بررسی موردی میزان امنیت کپچا در سایت های ایرانی

سعیدهاشم زاده[۱]،حنانه حسین نیا[۲]، محمد حسن شیرعلی شهرضا[۳]

[۱]دانشجوی کارشناسی ارشد، رشته مهندسی کامپیوتر، دانشگاه صنعتی امیرکبیر (پلی تکنیک تهران)

[۲]دانشجوی کارشناسی ارشد،رشته مهندسی کامپیوتر، دانشگاه صنعتی امیرکبیر (پلی تکنیک تهران)

[۳]دانشکده ریاضی و علوم کامپیوتر، دانشگاه صنعتی امیرکبیر (پلی تکنیک تهران)

Archive of SID

چکیده

درعصر فعلی وب سایت ها کاملا پویا بوده، و مدام در حال محاوره با کاربران خود باشند، لذا برای تحقق این امر نیاز به صفحاتی میباشد که داده ها را از کاربران دریافت نمایند. صفحات ورود به سایت و یا ثبت نام در یک سایت که معمولا متداول است، نمونه ای از این صفحات میباشند . یک فرستنده ی هرزنامه می تواند یک برنامه ساده طراحی کرده و به طور خود کار شروع به پر کردن فیلدها این گونه صفحات کند. برای مثال،یک فرم ثبت نام در سایت را در نظر بگیرید، یک برنامه ساده که توسط یک فرستنده هرزنامه نوشته شده است می تواند آن وب سایت را باسیل عظیمی از حساب های کاربری روبرو کند. راه حلی که برای جلوگیری از این گونه مشکلات ارائه شده است استفاده از کپچا میباشد. در واقع کپچا برای متمایز ساختن کاربر انسانی از ماشین به کار می رود. در کپچا سعی می‌شود که مساله‌ای مطرح شود که حل آن برای کاربر انسانی ساده بوده ولی حل آن برای یک برنامه کامپیوتری مشکل باشد. در این مقاله کپچای مورد استفاده در سایت سازمان سنجش آموزش کشور بررسی شده و مشخص شده که کپچاهای استفاده شده به راحتی توسط برنامه های کامپیوتری حل می‌شوند و در واقع این گونه کپچاها، روش مطمئنی برای تفکیک کاربران انسانی از برنامه‌های کامپیوتری نیستند.

کلمات کلیدی: **کپچا، پردازش تصویر،هرزنامه، شبکه عصبی، امنیت وب .**

۱-مقدمه

پویایی و محاوره در سایت های امروزی امری اجتناب ناپذیر است . سایت های بسیاری وجود دارند که در آنها نیاز به صفحاتی است که کاربر بتواند داده های مورد نیاز را وارد نماید. برای مثال سرویس های پست الکترونیکی رایگان مانند یاهو و جی میل که به کاربر اجازه می دهند، حساب کاربری رایگان داشته باشند را درنظر بگیرد، این گونه سایت ها به راحتی می توانند با سیل عظیمی از حساب های کاربری و یا هرز نامه هایی که توسط هرزنامه نویسان و نرم افزارهای ایجاد کننده ی هرزنامه تولید شده است، روبرو شوند. همچنین سایت هایی وجود دارند که در قسمت تماس با ما، کاربران می توانند پیام خود را ثبت کنند. یا شبکه های اجتماعی را در نظر بگیرید که کاربران می توانند پیام دلخواه خود را برروی سایت بفرستند. با وجود نرم افزارهای تولید هرزنامه نیاز به مکانیزمی است که بتواند کاربر انسانی را از ماشین تشخیص دهد . مکانیزمی که برای این موضوع مورد استفاده قرار گرفته شده

www.SID.ir

است استفاده از کپچا میباشد. در کپچا سعی می‌شود که مساله‌ای مطرح شود که حل آن برای کاربر انسانی ساده بوده ولی حل آن برای یک برنامه کامپیوتری مشکل باشد.

شایع ترین نوع کپچا استفاده از تصاویرحروف و اعداد می باشد که عمدتا چرخانده شده‌اند تا به راحتی توسط برنامه‌های کامپیوتری قابل تشخیص نباشند.. به طور کلی می توان انواع کپچا را به سه دسته کلی تقسیم کرد[2]:

الف) استفاده از تصاویری که شامل حروف و ارقام میباشد.

