



بررسی امنیت و کارایی کیچای متنی و مقایسه مدل‌های مختلف آن

گزارش پروژه کارشناسی مهندسی کامپیوتر

نام دانشجو: مطهره هژبر یوسفی

شماره دانشجویی: ۹۹۰۱۲۲۶۸۰۰۱۳

ایمیل: m.hojabryousefi@gmail.com

نام استاد: دکتر فرید فیضی

دانشگاه گیلان - تابستان ۱۴۰۳

۳ مقدمه
۴ کپچای متنی
۶ روش‌های افزایش امنیت کپچا
۶ کارایی کپچا
۷ روش ساخت کپچاهای استفاده شده
۹ نویسه‌خوان استفاده شده
۱۰ ارزیابی کارایی کپچا
۱۶ بررسی نتایج کپچاهای ۵ حرفی
۱۸ بررسی یک ادعا: آیا کپچای ترکیبی امنیت بالایی دارد؟
۲۰ نتایج بررسی
۲۲ منابع

مقدمه

کیچا^۱ برنامه‌ای است که توانایی تولید آزمون‌های خودکار بر پایه آزمون تورینگ دارد و هدف از ساخت آن، تشخیص و تمایز انسان و ماشین از یکدیگر می‌باشد. امنیت کیچا به معنای این است که با وجود آگاهی از روش کار کیچا، هیچ ماشین، کامپیوتر و برنامه کامپیوتری قادر به قبولی در این آزمون نخواهد بود.

هدف از ساخت کیچا، احراز هویت کاربران کامپیوتر هنگام استفاده از خدمات مربوط به فضای وب و اینترنت است. کیچا باید بتواند از دسترسی غیر مجاز به برخی صفحات وب جلوگیری کند و همچنین افراد غیر مجاز توانایی خوانش و فهم کد کیچا را نداشته باشند. از دیگر اهداف ساخت کیچا می‌توان به موارد زیر اشاره کرد:

۱. جلوگیری از هرزنامه در ایمیل‌ها

۲. افزایش امنیت اکانت کاربر و محافظت در برابر حمله‌هایی که سعی در دستیابی به رمز عبور دارند.

۳. احراز اصالت پرسش‌نامه‌های آنلاین و اطمینان از این که حتما یک انسان به سوالات جواب می‌دهد.

در موضوع امنیت شبکه، معیارهایی از جمله احراز هویت، احراز اصالت و محرمانگی اهمیت زیادی دارند و از نظر امنیتی، هیچ برنامه کامپیوتری با وجود آگاهی از روش تولید کیچا، نباید توانایی قبولی در آزمون کیچا را داشته باشد. نقش اصلی کیچا، تمایز قائل شدن بین انسان و ماشین است و کیچا باید برای انسانها قابل درک و برای ماشین غیرقابل فهم باشد. از زمان ساخت کیچا در سال ۲۰۰۰، روش‌های متعددی برای درک کدها و شکستن کیچا تولید و بررسی شده‌اند و برخی از آنها در فهم کدهای بر پایه متن و حتی تصویری موفق عمل کرده‌اند. به طور کلی، کیچا یک یا چند متن، عکس یا عدد را تولید کرده و از کاربر می‌خواهد که با توجه به آنها عملی را انجام دهد، برای مثال متن دیده شده را بنویسد یا روی بخشهایی از تصویر کلیک کند برای افزایش امنیت کیچای متنی، روش‌های مختلفی از جمله تغییر فونت، اعوجاج حروف و انجام تبدیل‌های هندسی روی متون وجود دارد و در سال‌های اخیر، برای ایمن تر کردن کیچا الگوریتم

¹ Captcha

های مختلفی معرفی شده است. [۱] در این گزارش نمونه‌هایی از کپچای متنی به صورت حروف انگلیسی و به صورت معادلات جبری و ترکیبی از آنها تهیه و بررسی شده است و نتایج بررسی از نظر امنیت و کارایی آورده شده است.

کپچای متنی

کپچای متنی به دلیل سادگی در تولید، الگوریتم‌های تولید سریع و هزینه پایین، بیشترین استفاده را دارد. تاکنون نرم‌افزارهای متعددی برای شناسایی کپچاهای متنی ساخته شده‌اند که اکثر آنها از سه مرحله برای شناسایی کپچا استفاده می‌کنند:

۱. انجام پیش پردازش روی تصویر

۲. قطعه‌بندی عکس به نواحی مختلف که هر ناحیه شامل تنها یک کاراکتر باشند.

