





دانشگاه علم و صنعت ایران

دانشکده مهندسی کامپیوتر

016702

۱۳۸۰ / ۱۲ / ۲۸

تصدیق هویت بر اساس نحوه کلیدزنی به روش آماری - کوواریانسی

پایان نامه برای دریافت درجه کارشناسی ارشد
مهندسی کامپیوتر - هوش ماشین و رباتیک

علی اصغر کاجی

استاد راهنما: دکتر مرتضی آنالوئی

بهار ۱۳۸۰

تقدیم بہ

پدر و مادر عزیزہ

کہ کلیۃً موفقیت ہایم را مدیون ایشان ہستم

چکیده

استفاده از خصوصیات بیولوژیکی و رفتاری افراد جهت کنترل دسترسی به منابع و حفظ امنیت سیستمها به شیوه‌های مختلف، همواره مورد توجه بوده است. روشهای مبتنی بر تشخیص گفتار، تشخیص چهره، تشخیص اثرانگشت و شناسایی از روی شبکیه چشم از جمله معروفترین این روشها هستند. روشی که در این تحقیق مورد توجه قرار گرفته است، استفاده از ویژگیهای کلیدزنی افراد است. در این تحقیق ویژگیهای اولیه زمان فشار کلیدها و مدت زمان بین کلیدها می‌باشد. با اخذ این ویژگیها از رشته‌های ثابت بعنوان داده‌های اولیه پس از انجام یک مرحله پیش‌پردازش اولیه با استفاده از یک کلاس‌بند ترکیبی کوواریانسسی که در اینجا پیشنهاد می‌شود در کلاس‌بندی کاربران بکار گرفته می‌شود. این روش با حفظ سادگی روشهای آماری و سهولت پیاده‌سازی نسبت به روشهای شبکه‌عصبی، خطای کمتری نسبت به سایر روشهای آماری نشان می‌دهد. همچنین در این روش می‌توان تغییراتی را که بطور مداوم و به کندی در جهت تثبیت الگوی کلیدزنی کاربران بوجود می‌آید، دنبال کرد. همچنین با توجه به مسئله تغییر صفحه‌کلید و نقش آن در تغییر الگوی کاربران روشی برای اعمال اثر آن در الگوی کلیدزنی کاربران پیشنهاد شده است.

کلمات کلیدی: کلیدزنی، تصدیق هویت، کلاس‌بند آماری، الگوی کلیدزنی.

سپاس خدایی را که سخنوران در ستودن او بمانند و شمارگران شمردن نعمت های او ندانند و
کوشندگان حق او را گزاردن نتوانند. خدایی که پای اندیشه تیزگام در راه شناسایی او لنگ است،
و سر فکرت ژرف رو به دریای معرفتش بر سنگ. صفت های او تعریف ناشدنی است و به وصف
درنیامدنی، و در وقت ناگنجیدنی، و به زمانی مخصوص نابودنی....

(خطبه اول نهج البلاغه)

از جناب آقای دکتر مرتضی آنالوئی استاد ارجمند که همواره از راهنمایی های ایشان در انجام این
پروژه بهره مند بوده ام سپاسگزارم،
همچنین لازم است از دوستان عزیز آقایان مهندس افشین ریاضی و مهندس بابک ناصرشریف که
از نظراتشان بهره مند شدم و دانشجویان دوره کارشناسی ارشد کامپیوتر در سایت کارشناسی ارشد و
آزمایشگاه شبکه های کامپیوتری که با کمال صبر اینجانب را در تهیه داده های مورد نیاز یاری
نمودند، کمال تشکر و تقدیر را بنمایم.