ب) استفاده ازتصاویری که شامل تصویر اشیاء و حیوانات و غیره میباشد و کاربر میبایست متن مربوط به آن تصویر را وارد نماید.

ج) استفاده از فایل های صوتی که کاربر میبایست متن بیان شده در آن را تایپ نماید.

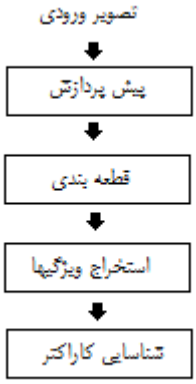
قابل ذکر است که این دسته بندی کاملا کلی بوده و هر کدام از این دسته بندی ها می توانند پیچیده تر شوند. مثلا در مورد دسته اول می تواند متن حاوی یک سوال باشد و کاربر می بایست جواب سوال را تایپ نماید.[2] گاهی از کپچا به عنوان معکوس تست تورینگ یاد میشود [1]. هدف از تست تورینگ تشخیص ماشین هوشمند از انسان است ولی هدف کپچا تشخیص کاربرانسانی میباشد.

تمرکز این مقاله بر روی دسته اول کپچا از دسته بندی بالا میباشد، سعی شده با ارائه روشی برای حل کپچا برپایه تکنیک های پردازش تصویر، اثبات شود که بسیاری از کپچاهای مورد استفاده در وب سایت های امروزی امنیت لازم را برای مقابله با ربات ها برآورده نمی کند.

در بخش ۲ مراحل انجام کار بررسی شده است. در بخش سوم نتایج بدست آمده آورده شده و در بخش چهارم ابتدا به شرح نتیجه گیری و سپس کارهای آتی پرداخته شده است.

۲-مراحل انجام کارو پیاده سازی الگوریتم

برای بررسی کپچاهای مورد استفاده در سایت‌های ایرانی سایت سازمان سنجش درنظر گرفته شد. تمامی مراحلی که در ادامه توضیح داده خواهد شد توسط زبان برنامه سازی c# در محیط visual studio 2015 پیاده سازی شده است. شکل ۱ مراحل انجام کار را به ترتیب نشان می دهد که در ادامه تمام مراحل به همراه نحوه پیاده سازی آنها توضیح داده خواهد شد.



شکل۱– مراحل شکستن کپچا

2-1-پیش پردازش

ابتدا باید پردازش های اولیه روی تصویر انجام شود تا تصویر برای قطعه بندی و مراحل بعدی آماده شود. پیش پردازش شامل سه مرحله است:

۱- تبدیل تصویر رنگی به تصویر سیاه سفید(Gray scale)
۲- باینری سازی
۳- حذف نویز

۲-۱-۱ - تبدیل تصویر رنگی به تصویر سیاه سفید و باینری سازی

شکل۲ – نمونه ای از کپچای رنگی

شکل ۲ یک نمونه از کپچای رنگی میباشد. همانطور که در شکل ۲ مشخص است تعداد رنگ های تصویر زیاد است. گام اول پیش پردازش ، تبدیل تصویر رنگی به یک تصویر سیاه سفید میباشد، منظور از تصویر سیاه سفید، تصویری است که فقط شامل ۲۵۶ رنگ خاکستری باشد، بدیهی است پردازش های مربوط به مراحل بعدی

Archive of SID
www.SID.ir

روی تصویری که تعداد رنگ کمتری دارد راحت تر از تصاویر رنگی میباشد.

در فایل های bmp رنگ هر پیکسل معمولا به صورت ۳ عدد ۸ بیتی، یعنی هر عدد بین ۰ و ۲۵۵ میباشد. برای تبدیل یک فایل bmp رنگی به خاکستری، میانگین این سه عدد را گرفته و برای هر پیکسل یک عدد بین ۰ تا ۲۵۵ در نظر می گیریم. بدین صورت رنگ هر پیکسل یک عدد بین ۰ تا ۲۵۵ میشود یعنی حداکثر ۲۵۶ رنگ خاکستری باقی می ماند .

