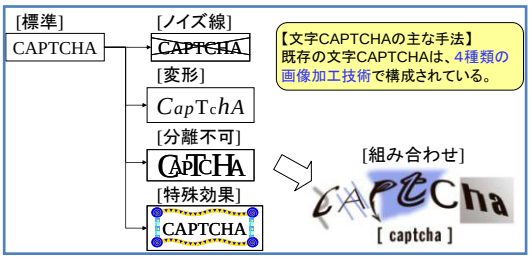

図 1：文字 CAPTCHA に使われている既存手法

図 1 は、文字 CAPTCHA に使われている既存の手法を示している。
文字 CAPTCHA に使われている既存の主な手法は、ノイズ線（arc）、変形（deform）、分離不可（undividable）、特殊効果（special effects）の、4種類の画像加工技術である。（図 1）
Oblivious CAPTCHA と Reading CAPTCHA のそれぞれに適したアルゴリズムを、上記の4種類の画像加工技術から選定または組み合わせて、開発した。

3．2 Oblivious CAPTCHA

Oblivious CAPTCHA は、CAPTCHA の新しい耐性強化軸として、真の文字列に“ダミー文字”を挿入し、真の文字列やその個数を答えさせる CAPTCHA である。
Oblivious CAPTCHA は、利用者が使えば使うほどダミーの精度が高まり、ダミーの基礎個数が増える仕組みを備える。

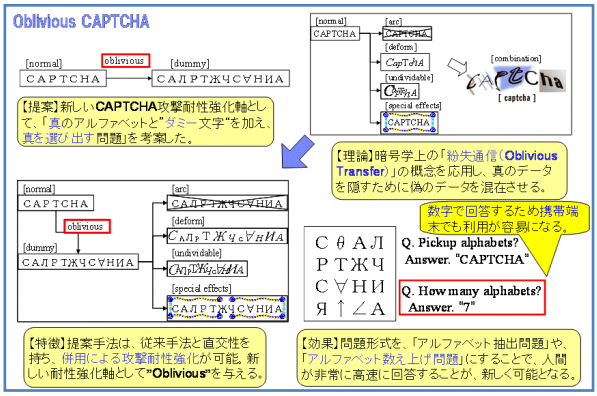

図 2：Oblivious CAPTCHA の概要

図 2 は、Oblivious CAPTCHA の概要を示している。
Oblivious CAPTCHA は、利用者が使えば使うほどダミーの精度が高まり、ダミーの基礎個数が増える仕組みを備える。利用者へダミーとして提示した記号の中で、多くの利用者がダミーと回答した記号は「良いダミー」として採用され、利用者により判断が分かれた記号はダミーから除外することでダミーの精度と品質を高める。また、ダミーか否か未知の記号を利用者に提示し、多くの利用者がその記号をダミーと回答すればダミーに採用され、利用者により判断が分かれた記号は却下することで、ダミーの基礎個数を増やすことができる。

3．3　Reading CAPTCHA

Reading CAPTCHA は、「人間は読めるが機械は読めない」という CAPTCHA の仕組みを利用して、機械が読み取れなかったスキャンされた紙媒体の文字を人間が代わりに読む CAPTCHA である。Reading CAPTCHA は、利用者が使えば使うほど紙媒体の文字の電子化が進み、また電子化の精度が高まる仕組みを備える。