مقدمه	۱
فصل اول . روشهای کنترل دسترسی	۷
۱-۱- مقدمه	۸
۱-۲- شناسایی کاربر	۹
۱-۳- تصدیق هویت کاربر	۹
۱-۴- تصدیق هویت تایید شده	۱۰
۱-۵- تصدیق هویت دوطرفه	۱۱
۱-۶- تصدیق هویت دوباره	۱۲
۱-۷- نمایش سطوح کنترل دسترسی با استفاده از شبکه‌های پتری	۱۴
۱-۷-۱- سطح اول و دوم	۱۴
۱-۷-۲- سطح سوم کنترل دسترسی	۱۴
۱-۷-۳- سطح چهارم	۱۶
۱-۷-۴- سطح پنجم کنترل دسترسی	۱۷
فصل دوم . تشخیص هویت بر اساس نحوه کلیدزنی	۲۰
۲-۱- مقدمه	۲۱
۲-۲- روش مدل آماری	۲۲
۲-۲-۱- دسته بندی اولیه	۲۳
۲-۲-۲- مدل آماری	۲۳
۲-۳- روش های احتمالات وزنی	۲۴

۲۶ روش فاصله اقلیدسی	۲-۳-۱
۲۷ روش احتمال بدون وزن	۲-۳-۲
۲۸ روش احتمال وزن دار	۲-۳-۳
۲۹ نتیجه‌گیری	۲-۴
۳۰ فصل سوم . روشهای تصدیق هویت بر اساس نحوه کلیدزنی	
۳۱ مقدمه	۳-۱
۳۲ روش فازی	۳-۲
۳۳ ارتباط بین ورودی و خروجی در سیستم فازی	۳-۲-۱
۳۶ استنتاج فازی	۳-۲-۲
۳۷ روشهای آماری تصدیق هویت	۳-۳
۳۸ کلاس بند غیر خطی	۳-۳-۱
۳۸ الگوریتم <i>K-means</i>	۳-۳-۲
۳۹ روش فاصله کسینوسی	۳-۳-۳
۴۰ روش کلاس بند بیز	۳-۳-۴
۴۰ کلاس بند یادگیرنده استنتاجی	۳-۳-۵
۴۲ روشهای شبکه عصبی	۳-۴
۴۲ شبکه <i>Backpropagation</i>	۳-۴-۱
۴۳ شبکه عصبی <i>Counterpropagation</i>	۳-۴-۲
۴۴ شبکه عصبی <i>Fuzzy ARTMAP</i>	۳-۴-۳
۴۴ شبکه عصبی توابع پایه شعاعی	۳-۴-۴
۴۵ شبکه عصبی با یادگیری کوانتیزاسیون برداری	۳-۴-۵

۴۵ شبکه عصبی Reinforcement ۳-۴-۶
۴۶ شبکه مجموع حاصلضرب ۳-۴-۷
۴۶ شبکه ترکیبی مجموع حاصلضرب ۳-۴-۸
۴۷ نتیجه گیری ۳-۵
۴۸ فصل چهارم . ارائه مدل تصدیق هویت
۴۹ ۴-۱- مقدمه
۴۹ ۴-۲- ایجاد کلاس بند
۴۹ ۴-۲-۱- نوع داده ها
۵۳ ۴-۲-۲- پیش پردازش داده ها
۵۶ ۴-۲-۳- ایجاد کلاس بند هر کاربر
۶۳ ۴-۳- تغییر الگوی کلیدزنی افراد
۶۳ ۴-۴- اثر تغییر ابزار ورودی در الگوی کلیدزنی افراد
۶۶ فصل پنجم . ارائه مدل تصدیق هویت
۶۷ ۵-۱- مقدمه
۶۹ ۵-۲- نتایج روش فاصله اقلیدسی و فاصله اقلیدسی وزن دار
۷۱ ۵-۳- نتایج روش فاصله کوواریانسی
۷۸ نتیجه گیری و پیشنهادات
۸۳ ضمیمه الف - نمودارهای کلیدزنی
۹۷ منابع و مراجع

فصل اول

- شکل (۱-۱). شبکه پتری سطوح اول و دوم کنترل دسترسی ۱۵
- شکل (۱-۲). شبکه پتری سطح سوم کنترل دسترسی ۱۶
- شکل (۱-۳). شبکه پتری سطح چهارم سطح چهارم کنترل دسترسی ۱۷
- شکل (۱-۴). شبکه پتری سطح پنجم کنترل دسترسی ۱۸
- شکل (۱-۵). شبکه پتری سطوح کنترل دسترسی بصورت ترکیب شده ۱۹