بعد از این که تصویر سیاه سفید شد ، می توان آن را به یک تصویر باینری تبدیل کرد. تصویر باینری تصویری است که پیکسل های آن فقط شامل دو رنگ میباشد. یعنی هر پیکسل یا سفید یاسیاه است.

باینری سازی تصویر باعث میشود کاراکتر ها از پس زمینه جدا شوند. برای تبدیل تصور سیاه سفید به تصویر باینری کافی است یک حد آستانه مانند x را درنظر گرفته و سپس پیکسل های بیش از x را سفید و پیکسل های کمتر از x را سیاه می شود.

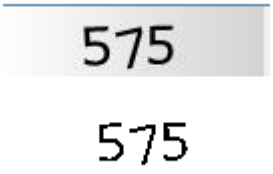
شکل ۳– نمونه ای از کپچای سازمان سنجش

شکل ۳ یک نمونه از کپچای سازمان سنجش را نشان می دهد ، همانطور که شکل۳ نشان می دهد تصویر به صورت سیاه سفید است و رنگ پس زمینه با رنگ متن از لحاظ روشنایی متفاوت است. این نوع کپچا سیاه سفید است اما چون به صورت bmp ذخیره شده است رنگ هر پیکسل ممکن از بازه ۰ تا ۲۵۵ که برای رنگ هر پیکسل تصویر سیاه سفید در نظر گرفتیم بیشتر باشد. می توان از خواصی که بیان شد برای کپچای انتخابی ، استفاده کرد و با یک روش ابتکاری مرحله تبدیل تصویر رنگی به تصویر سیاه سفید و باینری سازی تصویر را به صورت همزمان انجام داد.

روشی که برای ایجاد تصویر سیاه سفید و باینری سازی تصویر در این مقاله استفاده شده است به این صورت است که یک محدود ۲۰ پیکسلی از پس زمینه را انتخاب کرده و میانگین آن را به دست می آوریم. عدد بدست آماده را به عنوان حد آستانه در نظرمی گیریم و برای نشان دادن آن در این مقاله از x_t استفاده می کنیم. نکته ای که باید به آن اشاره کرد این است که x_t فقط یک بار برای یک تصویر محاسبه شده و از آن برای تمامی تصویرها استفاده میشود. الگوریتم زیر نحوه باینری سازی کپچای سازمان سنجش را نشان می دهد.

–به ازای هر پیکسل با رنگ c :

اگر $c>x_t$ بود رنگ آن پیکسل را سیاه و در غیر این صورت رنگ پیکسل را سفید در نظر بگیر.



شکل ۴– باینری سازی یک کپچای نمونه

شکل ۴ یک نمونه از تصویر باینری شده کپچا ی سایت سازمان سنجش را نشان می دهد.

۲-۱-۲- حذف نویز تصویر

نویز ها در کپچا ها می توانند به صورت نقطه های پراکنده یا به صورت خط های عمودی و افقی در تصویر باشند. این کار باعث میشود که عملیات قطعه بندی و شناسایی کاراکتر مشکل شود. اما همانطور که شکل ۳ نشان می دهد کپچای سایت سازمان سنجش دارای نویز هایی از این قبیل نمیباشد.

۲-۲ قطعه بندی

قطعه بندی یعنی این که کاراکتر ها ی داخل تصویر از هم جدا شده و هر کدام به صورت یک تصویر مجزا پردازش شود. برای این کار می توان از روش isolated connected component استفاده کرد[11] .

روش استفاده شده در این مقاله برای قطعه بندی به صورت زیر می اشد:

همانطور که شکل ۳ نشان می دهد و با بررسی نمونه های دیگر از کپچا های سایتهای انتخابی می توان به این نتیجه رسید که کاراکترهای موجود در کپچا به صورت جدا از هم هستند و فقط کافی است برای تفکیک کاراکتر ها و تبدیل هر کاراکتر به یک تصویر مجزا پیکسل های سیاه بهم پیوسته را شناسایی کرده و آن را به صورت یک تصویر درنظر بگیریم.

