1J3-1

Arnold's CATMAP のダイナミクスを用いた CAPTCHA システムの開発

Proposal of CAPTCHA system using the complex dynamics of Arnold's CATMAP

井上 聡[*1*2]
Satoru Inoue

[*1] 埼玉工業大学大学院 工学研究科
Graduate school of Engineering, Saitama institute of Technology

[*2] 埼玉工業大学 工学部
Faculty of Engineering, Saitama institute of Technology

In this paper, we propose and investigate the new CAPTCHA system based on the dynamics of Arnold's CATMAP. Our system is aimed to improved its performance through the tuning test which searches the readable words from the images changing periodically.

1. はじめに

インターネット上の WEB サービス利用のために新規にアカウントを作成する際やオンライン投票を行う際など，その応答者がコンピュータや人工知能技術を用いたエージェントではなく，実際に人間が応答していることを確認するために，CAPTCHA と呼ばれるチャレンジ/レスポンス型の認証が行われるサイトが近年増加している．基本的な CAPTCHA システムは，文字列を読み取る能力を問うことで人間と機械の判別を行うものである．文字列に歪みやノイズを付加した画像を Web ページ上に提示し，ユーザが判読できるか否かを試す．しかしながら文字認識技術やパターン分類技術の高度化により，この認証テストが OCR を代表とする文字読み取りエージェントにより突破されてしまうこともあり，それに対抗するように CAPTCHA の課題が過剰に難化し，人間自体もその課題を解くことができず，そのような認証がネットワーク利用時の障害となってしまい，本来の目的が果たせてない本末転倒な事例が散見される。また認証に利用する課題もセキュリティ確保の観点から大量に用意する必要があり，この技術を用いたシステム運用を困難にさせている要因の 1 つとなっている。そこで近年では文字列を用いた CAPTCHA とは異なる新しい方式の CAPTCHA が期待されている．本研究では画像を利用した CAPTCHA システムにカオス的ダイナミクスをもつ Arnold's CATMAP の画像変換を用いたシステムを提案し，CAPTCHA 課題の過剰な難化を避けつつも，認証システムの堅牢性を維持し，また運用コストを軽減するシステムの構築を目指す。

2. 関連研究

現在の CAPTCHA は機械に判読しにくい文字列の判読能力を試すことで，人間と機械を区別する方式が普及している．しかし，OCR をはじめとする文字読み取りアルゴリズムの向上により人間ではないエージェントが CAPTCHA の解答をすることが可能となった．そのため，マルチウェアによる不正利用を防ぐために文字列画像の難読化が繰り返し行われたが，過剰な難化により機械だけでなく人間にも判読が困難な文字列になってしまった．近年こういった問題を解消するために文字列を読み取るタイプの方式とは異なる CAPTCHA が提案された．以下の 2 つの CAPTCHA を示す．

連絡先：井上 聡，埼玉工業大学工学部，埼玉県深谷市普済寺 1690，inoue@sit.ac.jp

2.1 Assira

Asirra は画像認識能力を試す CAPTCHA の一つで，複数の画像を 2 つのカテゴリに分類するプロセスを介して，応答者の属性を判定する方式である[Elson 2007]．解答者は提示された 12 枚の猫と犬の画像から猫の画像を選び出す．提示される猫と犬の画像はラベル付けされており，解答者の答えがラベルと一致している場合に応答者が人間である判定する．猫と犬の分類は人間にとって直観的に行うことが可能であるため負荷が少なく利便性の高い CAPTCHA である．しかし Asirra に対する回避策として，機械学習を用いた画像認識攻撃とデータベース攻撃が挙げられている．2 クラス分類問題を得意とするサポートベクターマシン（SVM）などの機械学習判別器は Asirra にとって有効であり，また使用する画像データベースを短期間で大きく変化させることが出来ないためデータベース攻撃に対しては脆弱であると考えられる。