۳. شناسایی هر کاراکتر در هر ناحیه

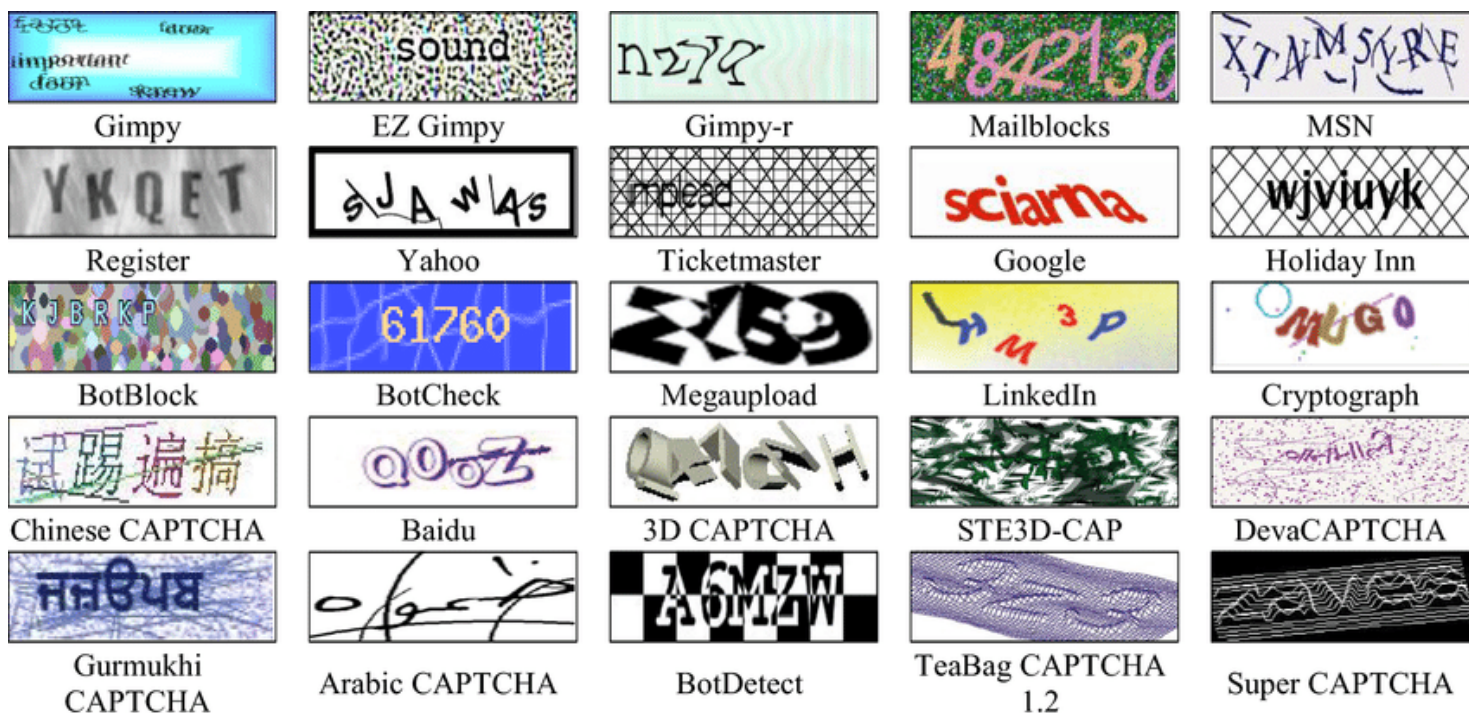
به منظور پیچیده کردن روند شناسایی کپچا برای نرم‌افزارهای کامپیوتری، برخی از روش‌های افزایش امنیت کپچا به کار برده می‌شود تا از خوانش و شناسایی آنها توسط ماشین جلوگیری شود. تاکنون الگوریتم‌های متعددی برای تولید کپچای متنی معرفی شده‌اند. گیمپی^۲ یکی از ابتدایی‌ترین الگوریتم‌های تولید کپچا است که از کلمات رایج انگلیسی با کمی تغییر روی حروف استفاده می‌کند که امنیت پایینی دارد و به کمک نویسه‌خوان‌ها^۳ قابل شناسایی و شکستن است. کپچای مگاآپلود^۴ از ترکیب حروف مجاور با یکدیگر استفاده می‌کند یا کپچای یاهو^۵ از حروف پراکنده و خطوط مورب برای افزایش امنیت کپچا استفاده می‌کند. در شکل ۱ نمونه‌های رایجی از کپچاهای متنی و ویژگی‌های آنها نمایش داده شده است.

^۲ Gimp

^۳ OCR

^۴ MegaUpload

^۵ Yahoo



شکل ۱ - انواع کپچای متنی

امنیت یک کپچا وابسته به توانایی جلوگیری از نفوذ و حمله باتها است. یک کپچا باید در ضمن قابل فهم بودن برای انسان، برای ماشین پیچیده و غیرقابل خوانش باشد. با این وجود، در صورتی که کپچا بسیار پیچیده باشد، ممکن است برای انسان غیرقابل درک بوده و از دسترسی افراد مجاز به اشتباه جلوگیری کند. سه عدد از ویژگی-های اصلی و مهم کپچا عبارتند از:

۱. قابل خوانش وقابل فهم توسط انسانها باشد.

۲. غیرقابل حدس زدن و پیشبینی باشد.

۳. از نظم مشخصی برخوردار باشد. برای مثال حروف انگلیسی در کپچا از چپ به راست خوانده شوند و این نظم در کپچا مشهود باشد. [۲]

همچنین از نظر امنیتی، کپچا نباید به برنامه کامپیوتری اجازه دهد تا در آزمون کپچا قبول شود و عبارت داده شده را شناسایی کند و حل آن باید تنها برای انسان آسان باشد. همچنین نباید معیارهای دیگر امنیت شبکه مانند محرمانگی اتصال کاربر را به خطر بیاندازد و نباید از داده‌های مربوط به کاربر برای تولید کپچا استفاده کند [۳].

روش‌های افزایش امنیت کپچا

برای افزایش امنیت کپچا و حفظ ویژگی‌های ذکر شده در کپچا، راه‌حل‌های مختلفی ذکر شده است که از بسیاری از آنها در تولید کپچاهای متنی امروزی استفاده می‌شود. در تولید کپچای متنی دو ویژگی کلی از کپچا هستند که معیار سنجش قدرت کپچا محسوب می‌شوند که عبارت‌اند از:

۱. کاراکترها: روش‌های منظم کردن، تغییر شکل و ویژگی‌های هر کاراکتر از کپچا بررسی می‌شود. مانند تغییر فونت، تغییر رنگ کاراکتر، چرخش حروف و... [۴].

۲. نویز: تغییرات ایجاد شده روی کاراکترها که برای انسان قابل فهم باشد. از انواع نویزهای قابل استفاده می‌توان به نویز در پس‌زمینه، تغییر رنگ پس‌زمینه، خطوط و اشکال مختلف در کپچا و... اشاره کرد. در شکل ۲ نمونه‌ای از یک کپچای متنی که در این گزارش استفاده شده، با تغییرات ذکر شده مشاهده می‌شود.