روشی که در این مقاله برای شناسایی پیکسل های بهم پیوسته استفاده میشود به صورت زیر میباشد:

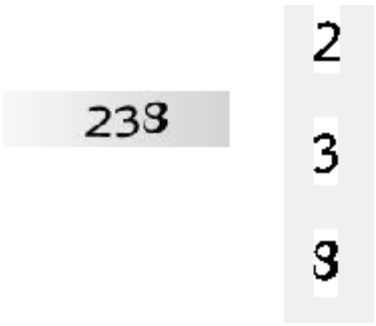
۱- به ازای هر ستون y_s از پیکسل ها :

اگر ستون y_s شامل پیکسل سیاه نبود مرحله ۱ را تکرار کن. در غیر این صورت ستون y_s را به عنوان اولین شروع کاراکتر در نظر بگیر.

۲- به ازای هر ستون y_e از y_s به بعد:

اگر ستون y_e شامل پیکسل سیاه بود مرحله ۲ را تکرار کن. در غیر این صورت y_e را ستون انتهای کاراکتر در نظر بگیر.

Archive of SID
www.SID.ir



شکل۵-نحوه جداسازی کاراکترها

شکل ۵ یک نمونه از قطعه بندی کپچای سایت سنجش را تو سط الگوریتم ارائه شده در این مقاله را نشان می دهد.

۲-۳-استخراج ویژگی تصویر

هر کاراکتر یک مجموعه ویژگی منحصر به فرد دارد، که آن کاراکتر را از دیگر کاراکترها متمایز می کند. از این ویژگی ها در مرحله شناسایی کاراکترها استفاده می شود.

ویژگی های مورد استفاده در این مقاله به شرح زیر میباشد[3][4]:

۱. حداکثر و حداقل تعداد پیکسل های سیاه که برای نمایش یک کاراکتر لازم است.
به عنوان مثال برای نمایش عدد ۱ طبق بررسی های انجام گرفته حدود ۵۶ پیکسل نیاز است.

۲. حداقل و حداکثر تعداد ستون لازم برای نمایش یک کاراکتر خاص. به عنوان مثال برای نمایش کاراکتر ۱ حداقل ۵ ستون و حد اکثر ۹ ستون لازم است.

۳. ایجاد بردار عمودی از تعداد پیکسل های سیاه، این بردار به شکل زیر تعریف میشود:

$$(x_0, x_1, x_2 \dots)$$

که در آن x_i معرف تعداد پیکسل های سیاه در ستون iام میباشد.

۴. ایجاد بردار افقی از تعداد پیکسل های سیاه. این بردار به شکل زیر تعریف میشود

$$(y_0, y_1, y_2 \dots)$$

که در آن y_i معرف تعداد پیکسل های سیاه از سطر iام میباشد.در ادامه از این سه ویژگی بالا برای شناسایی تصاویر استفاده میشود.

دو ویژگی اول می تواند تعیین کند که تصویر ورودی کدام کاراکتر نمیباشدو فضای جستجو را محدودتر کند تا از این طریق به سرعت بیشتر برسیم.

۲-٤ شناسایی کاراکتر ها:

شناسایی کاراکترها براساس ویژیگی های استخراج شده از تصویر انجام میشود. دو ویژگی اول بیان شده در قسمت قبل برای هر کاراکتر ثابت میباشد. ولی دو ویژگی دوم به دلیل دوران کاراکترها می تواند تغییر کند. معمولا برای دسته بندی کاراکترها از ماشین یادگیری و شبکه عصبی استفاده میشود.

در این مقاله به جای استفاده از ماشین یادگیری از روش دیگر استفاده شده است. که این روش به شرح زیر است:

همانطور که بیان شد دو ویژگی اول برای هر کاراکتر ثابت است ولی دو ویژگی دوم برای کاراکتر های مختلف متفاوت میباشد. بنابراین نیاز به دو ساختار مجزا داریم ، یک ساختار برای ذخیره سازی ویژگی های ثابت یک کاراکتر و یک ساختار برای ذخیره سازی ویژگی های متغیر کاراکتر، که در ادامه هر دو ساختار را تشریح می کنیم.