図 2-1. Assira の実行例

2.2 SS-CAPTCHA

画像認識能力試す CAPTCHA とは異なる方式として SS-CAPTCHA という「違和感を判別する能力」を試す CAPTCHA も提案されている．SS-CAPTCHA[山本 2009]は解答者に人間が作成した自然な文章と機械的に出力した不自然な文章を複数提示し，その中から自然な文章を選び出すことが出来るか否かで人間か機械かを判定する．文章に対する自然か不自然かの判別は人間にとって容易だが，機械にとって自然言語を完全に解釈することは困難である．しかし人間にとって自然な文章と不自然な文章を生成するためには人手が必要となるため，膨大なパターンを用意することが難しいと考えられる．

3. 提案手法

一般的な文字列 CAPTCHA は文字読み取りアルゴリズムなどにより解読されてしまう可能性はあるが，提示する文字列画像は機械的に生成可能なためデータベース攻撃に対しては耐性が高い．また提示された文字列を入力する作業は人間にとって

はきわめて単純であり，判読が可能であれば誤解答や誤操作は起こしにくい．一方，画像 CAPTCHA はデータベース攻撃や画像認識攻撃に対して脆弱な傾向はあるが， 人間と機械の画像認識の能力差は依然として大きく様々な方式の CAPTCHA が提案されており今後の発展が期待されている．以上を勘案したうえで本稿では Arnold's CATMAP のダイナミクスを用いた画像認識能力と文字列の読み取り能力を要求する CAPTCHA システムを提案する．

3.1 Arnold's CATMAP

Arnold's CATMAP とは，パイこね変換と呼ばれるカオス現象の一つで，ロシアの数学者 Vladimir I. Arnold が猫の画像を用いたことからそう呼ばれるようになった[Peterson 1997]．パイこね変換とは，パイ生地をこねるときの「伸ばす」，「折りたた む」という工程を繰り返し行うことで生地内の含有物が最も効率 よく混ざるということを数学的にモデル化したものである． CATMAP における線形写像は次式の通りである．

$$(x_{n+1}, y_{n+1}) = (x_n + y_n, x_n + 2y_n) \bmod N \qquad (1)$$

一辺が N ピクセルからなる正方形の画像を横方向に 2 倍，縦方向に 3 倍に引き伸ばし，元となる画像サイズに切り取り，折りたたむことで画像の変換が行われる．

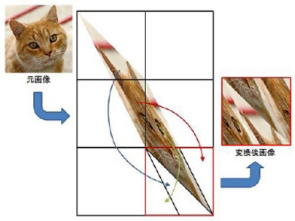

図 3-1. Arnold's CATMAP による画像変換

3.2 システムの概要

提案手法は CATMAP により変換された画像を解答者が簡易な操作で復元し，画像中の文字列を解答することで画像の判別能力と文字の読み取り能力の双方を問うことができる CAPTCHA システムである．本稿では，CATMAP により任意のステップ変換した画像を変換画像．変換前の画像を元画像と呼称する．提案手法の出題から解答までの手順は次の通りである．

出題者は出題する単語をランダムに複数選択し，正方形の画像に配置して元画像を作成する(図 3-2 左)．元画像を CATMAP の変換にかけ，変換画像を作成する．この時の変換回数は画像の判別が困難になるまで十分かつ，秘匿性を高めるためにランダムな回数行う． 作成した変換画像を質問画像として解答者に提示する(図 3-2 右)．次に解答者は提示された質問画像を CATMAP の変換を行うシステムで 1 ステップずつ変換し，元画像に書かれた文字列を読み取り解答する．最後に出題者は解答者の答えと選択された単語を確認し，一致している場合解答者を人間と判断する．

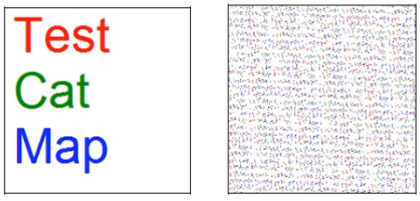

図 3-2 変換前画像(左)と初期提示画像(右)

3.3 CATMAP による変換画像作成

変換画像の作成は文字列を記述した元画像を CATMAP の変換則に従って作成する．元画像を CATMAP の変換に則り 1 ステップだけ変換した変換画像では，ある程度文字列を読み取ることが可能である(図 3-3)．