شکل ۲- کپچای متنی

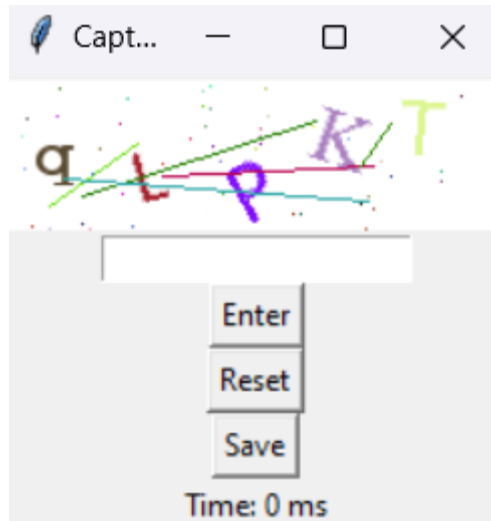
کارایی کپچا

مفهوم کارایی کپچا در ارتباط نزدیک با امنیت کپچاست، به این معنی که یک کپچای کارآمد و خوب از امنیت بالایی برخوردار می‌باشد و نویسه‌خوان‌ها قادر به شکستن آن نخواهند بود. کارایی یک کپچا در کنار ایمن بودن آن، به سهولت در استفاده و فهم توسط انسان نیز بستگی دارد، و یک کپچای غیرقابل درک توسط انسان دلیل استفاده از کپچا را به طور کامل زیر سوال خواهد برد. در نتیجه تنها از جانب ایمن بودن در برابر

نویسه‌خوان نمی‌توان یک کپچا را بررسی کرد، و باید توسط انسان پاسخ داده شوند و مدت زمان رسیدن به پاسخ صحیح، تعداد دفعات تکرار جواب و مواردی از این قبیل در ارتباط با کارایی کپچا بررسی شوند. کپچایی که بتوان در زمان کمتری توسط انسان به پاسخ آن دست پیدا کرد، کارایی بالاتری نسبت به کپچایی پیچیده دارد که کاربر در زمانی کوتاه نتواند پاسخ آن را بیابد.

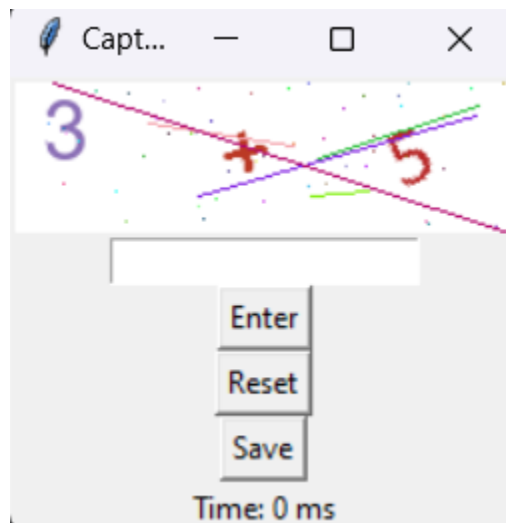
روش ساخت کپچاهای استفاده شده

در این بخش به الگوریتم تولید کپچا به زبان پایتون پرداخته خواهد شد. به طور کلی، برای ساخت کپچا نیاز به کتابخانه‌های مربوط به پردازش تصویر و متون خواهیم داشت که برای این پروژه از کتابخانه Pillow استفاده شده است. به دلیل نیاز به بررسی کارایی و امنیت هر کپچا، نیاز به ساخت یک GUI ساده با استفاده از tkinter برای بررسی ویژگی‌های مختلفی نیاز بوده که در این پروژه پیاده‌سازی شده است. بدین ترتیب، با اجرای برنامه پنجره‌ای باز می‌شود که یک کپچا به همراه یک باکس برای وارد کردن جواب و سه دکمه enter, reset, save و یک تایمر دیده می‌شود. کاربر پس از پاسخ به کپچا دکمه enter را می‌زند تا پاسخ او بررسی شود و پیام درست بودن یا نبودن پاسخ به او نشان داده می‌شود. همچنین امکان reset کردن کپچا در صورت ناخوانا بودن بعضی حروف وجود دارد، و برای ذخیره و استفاده از کپچا برای نویسه‌خوان یا موارد دیگر می‌توان دکمه save را استفاده کرد. تایمر نیز به میلی‌ثانیه زمان ایجاد کپچا را تا اولین پاسخ صحیح محاسبه می‌کند تا برای محاسبه کارایی کپچا بتوان از آن استفاده کرد. برای تولید کپچا و افزایش امنیت آن، از ۴ فونت مختلف برای تولید هر حرف استفاده شده است. همچنین رنگ نویز پس‌زمینه و رنگ هر کاراکتر، درجه چرخش هر حرف و بزرگ یا کوچک بودن هر حرف به طور رندوم در کپچا تعیین می‌شود. برای بعضی حروف انگلیسی مثل V که بزرگ و کوچک بودن آن مشابه است و ممکن است باعث خطای کاربر شود، مشخص می‌کنیم که اندازه فونت حروف کوچک، کمی کوچکتر از اندازه فونت حروف بزرگ باشند و تا جای ممکن هیچ دو حرف یکسانی کنار یکدیگر تولید نشوند. نمونه کپچای تولید شده و پنجره باز شده در شکل ۳ نشان داده شده است.



شکل ۳ - کپچای متنی با ۵ حرف انگلیسی

در کنار کپچای ۵ حرفی استفاده شده، یک کپچای دیگر نیز تولید شده است که به صورت یک معادله جبری ساده و جمع دو عدد تک رقمی می‌باشد و کاربر باید پاسخ معادله را در کادر وارد کند. بقیه اجزا مشابه کپچای قبلی هستند.