۲-٤-۱ ذخیره سازی ویژگی های ثابت کاراکتر:

ویژگی های ثابت کاراکتر مشخص می کند که تصویر ورودی کدام کاراکتر نمیباشد. بنابراین ذخیره و دقت در ثبت این نوع ویژگی در دقت شناسایی کاراکتر بسیار اهمیت دارد.

در این مقاله برای پیدا کردن حداقل و حداکثر تعداد پیکسل های لازم برای نمایش یک کاراکتر از برنامه ای که مقدار حداقل و حداکثر تعداد پیکسل های لازم را در جدولی در پایگاه داده ذخیره می کند استفاده شده است. برای حداقل و حداکثر ستون لازم برای برای نمایش یک کاراکتر از روشی مشابه استفاده شده است.

۲-٤-۱ ذخیره سازی ویژگی های متغیریک کاراکتر:

برای ذخیره سازی ویژگی های متغییر هر کاراکتر در این مقاله مانند ذخیره سازی مشابه ذخیره سازی ویژگی های ثابت برنامه ای نوشته شده است که برای هر کاراکتر مجموعه از تصاویر را دریافت کرده و سپس بردار های عمودی و افقی هر کاراکتر را ایجاد کرده و در پایگاه داده ذخیره می کند.

Archive of SID
www.SID.ir

۲-۵-کلاس بندی

برای دسته بندی و شناسایی تصاویر از شبکه عصبی استفاده شده است. شبکه های عصبی مصنوعی روشی عملی برای یادگیری توابع گوناگون نظیرتوابع با مقادیر حقیقی، توابع با مقادیر گسسته و توابع با مقادیر برداری میباشد.این گونه شبکه ها با موفقیت به مسائلی نظیر شناسایی گفتار و تصاویر، اعمال شده است. شبکه عصبی ای که در این مسأله به کارگرفته شده، شبکه پرسپترون است. در واقع نوعی از شبکه عصبی برمبنای یک واحد محاسباتی به نام پرسپترون ساخته میشود. یک پرسپترون برداری از ورودیها با مقادیر حقیقی را گرفته و یک ترکیب خطی از این ورودی ها را محاسبه می کند.[5]

اگر حاصل از یک مقدار آستانه بیشتر بود خروجی پرسپترون برابر با ۱ و در غیر اینصورت معادل ۱- خواهد بود.

قبل از توضیح نحوه پیاده سازی شبکه عصبی لازم به ذکر است که برای شناسایی تصویر قبل از این که داده ها را برای شبکه عصبی مهیا کنیم ، با استفاده از ویژگی های ثابت مشخص می کنیم که تصویر ورودی کدام کاراکتر نمیباشد. سپس کاراکترهای باقی مانده به همراه ویژگی های متغیرشان به ورودی شبکه عصبی ارسال میشود.

۲-۵-۱- نحوه پیاده سازی یک سلول شبکه عصبی پرسپترون

شکل ۶ – شکل کلی یک سلول پرسپترون

همانطور که درشکل ۶ مشاهده می کنید، یک سلول شبکه عصبی پرسپترون در این مسأله ، شامل سه ورودی و یک خروجی است. شرح ورودی ها به صورت زیر میباشد:

A ویژگی های تصویری است که باید به متن تبدیل شود. (تصویر اصلی که یکی از کاراکترهای کپچا میباشد)

B ویژگی های کاراکتری است که از پایگاه داده دریافت شده است.

C ویژگی های کاراکتری است که از پایگاه داده دریافت شده است.

ویژگی ها ی تصویر A هم با هم با ویژگی های کاراکتر B مقایسه میشودو هم با ویژگی های کاراکتر C، از بین کاراکتر های B و C کاراکتری انتخاب میشود که بیشترین شباهت را با تصویرA داشته باشد.

۲-۵-۲-نحوه مقایسه تصاویر

مقایسه تصاویر براساس ویژگی های استخراج شده از تصاویر انجام می گردد. برای مقایسه تصاویر کافی است بردار های عمودی و افقی باهم مقایسه شوند.

