

Arnold's cat map を用いた CAPTCHA システムの提案

Proposal of CAPTCHA using the Arnold's cat map

久保田 善則^{*1}
Yoshinori Kubota

井上 聡^{*1*2}
Satoru Inoue

^{*1} 埼玉工業大学大学院
Graduate School of Engineering, Saitama Institute of Technology

^{*2} 埼玉工業大学
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

近年 Web サービスの発展にともない、ボットと呼ばれる自動プログラムを用いたサービスの不正利用が問題となっている。そこで、人間と機械を判別するチューニングテストの有用性が高まっている。現在もっとも一般に普及している判別システムの一つに、CAPTCHA と呼ばれる方式がある。

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かけた画像を Web ページ上に提示し、ユーザが判読できるか否かを試す。



図 1-1 Google で使用されている CAPTCHA 例

文字列画像を用いた CAPTCHA は歪んだ文字の判読が機械には困難であることを利用したチューニングテストである。しかし OCR を代表とする文字読み取りアルゴリズムの向上により、既存の CAPTCHA における脆弱性が指摘されている。そこで、文字列に付加するノイズや歪みを大きくすることで機械による読み取りを阻止する CAPTCHA が提案された。しかし、文字列の過剰な難化により機械だけでなく人間による判読の妨げになるという問題も起きてしまっている。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。本稿では、カオス理論のパイコネ変換に着目し、Arnold's cat map の画像変換を用いた CAPTCHA を提案する。

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機械による CAPTCHA の解答ができるようになった。そのため、マル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の CAPTCHA 方式とは異なる CAPTCHA が提案された。以下の 2 つの CAPTCHA を示す。

2.1 Asirra

Asirra は画像認識能力を試す CAPTCHA の一つで、猫と犬の画像を分類する方式である[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場合に人間と判定する。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機は 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を防ぐことは困難であると考えられる。



図 2-1 Asirra の認証画面例

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敵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による回避策が有効だが、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

連絡先: 久保田 善則, 埼玉工業大学大学院, 埼玉県深谷市
普済寺 1690 埼玉工業大学大学院, m2003uop@sit.ac.jp

データベース攻撃や画像認識攻撃に対して弱い傾向があるが、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。本稿では Arnold's cat map を用いた画像認識能力と文字列の読み取り能力を試す CAPTCHA を提案する。

3.1 Arnold's cat map

Arnold's cat map (CatMap)とは、パイコ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]。パイコね変換とは、パイ生地をこねるとき「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。

$$(x, y) = (x + y, x + 2y) \bmod N \quad (1)$$

一辺が N ピクセルの正方形の画像を横に 2 倍、縦に 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。

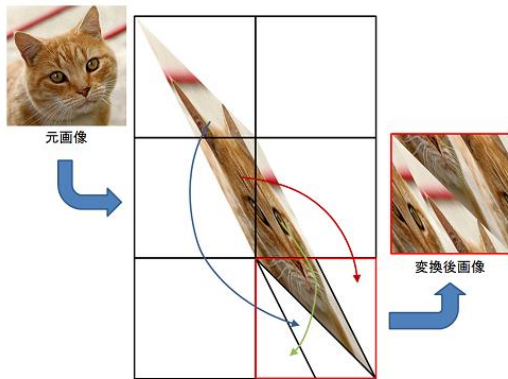


図 3-1 CatMap によるパイコね変換

図 3-1 の様に CatMap による変換を行った画像はピクセルの色の分布と画像サイズを保ったまま、オリジナルとは異なる画像に変換される。同様に変換後の画像にも再び CatMap のパイコね変換を行うと、画像はさらに細分化されオリジナルとは似ても似つかない画像へと変換される。繰り返し変換を行っていくにつれて画像はノイズのような状態に変化していくが、処理を一定回数行うと画像はオリジナルの状態に復元される(図 3-2)。



図 3-2 CatMap の繰り返し処理例

3.2 概要

提案手法は CatMap により変換された画像を解答者が復元し画像中の文字列を答えることで画像の判別能力と文字の読み取り能力を問う CAPTCHA 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する。提案手法の出題から解答までの手順は次の通り。

出題者は出題する単語をランダムに選択し、正方形の画像に単語を配置して元画像を作成する(図 3-3)。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。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 ステップずつ変換を行い、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。