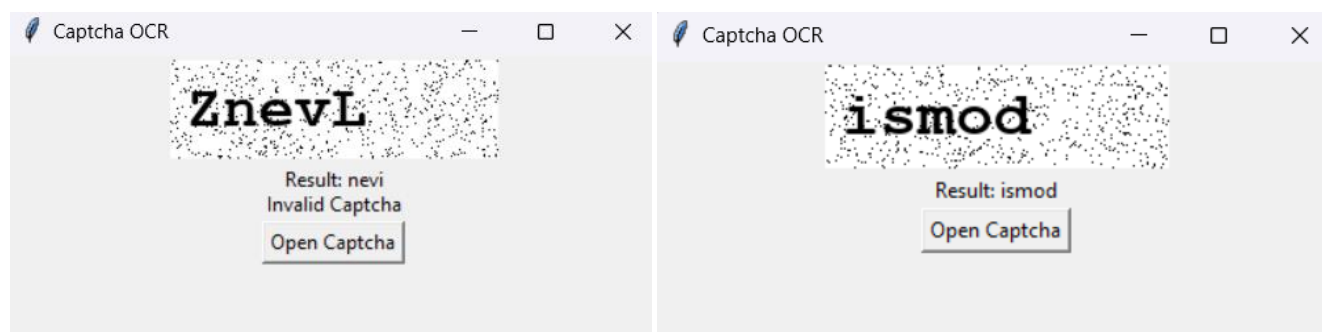


شکل ۴ - کپچای معادله جبری

نویسه‌خوان استفاده شده

برای ارزیابی امنیت کپچا نیاز به استفاده از یک نویسه‌خوان داریم که کپچاهای تولید شده را ورودی بگیرد و اقدام به شناسایی حروف موجود کند. برای این کار یک نویسه‌خوان معروف به نام Tesseract OCR استفاده خواهیم کرد که یکی از دقیق‌ترین مدل‌های موجود است و از ورژن ۵.۴۰۰ برای این کد استفاده خواهیم کرد. [۵] ابتدا تصویر یک کپچا را به فرمت png به آن ورودی خواهیم داد و سپس نویسه‌خوان اقدام به شناسایی کاراکترهای موجود خواهد کرد. در صورتی که موفق به شناسایی تمام حروف شود پیغام درست بودن کپچا را نشان می‌دهد، در غیر این صورت حروف شناسایی شده را نشان خواهد داد و پیغام نادرست بودن فرمت کپچا را نشان خواهد داد. (برای مثال اگر یک کپچای ۵ حرفی را ورودی بگیرد و تنها ۳ حرف آن را شناسایی کند، نمی‌تواند کپچا را به طور کامل بشکند و فرمت ۵ حرفی آن را شناسایی نکرده است، اما در صورت شناسایی تمام حروف می‌تواند کپچا را بشکند.)

برای ارزیابی قدرت نویسه‌خوان استفاده شده، از دو نوع کپچای ساده به سبک ez-gimpy استفاده میکنیم که به ترتیب شامل ۵ حرف انگلیسی و یک معادله جبری ساده هستند.



شکل ۵ - دو نمونه شناسایی حروف در کپچای ۵ حرفی ساده توسط نویسه‌خوان

ارزیابی کارایی کپچا

برای بررسی کارایی کپچا، به طور کلی سه متغیر خواهیم داشت. زمان ایجاد کپچا تا اولین پاسخ صحیح، تعداد دفعات تلاش کاربر برای رسیدن به اولین پاسخ صحیح و تعداد کاراکتر شناسایی شده توسط نویسه‌خوان. هر چه زمان طی شده تا اولین پاسخ صحیح کمتر باشد، کپچا برای انسان قابل فهم‌تر بوده است و کارایی بالاتری دارد. همچنین کپچایی که در اولین دفعه پاسخ داده شود به مراتب کارایی بیشتری نسبت به کپچایی است که کاربر را سردرگم کرده و او چندین دفعه مجبور به پاسخ می‌شود. همچنین شناسایی همه یا بعضی از حروف توسط نویسه‌خوان امنیت کپچا را زیر سوال می‌برد، بنابراین هرچه حروف کمتری شناسایی شوند کپچا از امنیت و کارایی بالاتری برخوردار است. در ادامه فرمول محاسبه کارایی کپچا آورده شده است:

$$E = \frac{1000}{T \cdot A \cdot (L + 1)^2}$$

E: میزان کارایی کپچا



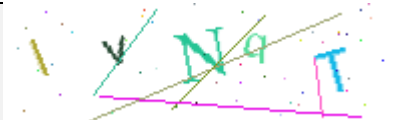
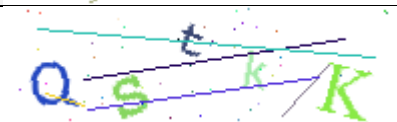





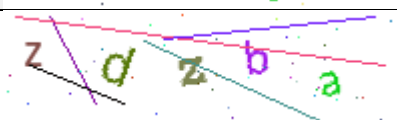
T: زمان محاسبه شده از ایجاد کپچا تا جواب صحیح توسط کاربر به ثانیه

A: تعداد دفعات وارد کردن پاسخ توسط کاربر تا اولین جواب صحیح


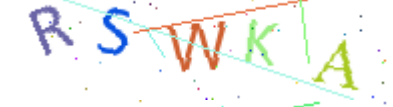








L: تعداد حروف تشخیص داده شده توسط OCR

به دلیل اهمیت بالای خوانده نشدن توسط نویسه‌خوان، از توان ۲ برای بالا بردن وزن آن در عبارت استفاده می‌کنیم. در ادامه، در جدول ۱ تا جدول ۵ برای ۱۰ نمونه از هر نوع کپچا، کارایی را محاسبه کرده و نتایج را بررسی می‌کنیم.