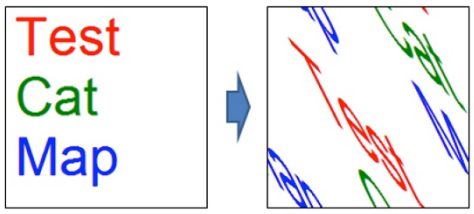

図 3-3 元画像と 1 ステップ変換後の画像

今回作成した CAPTCHA では一辺が 200 ピクセルの元画像を用いて変換画像を生成している．一辺が 200 ピクセルの正方画像での CATMAP 変換の場合，元画像に復元するまでに 150 ステップを要する．また，30 ステップ周期に図 3-6 に示すように 各単語が重なり合ったような画像となる．このことから出題に用いる変換画像の初期状態は元画像から 30〜120 ステップの範囲でランダムに変換を行ったものを使用した．

3.4 解答者への質問

提案手法では解答者が提示された変換画像に CATMAP の変換をかけ元画像への復元を行い，画像中の文字列の読み取り を行う．出題者は解答者の答えが正しい場合に解答者を人間と判別する．質問と解答までのプロセスをまとめると以下のとおりとなる。

【出題プロセス】

1. 一辺が 200 ピクセルの正方形画像に英単語を 3 つ選択し配置を行い，元画像を生成する．
2. 生成した元画像を CATMAP 変換にかける．変換回数は 30 から 120 ステップの範囲でランダムに行う．
3. 変換を行った画像を解答者に提示する(図 3-2 右)．

【解答プロセス】

I. 解答者は提示された画像の文字列の有無を確認する．

II. 判読可能な文字列が画像中に見つからない場合，解答者は画像を CATMAP 変換により次のステップへ変換する．

III. 変換を行った画像の文字列の有無を確認する．

IV. 以降，II, III を繰り返し，文字列を含む画像を探す．

V. 判読可能な画像に変換したら，画像から文字列を読み取り 3 つの英単語を答える． この時，解答者は必ずしも変換画像を元画像へ復元する必要はなく，ある頻度で出現する単語が重なり合った状態の画像から解答を行うことも可能である(図 3-4)．

図 3-4 低頻度で出現する文字読み取り可能な画像例

4. 検討

4.1 利便性についての検討

提案手法による利便性について検討を行う. 既存の文字列 CAPTCHA と提案手法を比較した場合, 解答者に提示する画像を生成するために若干の処理を必要とするが運用性に関しては大きな差はないと考えられる. CAPTCH は人間の解答者に大きな負担を与えないことが望まれる. 提案手法における解答者に対する負担を評価するため実験を行った. 実験では被験者に提案手法の CAPTCHA を解答してもらい, 解答時間と誤解答数を計測する. 実験には一辺が 200 ピクセルの画像中に 3 つの単語を配置した元画像を用いる. 元画像の初期変換回数は 30～120 ステップの間でランダムに行う. この評価実験では 10 人の被験者にそれぞれ 5 問ずつ解答してもらった.

図 4-1 実験に用いた解答インタフェース

実験を行った結果, 提案手法による CAPTCHA 課題への解答に要した時間は平均 34 秒となった. 既存の文字列 CAPTCHA の所要時間は約 10 秒から 18 秒程度である. 既存の文字列 CAPTCHA と比較して提案手法は解答時間が長くなっている. 解答までの所要時間は, 提示画像の初期変換ステップ数に依存するが, それを勘案してもこの結果は実用的な所要時間から大きく逸脱するものではないと考えられる. また誤解答数は 50 問中 4 回という結果が得られた. 誤解答となった解答のほとんどは解答者が復元途中の画像から単語を読み取ろうとした場合に発生している. また連続した誤解答は一度のみだった. この結果から, 提案手法は正答率 92%と既存 CAPTCHA システムと比較すると高い正答率を得られた. よって提案手法は誤解答による解答者への負担は少ないと考えられる.

