SURE 静岡大学学術リポジトリ
Shizuoka University REpository

4コマ漫画CAPTCHA：
マルウエアを排除する究極のチューリングテスト

メタデータ	言語: ja 出版者: 静岡大学 公開日: 2013-01-09 キーワード (Ja): キーワード (En): 作成者: 西垣, 正勝 メールアドレス: 所属:
URL	http://hdl.handle.net/10297/7002

様式Ｃ－１９

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成２４年５月１日現在

機関番号：１３８０１
研究種目：挑戦的萌芽研究
研究期間：2009～2011
課題番号：21650015
研究課題名（和文）４コマ漫画ＣＡＰＴＣＨＡ―マルウエアを排除する究極のチューリングテスト―
研究課題名（英文）Four-panel cartoon CAPTCHA - A turing test that is impossible for malwares to solve-
研究代表者
西垣　正勝（MASAKATSU NISHIGAKI）
静岡大学・創造科学技術大学院・教授
研究者番号：20283335

研究成果の概要（和文）：本研究では、WEB サービスを不正利用するマルウエア（悪意の自動プログラム）を排除するために、人間の最も高度な認知処理能力の一つである「ユーモアを解する能力」を利用した究極のチューリングテストを構築し、4 コマ漫画 CAPTCHA として実装する。近未来の技術を持ってしてもユーモアを解するレベルの自動機械（マルウエア）を実装することは不可能に近いと推測されるため、４コマ漫画 CAPTCHA の攻撃耐性は極度に高いと考えられる。また、漫画を読むことは人間にとって楽しい（エンターテイメント性を有している）ため、4 コマ漫画 CAPTCHA であれば、正規のユーザが利便性の低下を感じることなく、心地良く（楽しみながら）チューリングテストを受けることができる。

研究成果の概要（英文）：Conventional CAPTCHAs could be overcome by state-of-the-art malwares since the capabilities of computers are approaching those of humans. Therefore, CAPTCHAs should be based on even more advanced human-cognitive-processing abilities. In addition, it is also important to keep in mind that answering CAPTCHAs is an added annoyance for users. So, CAPTCHAs should be enjoyable for users. To cope with these issues, we focused on the human ability to understand humor which is considered one of the most advanced human cognitive processing abilities, and studied a new type of Turing test that uses four-panel cartoons, which would make CATPCHAs fun and enjoyable.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
2009 年度	900,000	0	900,000
2010 年度	1,100,000	0	1,100,000
2011 年度	1,000,000	300,000	1,300,000
総　計	3,000,000	300,000	3,300,000

研究分野：総合領域
科研費の分科・細目：情報学・計算機システム・ネットワーク
キーワード：チューリングテスト、WEB セキュリティ、マルウエア検知、CAPTCHA、ユーモア

１．研究開始当初の背景

WEB サービスの発展にともなって、人間と機械を識別するチューリングテストの有用性が益々高まっている。無料 WEB メールやブログなどのインターネットにおける WEB サービス提供サイトに対し、自動プログラム（マルウエア）を使って、大量にアカウントを不正取得する、多数のブログサイトにスパム記事を不正投稿する、大量に不正なサービス利用要求を行うなどのいわゆる「DoS：Denial of Service（サービス不能）」攻撃が定常的に頻発しているためである。チューリングテストは、このようなマルウエア（悪意の自動プログラム）と正規のユーザ

（人間）を識別するために必須の技術であり、現在、CMU の研究者によって開発された「CAPTCHA [1]」と呼ばれる方式が広く利用されている。

CAPTCHA の基本形態は、歪曲やノイズが付加された文字列画像を WEB ページに提示し、閲覧者がその文字を判読できるか否かを試すものである。この例を図 1 に示す。また、音声などを利用した CAPTCHA も利用されている。

図 1.Google で使用されている CAPTCHA

しかし，近年，既存の CAPTCHA における脆弱性が多くの研究者によって指摘されている．例えば，文字列の判読能力を試す CAPTCHA においては，すでに高機能な OCR（自動文字読取）機能を備えるマルウエアが出回るようになっている[2]。文字列に加える変形やノイズを大きくすることによってマルウエアを排除する確率を向上させることはできるが、そのような文字は人間にとっても難読度が高まるため、人間の正答率まで低下させてしまう。この問題に対し、人間の「より高度な知識処理」を利用して CAPTCHA を強化する方法が検討されてきた[3]。