فصل سوم

- شکل (۳-۱). نمودار ارتباط سیستم تصدیق هویت بر اساس نحوه کلیدزنی و سیستم سنتی تصدیق هویت ۳۲
- شکل (۳-۲). توابع عضویت مجموعه های فازی، ورودی ها و خروجی ۳۴
- شکل (۳-۳). قوانین فازی مرتبط کننده ورودی ها و خروجی ۳۵
- شکل (۳-۴). مجموعه های فازی استفاده شده در بررسی انطباق یک نمونه کلیدزنی با الگوی از قبل ذخیره شده بصورت (T_1, T_2, \dots, T_n) ۳۷

فصل چهارم

- شکل (۴-۱). مراحل ایجاد کلاس بند تصدیق هویت ۴۹
- شکل (۴-۳). وجود خطای مکث طولانی در رکوردهای یک کاربر ۵۵
- شکل (۴-۴). مراحل ایجاد کلاس بند تصدیق هویت ۶۵

فصل پنجم

- شکل (۵-۱). نمودار خطاهای کلمه PASSWORD مجموعه ۱ ۷۲
- شکل (۵-۲). نمودار خطاهای کلمه PASSWORD مجموعه ۱ پس از فیلترکردن ۷۳
- شکل (۵-۳). نمودار خطاهای کلمه PASSWORD مجموعه ۲ ۷۴
- شکل (۵-۴). نمودار خطاهای کلمه PASSWORD مجموعه ۲ پس از فیلترکردن ۷۵
- شکل (۵-۵). نمودار خطاهای کلمه USERNAME ۷۶
- شکل (۵-۶). نمودار خطاهای کلمه USERNAME پس از فیلترکردن ۷۷

ضمیمه الف

- شکل (۶-۱). نمودار زمانی کلمه PASSWORD کاربر ۱ ۸۵
- شکل (۶-۲). نمودار زمانی کلمه PASSWORD کاربر ۲ ۸۶
- شکل (۶-۳). نمودار زمانی کلمه PASSWORD کاربر ۳ ۸۷
- شکل (۶-۴). نمودار زمانی کلمه PASSWORD کاربر ۴ ۸۸
- شکل (۶-۵). نمودار زمانی کلمه PASSWORD کاربر ۵ ۸۹
- شکل (۶-۶). نمودار زمانی کلمه PASSWORD کاربر ۶ ۹۰
- شکل (۶-۷). نمودار زمانی کلمه PASSWORD کاربر ۷ ۹۱
- شکل (۶-۸). نمودار زمانی کلمه PASSWORD کاربر ۸ ۹۲
- شکل (۶-۹). نمودار زمانی نمونه‌های فیلترشده کلمه PASSWORD کاربر ۷ ۹۳
- شکل (۶-۱۰). نمودار زمانی نمونه‌های فیلترشده کلمه PASSWORD کاربر ۸ ۹۴
- شکل (۶-۱۱). نمودار زمانی کلمه USERNAME کاربر ۷ ۹۵
- شکل (۶-۱۲). نمودار زمانی کلمه USERNAME کاربر ۸ ۹۶

فصل دوم

- جدول (۲-۱). میزان خطای روش فاصله اقلیدسی در شناسایی افراد ۲۷
- جدول (۲-۲). میزان خطای روش احتمال بدون وزن ۲۸
- جدول (۲-۳). میزان خطای روش احتمال وزن دار ۲۹

فصل پنجم

- جدول (۵-۱). داده‌های مجموعه ۱ ۶۸
- جدول (۵-۲). داده‌های مجموعه ۲ ۶۹
- جدول (۵-۳). مقادیر خطای روش فاصله اقلیدسی ۶۹
- جدول (۵-۴). کاربران دارای نمونه‌های ایجاد خطا در روش فاصله اقلیدسی ۷۰
- جدول (۵-۵). مقادیر خطای روش فاصله کوواریانسی با مجموعه آموزش و آزمایش یکسان ۷۱
- جدول (۵-۶). مقادیر خطای روش فاصله کوواریانسی با روش Leaving-One-Out ۷۱

