بنام خداوند بخشنده مهربان

طرح احراز هویت با استفاده از كپچا (CPTCHA)

**چکیده**

احراز هویت فرآیندی است که طی آن درستی هویت یک فرد شناسایی و تأیید می گردد. بنابراین زمانی که کاربر بخواهد وارد سیستم شده و یا به منبعی دسترسی پیدا کند، ابتدا باید خود را اثبات کند. این شامل تأیید هویت یک شخص با اعتبار سنجی اثبات هویت کاربر یا تأیید صحت یک وب سایت با استفاده از گواهی های دیجیتال x.509 (یک استاندارد [ITU-T](https://fa.wikipedia.org/w/index.php?title=ITU-T&action=edit&redlink=1) است که به طور گسترده برای تعریف گواهی دیجیتال در [زیرساخت کلید عمومی](https://fa.wikipedia.org/wiki/%D8%B2%DB%8C%D8%B1%D8%B3%D8%A7%D8%AE%D8%AA_%DA%A9%D9%84%DB%8C%D8%AF_%D8%B9%D9%85%D9%88%D9%85%DB%8C) استفاده می‌شود) است. احراز هویت به چندین زمینه مربوط می شود. در هنر، تاریخ و مردم شناسی و ... . یک مشکل کلی این است که تأیید کنید مصنوع داده شده توسط یک شخص مشخص یا در یک مکان یا دوره مشخص از تاریخ تولید نشده است. در علوم کامپیوتر، تأیید هویت شخص اغلب برای دسترسی به داده ها یا سیستم های محرمانه مورد نیاز است. طرح احراز هویت اصلی استفاده از نامهای کاربری و رمزهای عبور عددی است. برای شکستن این رمزهای عبور ، بارها هكرها يا مهاجمان از برنامه های هوشمند استفاده می کنند. برای جلوگیری از این موضوع ، می توان از Captchas به عنوان یک اقدام امنیتی افزایشی برای بررسی اینکه آیا فرد تعامل پذیر انسان است یا یک ربات خودکار است، استفاده کرد. چندین روش وجود دارد که توسط نویسندگان مختلف پیشنهاد شده است. در این تحقیق، نویسندگان یک روش زیبا برای طراحی captcha پیشنهاد می کنند که از ورود ربات های هوشمند در هر سیستم اطلاعاتی جلوگیری می کند.

CAPTCHA به طور گسترده ای برای خنثی کردن هر عمل مكانيزه اي که برای اجراي شبيه سازي عملی که مختص انسان است و مكانيزه شده استفاده مي شود. آنها اصولاً برای خنثی کردن رباتهای هوشمند استفاده می شوند که حسابهای جعلی در وب سایتها ایجاد می کنند. CAPTCHA تابعي است که می تواند انسان متقابل باشد یا یک برنامه رایانه ای خودکار را ایجاد کرده و ارزيابي كند. انسان به راحتی می تواند از طریق این آزمون های ارزیابی، جایی که یک برنامه کامپیوتری خودکار توانايي عبور كردن از آن را ندارد عبور كند. از چنين برنامه هایی می توان برای تشخیص انسان از ربات های خودکار استفاده کرد و به طور گسترده ای در امنیت شبکه استفاده می شود.

شکل 1. نمونه ای از كپچاهاي ديداري

كپچاهاي صوتي از همان حروفي كه در كپچاهاي تصويري كاربرد دارند استفاده ميشود. این حروف براي اينكه خوانده ميشوند برای یک شخص دارای معلولیت جسمی بسیار مفید است.BotDetect CAPTCHA مدلي از كپچاي صوتي است که برای اطمینان از انسان بودن ورود كننده مورد استفاده قرار می گیرد و همچنین برای افراد نابینا و کم بینا كابرد زيادي دارد. این مدل از XHTML 1.1 و WCAG ، علامت گذاری سازگار با AAA استفاده می کند. به طور معمول، كپچاي صوتي BotDetect از عناصر Html5 Wav audio به طور کامل یا جزئی پشتیبانی نمیکند و به استفاده از عناصر <embed> و <object> برمی گردد. از انواع مرورگرهایی که كپچاهاي صوتي را پخش می کنند مانند پخش صوتی HTML5 انجام می دهند محافظت می کند. ساز و کار كپچاي صوتي BotDetect از 12 سبک صوتی مختلف استفاده می کند، در حالی که هر سبک از ترکیبی منحصر به فرد از جلوه ها و صدا استفاده می کند. این باعث می شود Captcha صوتی در هر تجزیه و تحلیل خودکار صوت بسیار قدرتمند باشد وقتی سبک صوتی به طور ناگهانی انتخاب می شود. مدل ساده كپچاي صوتی در شکل زیر نشان داده شده است