その代表的なものとして Asirra [4]がある。Asirra では、複数の動物の絵を表示し、その中から特定の動物の絵を選ばせる。例えば「猫を選べ」という質問に対し、猫の絵を正しく選択することができれば人間であるとして判定する。「絵の意味を理解する」ことは人間の高度な認知メカニズムの一つであり、マルウエアによる不正解答は不可能であると考えられていた。だが、最近になって、Asirra を破る自動プログラムに関する研究報告がなされ、研究者の間に衝撃が走った[5]。

マルウエアの能力の向上は留まるところを知らない。マルウエアがいかに高度になろうとも、マルウエアによる不正解答が根本的に不可能である「究極的なチューリングテスト（CAPTCHA）」がいよいよ必要とされる時代になってきた。

また、一方で、正規のユーザ（人間）にとっては、自分が人間であることをわざわざ示さなければいけないという意味では、チューリングテスト（CAPTCHA）に解答することは、本来は不要の「煩わしい手間」である。よって、チューリングテスト（CAPTCHA）は、正規のユーザ（人間）にとって「心地良い」ものでなければならないという要求も満たす必要がある。

参考文献：

[1] The Official CAPTCHA Site, http://www.captcha.net.
[2] J.Yan, A.S.E.Ahmad: Breaking Visual CAPTCHAs with Naïve Pattern Recognition Algorithms, 2007 Computer Security Applications Conference, pp.279-291, 2007.
[3] J.Elson, J.Douceur, J.Howell, J.Saul: Asirra: a CAPTCHA that exploits interest-aligned manual image categorization. 2007 ACM CSS, pp.366-374, 2007.
[4] P.Golle: Machine Learning Attacks Against the ASIRRA CAPTCHA, 2008 ACM CSS, 2008.
[5] CAPTCHA 認証は“終わった”技術なのか、月刊 Computerworld 2008 年 10 月号

２．研究の目的

WEB サービスを不正利用するマルウエア（悪意の自動プログラム）を排除するために、人間の最も高度な認知処理能力の一つである「ユーモアを解する能力」を利用した究極のチューリングテストを構築し、4 コマ漫画 CAPTCHA として実装する。近未来の技術を持ってしてもユーモアを解するレベルの自動機械（マルウエア）を実装することは不可能に近いと推測されるため、4 コマ漫画 CAPTCHA の攻撃耐性は極度に高いと考えられる。また、漫画を読むことは人間にとって楽しい（エンターテイメント性を有している）ため、4 コマ漫画 CAPTCHA であれば、正規のユーザが利便性の低下を感じることなく、心地良く（楽しみながら）チューリングテストを受けることができる。

３．研究の方法

(1) 総当り攻撃耐性の向上

4 コマ漫画 CAPTCHA の基本形（以下、基本方式）においては、4 コマ漫画の各コマをランダムに並べ替えて表示し、正しい順序を答えることができた者を人間として判定する。だが、この場合は、解答の組み合わせは 4!通りしかないため、1/24 の確率でマルウエアが偶然に正答を返すことができる。このため、より効果的な 4 コマ漫画 CAPTCHA の構成法を検討する必要がある。

具体的には、以下の 3 つの方法を検討した。

①改良方式 1：単純な CAPTCHA を独立に複数回繰り返す方法。

提示する 4 コマ漫画をランダムに変更しながら基本方式を規定のターン数（t）繰り返すことで、総当たり数を向上させる。CAPTCHA の総当たり数は、24 通りから 24^t 通りに増加する．t を大きくすればするほど総当たり数は増加し、マルウエアが CAPTCHA をパスすることが困難になる一方で、正規ユーザによる回答時間は t 倍に増