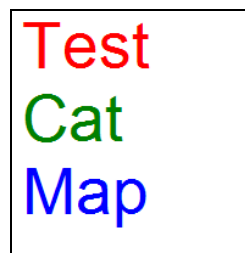


図 3-3 元画像

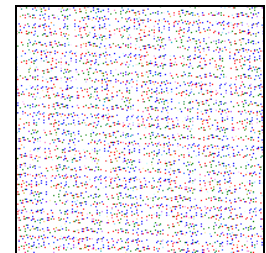


図 3-4 変換画像

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 3-5)。

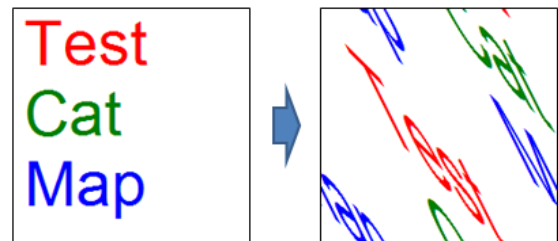


図 3-5 CatMap による画像変換(1 ステップ)

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 ピクセルの正方形画像での 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は 30 ステップから 120 ステップの範囲でランダムに変換を行った。

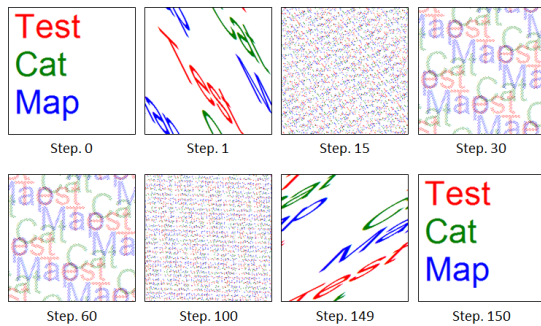


図 3-6 CatMap による画像変換(0~150 ステップ)

3.4 質問形式

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。出題から解答までの流れを以下に示す。

【出題】

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 3-7)。

【解答】

- 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、30 ステップ周期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。

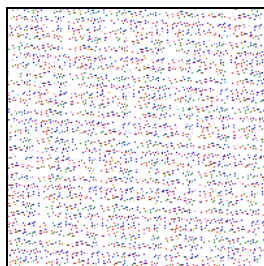


図 3-7 出題画像例



図 3-8 30 ステップ周期画像例

4. 検討

4.1 利便性

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。

CAPTCHA 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 単語を配置した元画像を用いる。元画像の変換回数は 30 ステップから 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 実験画面イメージ

実験を行った結果、提案手法の所要時間は平均 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が、実用的な所要時間から逸脱するものではないといえる。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%と画像 CAPTCHA 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

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、この時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 に有効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 に置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を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。

ボットなどを用いた自動プログラムによる攻撃のほかに、リレーアタックと呼ばれる人間の能力を用いた CAPTCHA の解読による Web サービスの不正利用も問題視されている。リレーアタックとは、図となる Web サイトで利用者を集め解読したい CAPTCHA を利用者に解答させる手法や、大量の人員を雇い解答を行わせる手法である。解答者が人間である以上、これらの手法を用いた攻撃に対して根本的な対策は見つかっていない。提案手法でも画像内に配置する単語を日本語などの別言語を用いることで解答できる人間の限定や、復元ステップ数の多いピクセル数の画像を用いて解答時間の遅延を図るなどの対策が可能だが本質的な解決を行うためには今後も検討を行う必要がある。CAPTCHA として解答者に提示される画像は提案手法の場合、そのままでは解答を行うことは出来ない。そのため CAPTCHA 内の画像のみをリレー(送信)する手法に対してはある程度の耐性を見込めるのではないかと考える。

5. まとめ

本稿では CatMap の画像変換を用いた CAPTCHA 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がある。

今後の課題としては、OCR を用いた文字情報の検索による攻撃に対する対策や、解答者の負担軽減などが挙げられる。また提案手法の発展形として元画像に背景を挿入する手法や文字列にノイズなどを付加する手法についても評価する必要がある。

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
- [Peterson 1997] Gabriel Peterson: Arnold's cat map. Math45-Linear algebra http://online.redwoods.cc.ca.us/instruct/darnold/maw/c_atmap.htm, 1997.
- [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.