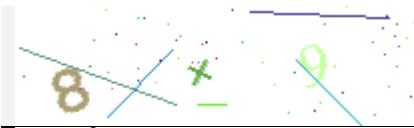
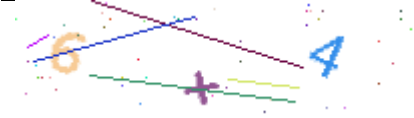
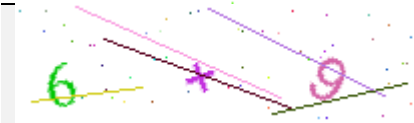
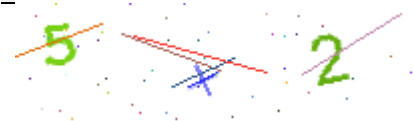
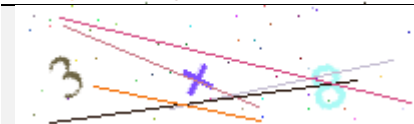

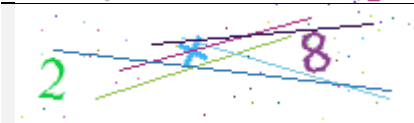
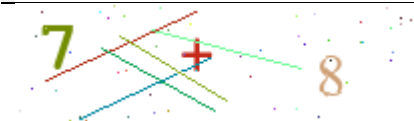

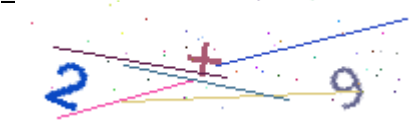
جدول ۱ - کارایی کیچاهای ۵ حرفی با حروف بزرگ و کوچک انگلیسی

کیچا	پاسخ کیچا	زمان پاسخ	تعداد حروف شناسایی شده توسط OCR	تعداد دفعات پاسخ	کارایی
	ObdDS	5.4	1	1	46.3
	aiIzf	6.3	0	1	158
	lvNqT	6.6	0	1	151
	QStkK	14.9	1	2	8.3
	axEBt	4.4	1	1	56.8
	gKC uF	13.4	0	2	37.3
	xhAoi	4.6	0	1	217
	wcQYu	4.8	0	1	208
	bYbLi	4.7	2	1	23.6
	zdzba	4.7	0	1	212.7

جدول ۲ - کارایی کپچاهای ۵ حرفی با حروف بزرگ انگلیسی

کپچا	پاسخ کپچا	زمان پاسخ	تعداد حروف شناسایی شده توسط OCR	تعداد دفعات پاسخ	کارایی
	DEHVJ	4.2	0	1	238
	RSWKA	4.2	1	1	59.5
	JXNOI	4.8	0	1	208
	QKPKA	3.6	3	1	17.3
	WZORE	4.0	2	1	27.7
	GAEUE	4.1	3	1	15.2
	CTRMI	3.2	1	1	78.1
	XAKRV	4.3	2	1	25.8
	RETFX	4.9	1	1	51
	KSQCE	3.8	1	1	65.8

جدول ۳ - کارایی کپچاهای معادله جبری

کپچا	پاسخ کپچا	زمان پاسخ	تعداد ارقام شناسایی شده توسط OCR	تعداد دفعات پاسخ	کارایی
	17	2.6	0	1	384.6
	10	1.6	0	1	625
	15	1.8	0	1	555.6
	7	1.7	0	1	588.2
	11	2.3	0	1	434.8
	12	1.5	0	1	666.7
	10	1.8	0	1	555.6
	15	1.8	0	1	555.6
	17	1.9	0	1	526.3
	11	1.7	0	1	588.2

جدول ۴ - کارایی کپچاهاى ۵ حرفى گيمپى با حروف بزرگ و کوچک انگليسى

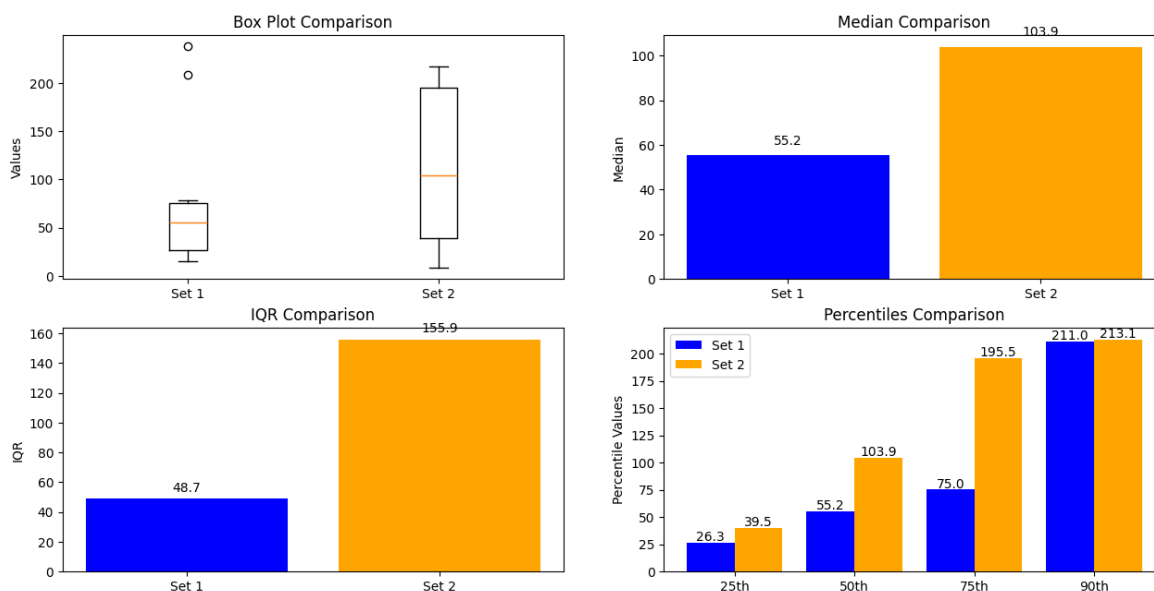
کپچا	پاسخ کپچا	زمان پاسخ	تعداد حروف شناسایی شده توسط OCR	تعداد دفعات پاسخ	کارایی
DHFai	DHFai	5.0	5	1	5.6
VTēEP	VTēEP	4.7	3	1	13.3
ismod	ismod	3.2	5	1	8.7
SoGRZ	SoGRZ	3.6	5	1	7.7
cZkDU	cZkDU	4.7	5	1	6.0
VGzKv	VGzKv	5.0	4	1	8.0
ZnevL	ZnevL	4.4	3	1	14.2
IEoYf	IEoYf	4.6	2	1	24.2
ObKwr	ObKwr	4.0	5	1	7.0
WgVEV	WgVEV	4.3	5	1	6.5