大する。
②改良方式 2：ダミー（囮）を混ぜる方法。
基本方式で用いられる 4 コマ漫画に d 個の異なるコマ（それぞれのコマは異なる 4 コマ漫画から 1 コマずつランダムに抽出）をダミーとして利用する。すなわち CAPTCHA 画面には 1 つの 4 コマ漫画から抽出された 4 コマおよび d 個のダミーのコマがランダムな順で並べられる。CAPTCHA の総当たり数は、24 通りから $_{d+4}P_4$ 通りに増加する。d を大きくすればするほど総当たり数は増加し、マルウエアが CAPTCHA をパスすることが困難になるが、その分、ユーザの識別負荷が増大する可能性がある.
③改良方式 3：複数の 4 コマ漫画を混ぜ合わせる方法。
異なる n 個の 4 コマ漫画をそれぞれランダムに混ぜ合わせ、計 4n 個のコマを一度にユーザに提示する。n 個の 4 コマ漫画全てを正しく並び替えることができたユーザのみ人間として判定される。CAPTCHA の総当たり数は、24 通りから(4n)!/n!通りに増える。

(2) 利便性に関する評価
4 コマ漫画 CAPTCHA（基本方式）におけるエンターテイメント性についての評価が未実施であったため、アンケートを通じて 4 コマ漫画 CAPTCHA の利便性を調査する必要がある。
具体的には、以下の 4 つの項目について 5 段階評定のアンケート調査を行った。
①理解のし易さ：
CAPTCHA の問題を認識すること、および、その問題に回答することが、ユーザにとって容易であるか。
②煩わしさ：
CAPTCHA に回答する作業が、（例えば時間がかかる等の理由で）ユーザは面倒に感じるか。
③何度もやりたいか：
CAPTCHA に回答する作業を複数回行うことに関して、ユーザが許容できるか。
④楽しさ：
CAPTCHA に回答する作業が、ユーザにとって楽しいか。

４．研究成果
(1) 総当り攻撃耐性の向上
改良方式 1～3 に対する実験結果を表 1 に示す。比較のために、文字 CAPTCHA および基本方式の実験も行っている。表中「成功率」は、各方式において CAPTCHA の回答に成功した割合である。「回答時間の平均」および「回答時間の標準偏差」は, CAPTCHA 画面が表示されてから被験者が回答をし終えたまでの時間の平均と標準偏差を、方式ごとにそれぞれ示したものである。「総当たり数」とは，各方式における入力の組み合わせ総数を示している。
改良方式 1 の「成功率」および「回答時間の平均」は，基本方式の「成功率」を t 乗、「回答時間の平均」を t 倍することで、それぞれ試算することができる。ユーザに提示される全コマ数が改良方式 2（d=4）および 3（n=2）と同等になるよう、改良方式 1 における繰り返し回数 t を 2 とする。
表 1 より、基本方式に改良を加えることで、総当たり数を増やすことに成功している一方、回答に要する時間も増えていることが見て取れる。成功率に関してはいえば、改良方式のほうが基本方式よりも優れた結果が得られているケースもあることがわかる。しかし、4 コマ漫画の種類や実験の順序効果の影響に引きずられている可能性もあり、本結果からだけでは、改良方式により成功率が向上したとは言い切れない。
今後、更なる実験を通じ、4 コマ漫画 CAPTCHA の構成法を検討していきたい。

表 1.実験結果

		成功率 (%)	回答時間の平均 (秒)	回答時間の標準偏差 (秒)	総当り数
文字 CAPTCHA		92.86	12.63	6.72	26^7
基本方式		82.14	26.59	14.29	24 (4!)
改良方式 1	t=2	67.47	53.18		576 (4!×4!)
改良方式 2	d=1	96.43	28.24	11.35	120 ($_5P_4$)
	d=2	96.43	35.89	18.00	360 ($_6P_4$)
	d=3	67.86	42.84	19.94	840 ($_7P_4$)
	d=4	78.57	41.70	18.73	1680 ($_8P_4$)
改良方式 3	n=2	82.14	51.10	21.17	20160 (8!/2!)

(2) 利便性に関する評価
アンケート結果を表 2 に示す。比較のために、文字 CAPTCHA に対するアンケートも行っている。
表 1 より 4 コマ漫画 CAPTCHA は文字列 CAPTCHA に比べ 2 倍～4 倍も回答に時間を要していることが見て取れ、アンケート調査（表 2）で 4 コマ漫画 CAPTCHA に対し「煩わしい」と回答した被験者が多くいることが

理解できる。しかし、「煩わしさ」以外の項目においては、文字列 CAPTCHA と同程度か、より優れているという結果が得られており、全体的には 4 コマ漫画 CAPTCHA のほうが良い結果となっている。