4.2 安全性についての検討

提案手法は画像認識能力と文字読み取り能力の双方を試す CAPTCHA である. 3. 4 で述べた通り提案手法では解答者が提示されたノイズのような画像を CATMAP 変換にかけ, 文字情報を含んだ画像か否かを判断する. 文字情報を持たない画像と判断した場合, 再度 CATMAP 変換を行う. 変換は解答者が文字情報を含む画像を発見するまで繰り返し探索する. 解答者は発見した文字情報画像の文字列の読み取りを行うが, 画像中の文字列が重なり合って読み取りが困難である場合, 再度探索 を行い読み取り可能な画像を探す. 読み取りが可能な文字列を含んだ画像から文字列を読み取り解答を行う. そのため安全性において画像認識と文字読み取りの基本的な耐性を備えていると言える. まず多くの画像認識を用いた CAPTCHA に脅威とされるデータベース攻撃に対する提案手法での攻撃耐性を示す. デー タベース攻撃は CAPTCHA が出題する問題と答えを記録したデータベースを構築し, 記録された問題が出題されたときにそのデータベースを利用して機械的に解答する攻撃である. 画像認識能力を問う CAPTCHA に対しては, 出題された問題を取り込み, 画像の分類を行い解答用のデータベースを構築する. 画像 CAPTCHA において問題に使用される画像は解答者に 提示しなければならないため, 攻撃者のデータベース構築を防ぐことが困難である. 攻撃に対する対抗策としては, 使用する画像の枚数を多くすることでデータベース構築の労力を増大させることが効果的だが, そのために必要な画像収集などサービス の提供者側の労力も増大してしまう. 提案手法において提示画像の生成には文字列を配置した元画像を用いている. そのため写真などを用いる画像 CAPTCHA とは異なり, 出題ごとに異なる画像を生成できるためデータベース攻撃に高い耐性を持つと言える.

次に OCR などの文字読み取りアルゴリズム用いた攻撃に対する耐性について示す. 一般的な文字列 CAPTCHA は文字列 にノイズや歪みなどを付加した画像を解答者に提示し, その解答の正誤により人間か機械かを判別する. 文字列 CAPTCHA は画像中の文字列を判読できるか試すものであるため, OCR など文字読み取りアルゴリズムを用いた攻撃によって突破される危険性がある. 提案手法では解答者に対して提示される画像は文字列が判読不可能な状態で提示される. そのため提示された状態の画像に OCR を用いても文字を読み取ることはできない. また画像の CATMAP 変換を行い文字情報の検索を行うことも可能だが, 文字情報を持つ画像の判別, また復元の途中に現れる文字の重なり合った状態の画像と元画像の判別を行わなくてはならないため, 機械による既存の攻撃方法では提案手法の解答は困難であると考えられる.

5. まとめ

本稿では CATMAP の画像変換を用いた CAPTCHA システムを提案した. 提案手法は画像認識能力と文字読み取り能力の両方を試すため, データベース攻撃や画像認識攻撃に対して耐性があると考えられる. 利便性においてはリレーアタックなどの人間による攻撃を考慮したうえで, 難易度の設定やユーザインタフェースの開発を行う必要性がある. またシステムの堅牢性を確保するために, 現在は固定されている画像の縦横ピクセル数を可変なものにしたり, 実装時の有用性を考慮し, 画像のアスペクト比を 1:1 ではないもののアルゴリズム構築などが今後の課題といえる.

参考文献

[Elson 2007] J.Elson, J.Douceur, J.Howell, J.Saul: Asirra; a CAPTCHA that exploit interest-aligned manual image categorization, 2007 ACM CSS, pp.535-542, 2007.

[山本 2009] 山本 匠, 西垣 正勝, J.D.Tygar: 機械翻訳の違和感を用いた CAPTCHA の提案, 情報処理学会研究報告. CSEC, [コンピュータセキュリティ]. 2009(37), p. 1-8, 2009

[Bursztein 2010]E. Bursztein, S. Bethard, J. C. Mitchell, D. Jurafsky, and C. Fabry, “How good are humans at solving CAPTCHAs? a large scale evaluation,” Proc. IEEE S&P, 2010.

[小林 2010]小林司, 藤堂洋介, 森井昌克:画像認識の困難性を利用した CAPTCHA 方式の提案, 電子通信情報学会信学技報, pp37-42, 2010.