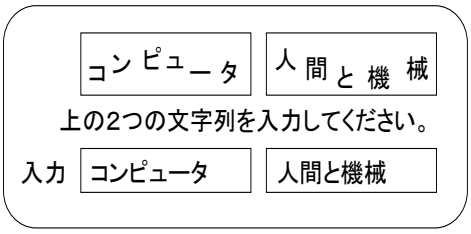

図 3：Reading CAPTCHA

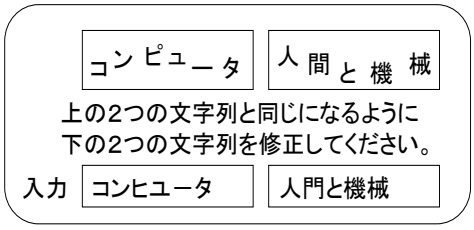

図 4：誤り文字を修正する Reading CAPTCHA

図 3 は、Reading CAPTCHA の概要を示している。
Reading CAPTCHA は、2つの文字列を表示する。両方とも紙媒体からスキャンされた文字であるが、正しい文字列が分かっている文字列は片方だけで、もう片方は OCR プログラムにより正しく読み取れなかったと警告された文字列である。利用者は両方の文字列を入力する必要があるが、スパム対策の確認に用いられるのは片方だけである。もう片方は OCR の結果の修正情報として利用する。正確さを保つため、正しい文字列が分かっていない文字列は複数の利用者に提示し、それらの入力を総合して正しい文字列を判断する。
既存の手法が利用者の労働力を人間確認だけに利用するのに対し、Reading CAPTCHA は人間確認と同時に OCR の修正という有用な目的に活用しており、既存手法と本質的に異なるアプローチである。
図 4 は、誤り文字を修正する形態の Reading CAPTCHA である。誤りの部分だけを入力することで、利用者の作業負担を軽減する。複数の OCR で事前実験した結果、特に日本語文は漢字など1文字単位の認識誤りの割合が高く、修正を入力する形態で効率的な Reading CAPTCHA が構成できる。
文字列は、適切な単位で区切り利用者に提示する。文字列が欧文など明確な区切りのある言語の場合、単語の区切りの空白を利用して単語単位で区切る。文字列が日本語など明確な区切りが無い言語の場合、適切な長さの文字数や句読点を単位として区切る。
利用者の負担を軽減するために、解く問題の選択権を利用者に与える。利用者は、提示された Reading CAPTCHA が解けない場合、別の Reading CAPTCHA の問題や、通常の文字方式などの CAPTCHA の中から自分が解きたい（解くことが可能な）CAPTCHA を自由に選択することができる。

3．4　CAPTCHA サーバと Web・携帯プラグイン

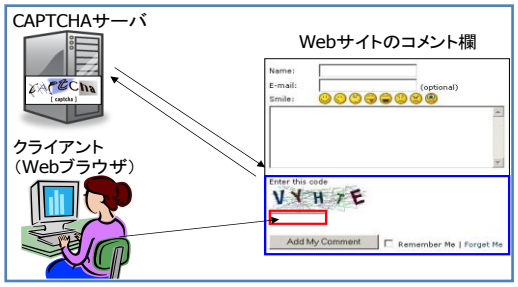

図 5：利用シーン

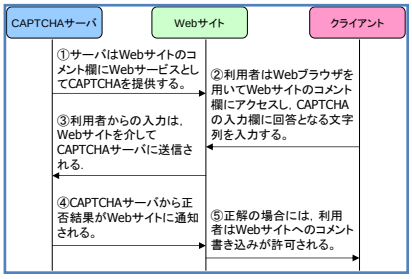

図 6：利用シーン（シーケンス図）

新たに開発した CAPTCHA は Web サービスとして提供される。また、サーバが提供する Web サービスを、Web サイト・携帯電話向け Web サイトが容易に利用・組み込み可能な形にするための Web サイト用の拡張機能として、Web プラグイン・携帯プラグインを開発した。日本で広く普及している携帯電話向け Web サイトにも対応することで、CAPTCHA サービスの利用者数を増やし、利用者層を広げる。
図 5 は利用シーンのイメージ図である。また、図 6 はそのシーケンス図である。

4. 従来の技術（または機能）との相違

開発した CAPTCHA は以下の特徴を持つ。

1) CAPTCHA の利用が CAPTCHA の強化につながる、持続可能なスパム防止基盤を構築する。
2) CAPTCHA を解く人間の労働力を有用な目的に役立てることで、人資源の有効活用をはかる。

5. 期待される効果

従来技術との比較において、当成果が解決する課題と優位性は以下の通りである:

課題1:従来 CAPTCHA は、アルゴリズムが固定のため、時間の経過により、強度が劣化してくる

当成果:使用することでより強固に成長し、セキュリティ強度が向上する「Oblivious CAPTCHA」手法を開発

課題2:従来 CAPTCHA は、設置に人件費やシステムコストの負担が大きい

当成果:CAPTCHA を解くことを出版物の電子化に役立てる「Reading CAPTCHA」手法を開発

- インターネット利用者は1日2億個の CAPTCHA を解いている
- 1個10秒とすると1日15万時間以上の労働力
- 時給 1000 円とすると1．5億円分の労働力

6. 普及（または活用）の見通し

想定するビジネス（収益）モデルとして、以下を考えている。

- Web プラグイン課金モデル（Web サイト提供事業者）
- 紙媒体の原稿の電子化手数料モデル（原稿提供者）
- 広告モデル（広告提供者）
- 電子メールアドレスの非クロール化サービス

7. クリエータ名（所属）

小瀬木　浩昭　（内閣府認証特定非営利法人ITプロ技術者機構）