۲-۵-۳- نحوه مقایسه دو بردار

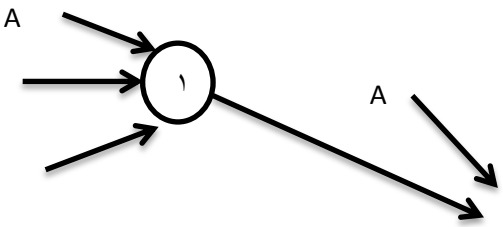
برای مقایسه دو بردار از روش Chi-Square استفاده شده است. [5] روش Chi-Square به صورت زیر تعریف میشود.

اگر ما دو بردار $A=(a_1,a_2,a_3,...a_n)$ و $B=(b_1,b_2,b_3,...b_n)$ داشته باشیم میزان فاصله دوبردار به روش Chi-Square از فرمول زیر به دست می آید.

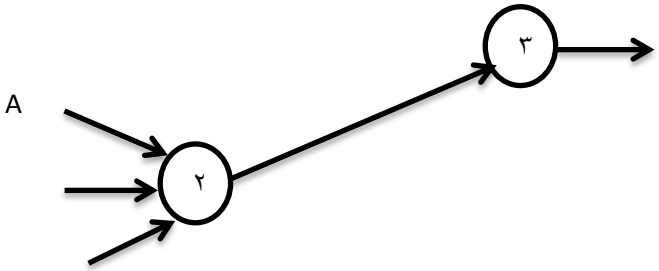
$$\sum \frac{(ai-bi)^2}{ai} \quad (۱)$$

۲-۵-٤-نحوه اتصال سلول های شبکه عصبی

برای ایجاد یک شبکه عصبی پرسپترون باید سلول های پرسپترون به یکدیگر متصل شوند . برای اتصال این سلولها کافی است، در مرحله اول ورودی اول همه سلول ها A باشد، و دو ورودی دیگر از پایگاه داده دریافت شود. درمرحله ی بعد بازهم ورودی اول تصویر A باشد، دو ورودی دیگر هر سلول هم باید از خروجی سلول های مرحله ی قبلی انتخاب شوند.



Archive of SID
www.SID.ir



شکل ۷ – شبکه عصبی استفاده شده

در شکل ۷ هر دایره نشان دهنده یک سلول شبکه ی عصبی میباشد، همان طورکه شکل ۶ نشان می دهد سلول های ۱ و ۲ سلول های سطح اول این شبکه ی عصبی میباشند، که در مرحله اول ورودی سلول ها از ماشین یادگیری دریافت میشود و در مرحله بعد خروجی سلول های مرحله قبل به عنوان وروی سلول های مرحله بعدی استفاده میشوند و این عمل تازمانی ادامه می یابد که به یک خروجی واحد برسیم. که این خروجی واحد شبه ترین تصویر به A میباشد. که ماشین یادگیری می تواند تعیین کند که تصویر خروجی شبکه عصبی برابر چه متنی مباشد.

برای بدست آوردن دقت بیشتر می توان شبکه عصبی را طوری اصلاح کرد که برای میزان شباهت بیان شده در بخش ۲-۵-۱ یک حد آستانه در نظر گرفته و هر تصویرکه میزان شباهتش با تصویر اصلی (A) از حد آستانه کمتر بود به عنوان خروجی شبکه عصبی انتخاب شود در این صورت شبکه عصبی بیش از چند خروجی خواهد داشت که برای انتخاب خروجی مناسب می توان تبدیل فوریه[3] [4] تصاویر را محاسبه کرده و با تصویر اصلی مقایسه کرد هر تصویری بیشترین شباهت را با تصویر اصلی داشته باشد، به عنوان خروجی شبکه عصبی انتخاب خواهد شد.

۳-نتایج

برای بررسی کپچاهای مورد استفاده در سایت های ایرانی ۵۰ کپچا برای یادگیری سیستم و ۱۰۰ کپچا برای آزمایش سیستم از سایت سازمان سنجش انتخاب شد. با توجه به آزمایش انجام گرفته برروی تصاویر کپچای سازمان سنجش با روش بیان شده در این مقاله، نتایج در جدول شماره یک نشان داده شده است.