جدول ۵ - کارایی کپچاهای معادله جبری گیمپی

کپچا	پاسخ کپچا	زمان پاسخ	تعداد ارقام شناسایی شده توسط OCR	تعداد دفعات پاسخ	کارایی
9 + 3	12	1.9	1	1	131.6
8 + 1	9	1.4	0	1	714.2
6 + 3	9	1.7	1	1	147.1
2 + 9	11	1.6	2	1	69.4
7 + 1	8	1.5	0	1	666.7
6 + 5	11	1.2	2	1	92.3
8 + 8	16	1.7	0	1	588.2
7 + 8	15	1.7	0	1	588.2
1 + 7	8	1.7	0	1	588.2
2 + 2	4	1.5	2	1	74.1

بررسی نتایج کپچاهای ۵ حرفی

با بررسی میزان کارایی کپچا، می‌توانیم ابتدا مقایسه را بین دو کپچای ۵ حرفی انجام دهیم که یکی شامل حروف بزرگ و کوچک انگلیسی است و دیگری تنها شامل حروف بزرگ می‌باشد. با توجه به پراکندگی مقادیر، استفاده از میانگین برای ارزیابی کارایی شاخص قابل قبولی نمی‌باشد. برای این کار از ۴ شاخص مورد استفاده در آمار استفاده میکنیم که شامل میانه^۶، دامنه بین چارکی^۷ یا IQR، صدک^۸ و نمودار جعبه‌ای^۹ استفاده میکنیم. با استفاده از میانه می‌توان داده‌ها را مرتب کرد و چون تعداد آنها زوج است، میانگین داده ۵ و ۶ را به عنوان میانه محاسبه کرد. برای IQR فاصله بین چارک اول و سوم بررسی خواهد شد، همچنین از صدک‌های ۲۵، ۵۰، ۷۵ و ۹۰ برای بررسی متغیرها (شامل چارک اول، دوم، سوم) استفاده خواهیم کرد و نمای کلی پراکندگی داده‌ها را در نمودار جعبه‌ای نشان داده ایم. [۶، ۷] در شکل ۶ نتایج بررسی کارایی دو کپچای ۵ حرفی نشان داده شده است که Set 1 کپچای حروف بزرگ، و Set 2 کپچای شامل حروف بزرگ و کوچک می‌باشد.



شکل ۶ - نتایج بررسی کارایی به صورت نمودار

^۶ Median

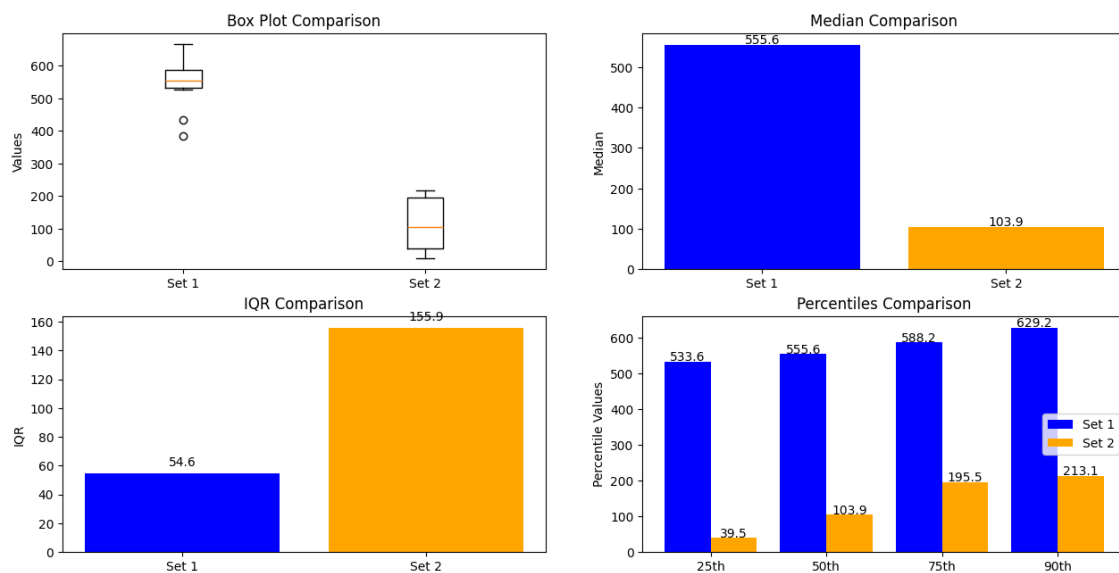
^۷ Interquartile range

^۸ Percentile

^۹ Box plot

با توجه به نتایج بالا، می توان گفت که کپچای ۵ حرفی شامل حروف بزرگ و کوچک، به طور کلی کارایی بالاتری نسبت به کپچای شامل حروف بزرگ دارد، زیرا با توجه به نمودار پراکندگی عناصر ست دوم بیشتر و مقادیر بالاتری در چارک سوم به بعد خود دارد و میانه بیشتری نسبت به ست اول داراست.