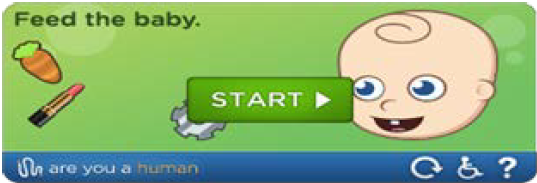
شكل 2: كپچاي مدل صوتي كه با گفتن حروف و تلفظ آن توسط كاربر كار مي كند

دسته بعدي كپچا، كپچاهاي تصويري است. نامگذاری CAPTCHA شش تصویر را به کاربر نشان مي دهد. هر تصویر با نام متناظر برچسب گذاری شده است. اگر کاربر نام صحیحی از تصویر را که قبلاً با تصویر برچسب گذاری شده است وارد کند، کاربر می تواند مرحله تاييد را بگذراند.

**بررسي انواع كپچا**

كپچاها به روش احتمالی غیر قابل شکستن و نفوذ تولید می شود که از هر الگوریتمی برای ایجاد اختلال در آنها جلوگیری می کند، اما باز هم کارای قابل توجهی برای نفوذ و شکستن آن انجام شده است. در سال 2003 ، موری و ماليک یک الگوریتم تطبیق شکل ایجاد کردند تا كپچاهاي Gimby و EZ-Gimpy را مختل کنند. آنها در مورد EZ-Gimpy 92% و در مورد Gimpy به 33٪ موفقیت دست يافتند. در سال 2004، موی و همکاران از روش تخمین تحريف برای شکستن كپچاهاي EZ-Gimpy استفاده کرد و همچنین به یک موفقيت بزرگ در زمينه كپچاهاي EZ-Gimpy دست پيدا كردند. اين کار نشان می دهد "تقسیم بندی" یک مسئله بسیار دشوارتر از "تشخیص" است زیرا الگوریتم های یادگیری ماشین می توانند به طور صحیح مسئله شناسایی را حل کنند، اما در حال حاضر هیچ الگوریتم موثری برای رمزگشایی مشکل تقسیم بندی ناشی از این بی نظمی های اضافه شده وجود ندارد. تکنیک های باز کردن و برچسب گذاری تصویر برای توسعه یک الگوریتم تقسیم بندی استفاده می شود. CAPTCHA های EZ-Gimpy و Gimpy را می توان به چهار نوع تقسیم کرد: زمینه های ساده (بدون مش) ، پس زمینه مش سیاه ، پس زمینه مش سفید و کاراکترهای متصل به هم. این روش کلمات را از یک فرهنگ لغت با 850 کلمه می گیرد، بنابراین می توان به راحتی این الگوریتم را شکست. یک الگوریتم همبستگی برای شناسایی کلمه دقیق EZ-Gimpy CAPTCHA در 99% مواقع طراحی شده است.

شكل 3: كپچاي مدل Gimpy

در مدل بدون مش كاراكترها در یک پس زمینه ساده نوشته شده اند در حالی که در مدل مش سياه، كاراكترها روی زمينه سياه و در مدل مش سفيد كاراكترها روي زمينه سفيد نوشته شده اند. بيشتر وب سرورها از كپچا بعنوان معیار امنیتی برای تشخیص کاربران انسانی از رباتها استفاده مي كنند. در متن CAPTCHA، کاراکترها عمداً تحریف می شوند و به خنثی کردن شناسایی توسط Bots متصل می شوند. امنیت CAPTCHA متن را می توان با افزودن علمی نویز ، تحريف و چیدمان بیشتر كاراكترها افزایش داد، نگرانی قابل توجهی در طراحی متن CAPTCHA وجود دارد. نمونه هایی از كپچاهای متنی شامل روش Gimpy، كپچاي دست نویس، روش متنBaffle ، روشPayPal، روشHotmail ، الگوهای تصویری پویا و چاپ دهدهی است. متن موفق CAPTCHA که توسط مایکروسافت، گوگل و یاهو استفاده شده است با استفاده از کمان های تصادفی، خطوط تصادفی متصل و نوشته هاي تو در تو با تکنیک هایی مقاوم در برابر حملات تقسیم بندی کار می كند. روش CAPTCHA بازی با یک پایگاه داده از مینی بازی های کارتونی کار می کند که برای کاربران نهایی با برخی از مشکلات دسترسی بسیار جالب و پشتیبانی می شود. این كپچاها به احتمال زیاد برای پایانه های دسک تاپ مناسب نیستند، اما بیشتر مناسب دستگاه های موبایل و صفحه لمسی هستند و در شکل 4 نشان داده شده است.