جدول شماره یک – نتایج آزمایش شناسایی کپچا

تعداد کپچاهای استفاده شده برای یادگیری سیستم	۵۰ عدد (برای هر رقم ۵ نمونه)
تعداد کپچاهای استفاده شده برای آزمایش	۱۰۰ عدد
تعداد کپچاهایی که کاملا درست تشخیص داده شده اند	در ۸۳ کپچا هر سه رقم کاملا درست تشخیص داده شده است.
تعداد تشخیص نادرست	۱۷ عدد کپچا
تعداد کپچا هایی که فقط دو رقم آن درست تشخیص داده شده است	۱۰ عدد
تعداد کپچا هایی که فقط یک رقم آن درست تشخیص داده شده است	۵عدد
تعداد کپچا هایی که هیچ یک از ارقام آن درست تشخیص داده نشده است	۲ عدد

باتوجه به جدول۱ می توان نتیجه گرفت که روش بیان شده در این مقاله توایی تشخیص ۹۱.۳۳ درصد از کاراکترها را دارد.

با در نظر گرفتن این موضوع که در صفحات وب تصاویر کپچا را می توان refresh کرد ، این درصد تطابق می تواند اثبات کند که استفاده از کپچاهای ساده همچون کپچای سایت سازمان سنجش روش مناسبی برای تفکیک کاربران انسانی از کامپیوتر نیست.

٤-نتیجه گیری

هدف مقاله بررسی کارایی کپچاهای استفاده شده در سایت های ایرانی بود. برای این منظور کپچای مورد استفاده در سایت سازمان سنجش آموزش کشور که یکی از پربیننده ترین سایت های کشور می‌باشد، مورد بررسی قرار گرفت. بدست آمده نشان داد که کپچاهای استفاده شده توانایی لازم برای مقابله با برنامه های کامپیوتری را نداند.

راه حل این مشکل می‌تواند استفاده از کپچاهای پیچیده تر باشد ولی آن کپچاها نیز با روش های پیشرفته تر شکسته می‌شوند. بنابر این وجود این نوع کپچاها در سایت ها فقط باعث به زحمت افتادن کاربران بوده و در مقابل برنامه های کامپیوتری شکست می‌خورند.

Archive of SID
www.SID.ir

مراجع

[1] Baird, Henry S., and Mark Luk, "Protecting Websites with Reading-Based CAPTCHAs," *Second International Web Document Analysis Workshop,* 2003.

[2] Antialiasis, No Spam! - an alternative to CAPTCHA images www.vbulletin.org/forum/showthread.php?t=124828 آخرین مشاهده: ۱۶ اسفند ۱۳۹۴.

[3] Gonzalez, Rafael C., and Richard E. Woods, Digital image processing, 3rd edition, 2008.

[4] David, Vasantha Kalyani, and Sundaramoorthy Rajasekaran. *Pattern recognition using neural and functional networks*. Springer, 2009.

[5] Sokolove, Phillip G., and Wayne N. Bushell, "The chi square periodogram: its utility for analysis of circadian rhythms," *Journal of Theoretical Biology,* 72.1, 1978, pp:: 131-160.

[6] Rojas, Raúl, *Neural networks: a systematic introduction,* Springer Science & Business Media, 2013.

[7] Mori, Greg, and Jitendra Malik, "Breaking a visual CAPTCHA," *Unpublished Manuscript,* 2002.

[8] Von Ahn, Luis, Manuel Blum, and John Langford, "Telling humans and computers apart automatically," *Communications of the ACM,* 47, 2, 2004, PP: 56-60.

[9] Karthik, C. H. B. L. P., and Rajendran Adria Recasens, "Breaking Microsoft's CAPTCHA," 2015.

[10] Mitianoudis, Nikolaos, and Nikolaos Papamarkos, "Document image binarization using local features and Gaussian mixture modeling," *Image and Vision Computing,* 38 (2015), pp: 33-51.

[11] Cruz, Francisco, and Oriol Ramos Terrades. "Handwritten line detection via an EM algorithm," *12th International Conference on Document Analysis and Recognition (ICDAR), 2013*.

Archive
www.SID.ir