اکنون می توانیم کپچای مربوط به معادله جبری را با کپچای ۵ حرفی شامل حروف بزرگ و کوچک بررسی کنیم تا نتیجه بگیریم بین ۳ نوع کپچای ساخته شده، کدام یک از کارایی بالاتری برخوردار است. شکل ۷ نتایج بررسی کپچای جبری را با کپچای ۵ حرفی نشان می دهد که ست ۱ مقادیر کارایی کپچای جبری و ست ۲ کارایی کپچای ۵ حرفی را نمایش می دهد.

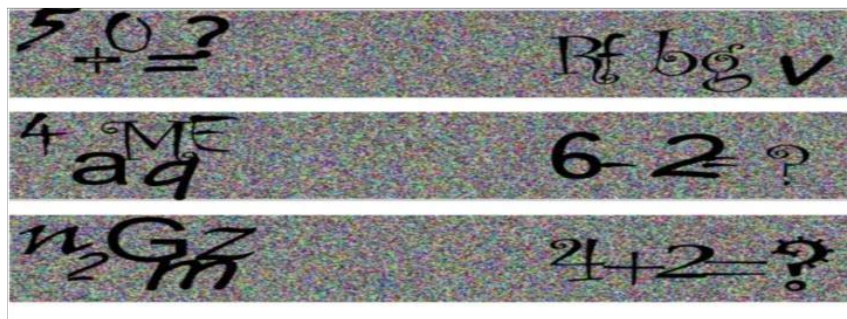


شکل ۷ - نتایج بررسی کارایی به صورت نمودار

با توجه به نمودار بالا، ست اول یعنی کپچای جبری با اختلاف زیادی از کپچای ۵ حرفی کارایی بهتری دارد. به طور کلی هنگام پاسخ به کپچای جبری، عمدتاً زمان پاسخ کمتر از ۲ ثانیه بوده است و OCR موفق به شناسایی اعداد نشده است که این ویژگی ها باعث افزایش کارایی این نوع کپچا شده است.

بررسی یک ادعا: آیا کپچای ترکیبی امنیت بالایی دارد؟

با ارجاع به مقاله‌ای [۸] در زمینه تولید کپچای متنی با امنیت بالا، این مقاله کپچای ترکیبی به کمک یک عبارت ۵ حرفی با ترکیبی از اعداد و حروف، و یک معادله جبری ساده شامل جمع یا تفریق دو عدد تک رقمی ساخته بود که از نظر امنیتی در برابر نویسه‌خوان امنیت بسیار بالایی نسبت به الگوریتم‌های گیمپی، مگا آپلود و ... داشت. پس از بررسی ۱۰۰ نمونه کپچای تولید شده، درصد شکسته شدن کپچا ۴ درصد بوده است که در مقایسه با بقیه الگوریتم‌ها امنیت بالایی دارد، اما در این مقاله در ارتباط با کارایی این کپچا و توانایی پاسخ توسط انسان مطلبی آورده نشده است. در شکل ۸ نمونه‌ای از کپچای تولید شده از الگوریتم جدید آورده شده است که در اولین نگاه، به نظر می‌رسد شناسایی کاراکترها ممکن است کاربر را با خطا مواجه کند. در این بخش می‌خواهیم نمونه‌ای ساده‌تر از کپچای ترکیبی که شامل عبارتی ۵ حرفی از حروف انگلیسی و یک معادله جبری شامل جمع دو عدد می‌باشد را ساخته و کارایی آن را بررسی خواهیم کرد و به این سوال پاسخ خواهیم داد که آیا کپچای ترکیبی از کپچاهایی که قبلاً ذکر شده‌اند، امنیت و کارایی بالاتری دارد؟ در جدول ۶ کارایی کپچاهای ترکیبی آورده شده است.



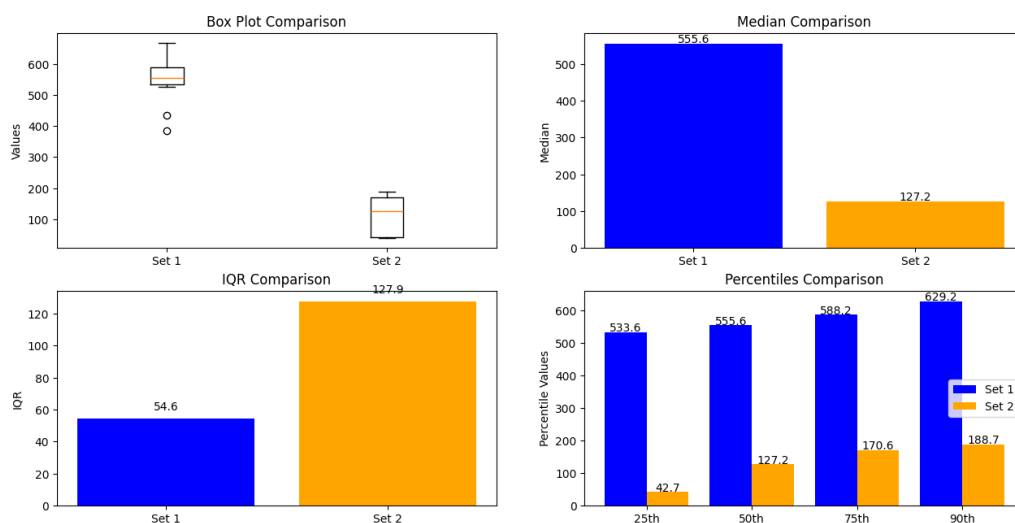
شکل ۸ - کپچای ترکیبی تولید شده توسط الگوریتم گفته شده در مقاله