شكل 4: نمونه اي از Game CAPTCHA

كپچاهاي مبتنی بر تصویر به کاربران نهایی نیاز دارند که تصاویر دارای برچسب یا چرخش را تشخیص دهند. در كپچاي شناسایی تصویر، به کاربران نهایی یک مجموعه متناهی از تصاویر برای نامگذاری، تشخیص یا شناسایی ناهنجاری ها در آنها ارائه ميدهد. كپچاهاي ضمنی، کاربران نهایی نیازی به خواندن یا تایپ هر چیزی ندارند و با کلیک های ساده بر روی نقاط داغ کلیک می كنند. طراحي كپچا نقاط مختلفی را روی صفحه ای با پس زمینه داراي نويز ایجاد می كند. شكل زیر نمونه ای از نقاشی كپچا را نشان می دهد.

شكل 5: Drawing CAPTCHA

OCR با شناسایی خودکار عوامل مختلف در یک تصویر سند کار می کند که منجر به شناسایی، تجزیه و تحلیل و درک محتوای اسناد و مدارک می شود. سیستم OCR منطقه متن captcha را به خطوط متن، خطوط متن را به کلمات و سپس کلمات را به خصوصیات تقسیم می کند.

سپس این کاراکترها توسط کاربر شناسایی می شوند. علامت شناسایی را می توان به طور کلی به دو نوع تقسیم کرد: داده های چاپی ماشینی و داده های دست نویس. پیش پردازش، تقسیم بندی، استخراج ویژگی، بازشناسی و پردازش پس از آن مراحل اصلی در ضبط OCR هستند. میزان موفقیت یا شکست یک سیستم OCR به توزیع و عملکردهای استخراج ویژگی بستگی دارد.

CAPTCHA های مبتنی بر OCR بیشتر CAPTCHA های متنی هستند که در آنها تصاویر تحریف شده از حروف و یا رقم به کاربر نشان داده می شود. كاربر باید آنها را تشخیص داده و پاسخ را در قسمت متن داده شده تایپ كند. اما، این CAPTCHA ها یک اشکال داخلی دارند.

CAPTCHA های مبتنی بر OCR برای تلفن های همراه و دستگاه هایی مانند آن بسیار مشکل ساز هستند، زیرا ممکن است استفاده از صفحه کلید غیرقابل اجرا یا دشوار باشد. CAPTCHA براي کاربران واقعی دشوارتر شده و درعین حال مهاجمان در شکستن CAPTCHA های موجود نیز پیشرفت بهتری دارند.

**روش پیشنهادی**

روش بهبود یافته به گونه ای طراحی شده است که مقاومت روش CAPTCHA تصویر را در برابر حمله هکرها افزایش دهد. ساختار آن به شرح زیر است:

الف) این روش یک سری تصاویر سه بعدی را در سمت چپ و راست صفحه نمایش مي دهد.

ب) تصاویر بدون ترتیب و با مدلهای متفاوت ولي به صورت تصاویر هم نوع و هم موضوع نمایش داده می شوند.

ج) حال برنامه رایانه ای از کاربر می خواهد که تصویر سه بعدی را با تصویر مربوط به همان نوع انتخاب كند. اين تصاویر لازم نیست یکسان باشند اما شامل تصاویری از نوع مشابه هستند.

د) اگر کاربر به درستی تمام این تصاویر مشابه را انتخاب كند، مجاز است نام این تصاویر را به ترتیب متن نمایش داده شده در جعبه متن داده شده وارد کند.

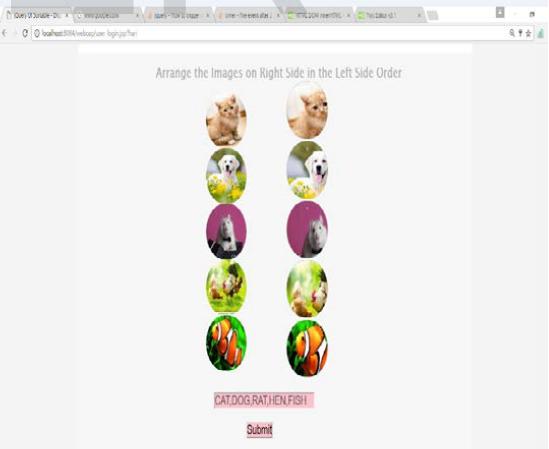
ه) اگر توالی نام و يا تصوير وارد شده صحیح باشد، می توان حدس زد که کاربر انسانی است و ربات نیست.

