

Evaluation and Development of Onomatopoeia CAPTCHAs

Riko Shigeno, Michihiro Yamada, Hiroaki Kikuchi (Meiji University, Japan)
Maki Sakamoto (The University of Electro-Communications, Japan)



Background

CAPTCHA

(Completely Automated Public Turing Test To Tell Computers and Humans Apart)

Type of challenge response test used for avoiding from automated registration made by malicious software.

(ex)CAPTCHA to authenticate characters



Requirement of CAPTCHA

- (1) Human can easily solve a CAPTCHA.
- (2) Computer can't solve a CAPTCHA.
- (3) CAPTCHAs can be generated automatically.

Proposal on "Onomatopoeia CAPTCHA"

Onomatopoeias

Words that mimicked sounds and conditions of things.



	Japanese onomatopoeia	English onomatopoeia
Dog bark	WanWan	arf, bow wow
Sound to eat	MoguMogu	chomp, munch

Human → Easy : Human usually understand onomatopoeia unconsciously and use it

Computer → Difficult : The mechanism to recognize onomatopoeia is not very clear still now

Use of onomatopoeias may lead to effective results.

Example proposed method

Scheme 1E(4 choices)	Scheme 2E(Image selection)	Scheme 5E (English Manga Onomatopoeia CAPTCHA)
Choose one of the best onomatopoeia representing meaning (ex)Which word feel like Fast? • Kakaka • Sassa • Bo- • Kori (Answer)Sassa	Search the image based on the dictionary and select the onomatopoeia that fits the image (ex)  • Togetoge • Yusayusa • Botoboto • Uhu (Answer)Uhu	Give hidden onomatopoeia used for Japanese comics, select one suitable onomatopoeia from the choices (ex)which word that is likely to fit in the red painted image? • Hop • Amble • Whir • Stickly  (Answer)Hop

What is good CAPTCHA?

HAR(Human Acceptance Rate) (1)

Probability that CAPTCHA correctly accepts humans
→ taking the average correct answer rate: subjects(54 Japanese, 63 foreigners)

MAR(Machine Acceptance Rate) (2),(3)

Probability that CAPTCHA wrongly accept computers
→ Google search attack(MAR_g), Brute force attack(MAR_b)

Good CAPTCHA → HAR is high , MAR is low

(2)Attack using search engine MAR_g

Perform AND search on meaning words with their respective choices and choose the option with the highest search hit count .

Which word feel like "walk"?	Search words	Number of hits
• Bakubaku • Bosoboso • Tokotoko • Gyahaha (A)Tokotoko	Walk , Bakubaku	3140
	Walk , Bosoboso	22800
	Walk , Tokotoko	56900
	Walk , Gyahaha	2770

(3)Brute Force Attack MAR_b

Choose from their respective choices at random.

Which word feel like "walk"?	Choices	Probability of selection
• Bakubaku • Bosoboso • Tokotoko • Gyahaha (A)Tokotoko	Bakubaku	1/4
	Bosoboso	1/4
	Tokotoko	1/4
	Gyahaha	1/4

Result

Styles	1 (4-out-of-1)		2		3	4 (Image selection)		5 (Manga)		
	Dictionary	SK	Dictionary	SK	Dictionary	Dictionary	SK	Dictionary (Japanese)	Dictionary (English)	Dictionary (English)
Subjects	Japanese					Japanese	Japanese	Japanese	non-Japanese	
Number of subjects	63							57	63	
Number of queries	18	3	18	3	12	24	3	10		
(3) MAR _b	1/4				1/16	-	-	1/4		
(2) MAR _g	0.400	0.333	0.277	0.333	-	-	-	-	-	-
(1) Mean HAR	0.891	0.757	0.705	0.444	0.680	0.894	0.513	0.895	0.452	0.410
Mean answering time [sec]	00:07	00:08	00:08	00:09	00:07	00:07	00:09	00:14	00:26	00:18