「理解のし易さ」に関しては、4 コマ漫画 CAPTCHA も文字列 CAPTCHA も同程度の結果が得られている。しかし、昨今の高度な OCR 機能を持ったマルウエアに対抗するためにも、文字列 CAPTCHA の難読度は日に日に高まってきており、文字列 CAPTCHA の「理解のし易さ」は今後急速に低下していくのではないかと予想される。

一方、「何度もやりたいか」や「楽しさ」の項目については、4 コマ漫画 CAPTCHA のほうが文字列 CAPTCHA に比べ、高い評価を得ている。このことから、4 コマ漫画 CAPTCHA は、正規のユーザが若干の利便性の低下を感じる（煩わしさを感じる）ものの、心地良く（楽しみながら）チューリングテストを受けることができるといえる。

以上より、4 コマ漫画 CAPTCHA は、回答時間の長さや煩わしさが依然として解決すべき課題ではあるものの、それらを考慮しても、ユーザが心地良く（楽しみながら）、何度もやりたいと思える CAPTCHA であるということが確認できた。

表 2.アンケート結果

	文字 CAPTCHA		4 コマ漫画 CAPTCHA	
	平均	標準偏差	平均	標準偏差
理解のし易さ：1 し易い⇔し難い 5	2.57	1.12	2.21	0.67
煩わしさ：1 簡単⇔面倒 5	3.07	0.82	3.93	0.62
何度もやりたいか:1 やりたい⇔やりたくない 5	3.86	0.96	2.43	0.70
楽しさ:1 楽しい⇔楽しくない 5	4.43	1.06	1.57	0.98

５．主な発表論文等
（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計 3 件）

① Takumi Yamamoto, Tokuichiro Suzuki, Masakatsu Nishigaki：A Proposal of Four-panel cartoon CAPTCHA, Proceedings of IEEE International Conference on Advanced Information Networking and Applications 2011, 査読有, 巻無し, 2011, pp. 159-166
② Takumi Yamamoto, J.D.Tyagr, Masakatsu Nishigaki：CAPTCHA Using Strangeness in Machine Translation, Proceedings of IEEE International Conference on Advanced Information Networking and Applications 2010, 査読有, 巻無し, 2010, pp.430-437
③ Takumi Yamamoto, Tokuichiro Suzuki, Masakatsu Nishigaki：A Proposal of Four-panel cartoon CAPTCHA: The Concept, Proceedings of 2010 International Workshop on Trustworthy Computing, 査読有, 巻無し, 2010, pp. 575-577

〔学会発表〕（計 6 件）

① 可児潤也, 上松晴信, 西垣正勝：ワンモア CAPTCHA の提案, 2012 年暗号と情報セキュリティシンポジウム, 2012.2.1, 金沢エクセルホテル東急（石川）
② 上原章敬, 鈴木徳一郎, 山本匠, 西垣正勝：4 コマ漫画 CAPTCHA の検討, 情報処理学会研究報告, 2011-CSEC-52-13, 2011.3.3, 関西大学（大阪）
③ 山本匠, 鈴木徳一郎, J.D.Tygar, 西垣正勝：人間の高度な認知処理に基づく CAPTCHA の提案, 映像メディア学会技術報告, ME2010-173, 2010.12.16, 首都大学東京（東京）
④ 鈴木徳一郎, 山本匠, 西垣正勝：リレーアタックに耐性をもつ CAPTCHA の提案, 情報処理学会コンピュータセキュリティ研究会, 2010.3.4, 東北大学（宮城）
⑤ 山本匠, J.D.Tygar, 西垣正勝：機械翻訳 CAPTCHA（その 2), コンピュータセキュリティシンポジウム 2009, 2009.10.26, 富山国際会議場（富山）
⑥ 山本匠, J.D.Tygar, 西垣正勝：機械翻訳の違和感を用いた CAPTCHA の提案, 情報処理学会コンピュータセキュリティ研究会, 2009.7.3, 秋田大学（秋田）

〔その他〕
ホームページ：
http://minamigaki.cs.inf.shizuoka.ac.jp

６．研究組織
(1)研究代表者

西垣　正勝（MASAKATSU NISHIGAKI）
静岡大学・創造科学技術大学院・教授
研究者番号：20283335

(2)研究分担者
なし
(3)連携研究者
なし