اكثرا جعبه متن يا كادر مربوطه غیر فعال می ماند تا زمانی که کاربر به درستی همه تصاوير مشابه با هم را انتخاب كند، سپس فقط جعبه متن فعال است و کاربر می تواند ترتیب صحیح نام تصاویر نقشه برداری شده را در جعبه متن وارد كند. در این روش، رایانه چهار توانایی لازم برای قبولی در آزمون را مجدداً لازم دارد.

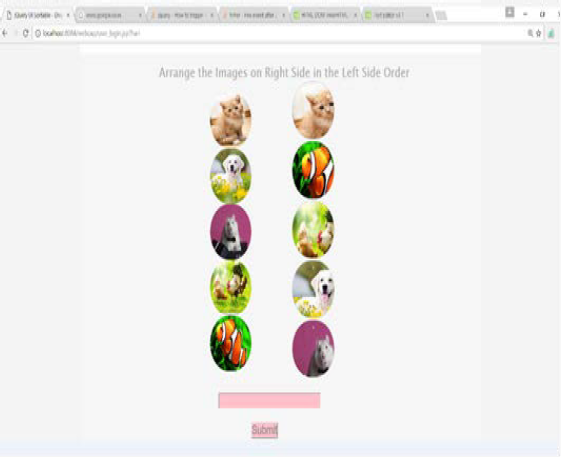
1. کاربر باید شکل تصویر داده شده را مشخص کند.
2. کاربر باید همه این تصاویر را با همان نوع تصاویر مربوطه در سمت راست مطابقت دهد.
3. سرانجام کاربر باید نام این تصاویر را پیدا کند و باید با استفاده از ویرگول به عنوان جدا کننده این نام ها، در جعبه متن داده شده تایپ کند.

برای وارد کردن نام تصویر توسط یک ربات هوشمند در جعبه متن، كاري دشوار دشوار است که این وظایف را به ترتیب صحیح انجام دهد، فقط یک کاربر انسانی می تواند شی مربوط را تشخیص داده و انتخاب كند. اين روش توسط جاوا یا هر زبان برنامه نویسی دیگری قابل پیاده سازی است. اجراي آن با برخی تفاوت ها از نوع CAPTCHA ساده و ساده است. برنامه CAPTCHA شش تصویر (شی) را به طور تصادفی انتخاب می کند که باید با تصاویر قبلی متفاوت باشد.

**نتیجه تجربی**

این روش مجموعه ای از تصاویر سه بعدی را در سمت چپ و راست صفحه نمایش می دهد. تصاویر به صورت تصادفي با بطور عاميانه درهم و برهم با مدلهای مختلفی از تصاویر یکسان نمایش داده مي شوند. حال برنامه رایانه ای از کاربر میخواهد که تصویر سه بعدی را با تصاوير مشابه يا هم نوع خود انتخاب كند. اين تصاویر لازم نیست یکسان باشند اما شامل تصاویری هم نوع و هم موضوع هستند.

شكل 6: CAPTCHA Verification

اگر کاربر به درستی با تمام این تصاویر مطابقت داشته باشد، کاربر مجاز است نام این تصاویر را به ترتيب متن نمایش داده شده در جعبه متن داده شده وارد كند. اگر توالی نام وارد شده صحیح باشد، می توان حدس زد که کاربر انسانی است و ربات نیست.

شكل 4: CAPTCHA Analysis

**نتيجه گيري**

برای دسترسی به داده ها یا سیستم های محرمانه ، در شبکه های توزیع شده ، تأیید هویت شخص از ده مورد لازم است. طرح احراز هویت اصلی استفاده از نامهای کاربری و رمزهای عبور عددی است. برای شکستن این رمزهای عبور ، بارها مهاجمان از برنامه های هوشمند استفاده می کنند. برای جلوگیری از این موضوع ، می توان از CAPTCHA به عنوان یک اقدام امنیتی اضافی برای بررسی اینکه انسان تعامل برقرار است، یک انسان است یا یک ربات خودکار است. چندین روش وجود دارد که توسط نویسندگان مختلف پیشنهاد شده است. در این تحقیق، نویسندگان روشی زیبا برای طراحی CAPTCHA پیشنهاد داده اند که از ورود ربات های هوشمند در هر سیستم اطلاعاتی جلوگیری می کند. کپچای پیشنهادی امنیت بیشتری نسبت به سایر روش ها ایجاد می کند. برای وارد كردن نام تصوير توسط یك ربات هوشمند در جعبه يا كادر متن، تحقق این كارها به ترتیب صحیح برای كامپیوتر دشوار است، فقط یك کاربر انسانی می تواند شی مورد نظر را تشخیص و انتخاب کند. اين روش را می توان توسط جاوا یا هر زبان برنامه نویسی دیگری پیاده سازی کرد. اجرا با برخی تفاوت ها مشابه روش اصلی CAPTCHA است.