جدول ۶ - کارایی کپچاهای ترکیبی

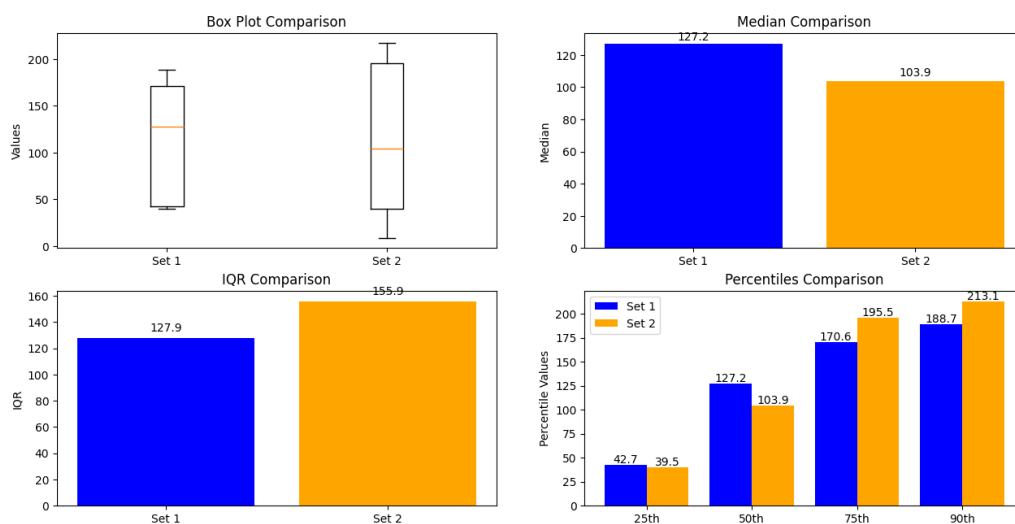
کارایی	تعداد دفعات پاسخ	تعداد شناسایی شده توسط OCR	زمان پاسخ	پاسخ کپچا	کپچا
113.6	1	0	8.8	GJCGp 10	
42.3	1	1	5.9	pbAZp 12	
175.4	1	0	5.7	WmYNc 15	
188.7	1	0	5.3	Xysmu 1	
188.7	1	0	5.3	cWnjQ 6	
42.0	2	0	11.9	lXviu 11	
140.8	1	0	7.1	GyEca 2	
39.6	2	0	12.6	Itdzq 11	
156.3	1	0	6.4	tZJPV 6	
43.9	1	1	5.7	foNUi 9	

نتایج بررسی

مقایسه کپچای ترکیبی با کپچای جبری در شکل ۹ و مقایسه کپچای ترکیبی با کپچای ۵ حرفی با کپچای بزرگ و کوچک در شکل ۱۰ آورده شده است.



شکل ۹ - مقایسه کارایی کپچای جبری (ست ۱) و کپچای ترکیبی (ست ۲)



شکل ۱۰ - مقایسه کارایی کپچای ترکیبی (ست ۱) با کپچای ۵ حرفی (ست ۲)

با توجه به نتایج بدست آمده، کپچای جبری با اختلاف زیادی کارایی بالاتری نسبت به کپچای ترکیبی دارد. همچنین در مقایسه با کپچای ۵ حرفی، میانه بیشتری دارد اما فاصله بین چارکی آن کمتر می‌باشد. در بعضی موارد مخصوصاً از چارک سوم به بالا، کارایی کپچای ۵ حرفی بهتر از کپچای ترکیبی می‌باشد. از نظر بررسی امنیت، میانگین تشخیص حروف توسط نویسه‌خوان در ۱۰ نمونه از کپچای ترکیبی که هر نمونه شامل ۷ عدد کاراکتر (۵ حرف انگلیسی و ۲ رقم) می‌باشد، برابر با $\frac{2}{70}$ می‌باشد که این عدد برای کپچای ۵ حرفی با همین تعداد نمونه برابر با $\frac{1}{10}$ می‌باشد که نشان می‌دهد کپچای ترکیبی از کپچای ۵ حرفی امنیت بالاتری دارد. در ۱۰ نمونه کپچای جبری، هیچ رقمی توسط نویسه‌خوان تشخیص داده نشده است، اما به معنی ایمنی کامل در برابر نویسه‌خوان نمی‌باشد و در تعداد نمونه های بالاتر ممکن است بعضی ارقام تشخیص داده شوند. به طور کلی، می‌توان نتیجه گرفت کپچای جبری امنیت بالاتر و همچنین کارایی بالاتری نسبت به بقیه کپچاها دارد و درک و پاسخ به آن برای کاربر انسانی ساده‌تر است. در ارتباط با کپچای ترکیبی و صحت‌سنجی ادعای مقاله، به این نتیجه رسیدیم که امنیت بالایی در مقایسه با کپچای ۵ حرفی دارد، اما تفاوت زیادی در کارایی با کپچای ۵ حرفی ندارد و در بعضی موارد، کپچای ۵ حرفی کارایی بالاتری را ارائه داده است.

1. Kato, N., Suzuki, M., Omachi, S., Aso, H., Nemoto, Y., A handwritten character recognition system using directional element feature and asymmetric Mahalanobis distance” IEEE Trans. on Pattern Analysis and Machine Intelligence, (1999)
2. Hsu, C., Ying-Lien, L., “Usability study of text-based CAPTCHAs”, Elsevier, vol.32, 81–86 (2011)
3. Brodic, D., Amelio, A.: The CAPTCHA: Perspectives and Challenges, Perspectives and Challenges in Artificial Intelligence, Vol.162, Springer, (2020)
4. Burxstein, E., Martin, M., Mitchell, J. C., Text-based CAPTCHA Strengths and Weaknesses, Stanford University (2014)
5. Smith, R., An Overview of the Tesseract OCR Engine, Google Inc.
6. Basu, S., DasGupta, A., The Mean, Median and Mode of Unimodal Distributions: A Characterization, Purdue University Department of Statistics (1992)
7. Whaley, D. L., The Interquartile Range: Theory and Estimation, East Tennessee State University, Electronic Theses and Dissertations (2005)
8. Kaur, K., Behal, S., Designing a Secure Text-Based CAPTCHA, 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015), Elsevier (2015)