مقدمه

استفاده از خصوصیات بیولوژیکی افراد جهت شناسایی و تأیید آنها در ورود و استفاده از کلیه سیستمها، همواره بمنظور کنترل دسترسی و حفظ امنیت مورد توجه بوده است. روشهای تشخیص گفتار [1]، تشخیص چهره [2]، تشخیص اثر انگشت [3] و تشخیص از روی شبکه چشم [4] از جمله معروفترین روشهایی هستند که از این خصوصیات در شناسایی افراد استفاده می کنند. روشی که ما مورد توجه قرار داده ایم، استفاده از ویژگیهای کلیدزنی افراد است.

عمومی ترین وسیله ارتباطی یک کاربر با کامپیوتر صفحه کلید است. در هنگام کلیدزنی می توان داده های نظیر زمانهای کلیدزنی، میزان فشار وارد شده بر کلیدها، دمای سر انگشت کاربر و جهت ضربه زدن به کلیدها را علاوه بر کاراکترهای وارد شده از کاربر اخذ نمود. با توجه به سخت افزار صفحه کلیدهای موجود عادی ترین داده های قابل اخذ از کاربر زمانهای کلیدزنی شامل زمان فشار یک کلید و زمان رهاسازی کلید می باشد. استفاده از این داده ها علاوه بر سهل الوصول بودن و عدم نیاز به تغییر سخت افزار، نسبت به سایر روشهایی که از ویژگیهای بیولوژیکی افراد استفاده می کنند بدلیل شفاف بودن و در نتیجه مخفی ماندن از دید کاربر دارای مزیت می باشد.

استفاده از این زمانها می تواند همراه با دو رشته کاراکتری وارد شده در هنگام ورود به سیستمها شامل شناسه کاربر¹ و کلمه عبور² در تأیید هویت کاربران استفاده شده و امنیت بالاتری را تأمین می کند. همچنین با دریافت این داده ها از کاربر پس از ورود به سیستم و در هنگام کار دائمی با شناسایی کاربر

¹ UserID

² PassWord

می‌توان یک نظارت دائمی را بر سیستم اعمال نمود. در سیستم‌هایی که حفظ امنیت آنها قابل توجه می‌باشد این مراقبت می‌تواند مانع از سوءاستفاده افراد غیرمجاز، پس از ترک سیستم و باز ماندن آن توسط افراد مجاز شود. در هر دو وضعیت، زمانهای کلیدزنی در دو دسته قابل‌تمایز قرار می‌گیرند. زمان فشارکلید مدت زمان بین فشار یک کلید تا رهاکردن آن کلید می‌باشد و زمان بین دو کلید مدت‌زمان بین رهاکردن یک کلید تا فشار کلید بعدی می‌باشد. در کلیه روشهایی که برای دو موضوع شناسایی کاربران و تایید هویت آنها مطرح می‌باشد از یک یا هر دوی این ورودی‌ها [5][6][8] و در بعضی از روشها نیز ورودی دیگر نظیر سختی کلیدزنی [7] استفاده می‌شود.

در مبحث تصدیق‌هویت بر اساس نحوه کلیدزنی، از سال ۱۹۹۰ تاکنون کارهایی انجام شده‌است

که گزارش تعدادی از آنها را در دست داریم و بطورمختصر روشها و نتایج آنها را مرور می‌کنیم.

اولین کار منتشرشده، در سال ۱۹۹۰ توسط آقایان Joyce و Gupta بوده‌است [5] که بعنوان

یک کار پایه در تحقیقات دیگران به آن اشاره شده است. در این تحقیق با استفاده از معیار فاصله

اقلیدسی یک دسته ویژگی‌های زمانهای بین کلیدهای متوالی (Latency) با یکدیگر مقایسه شده‌اند. این

کلاس‌بند با جمع‌آوری داده‌هایی از ۳۳ کاربر بر روی ایستگاههای کاری SUN و با دقت ۱۰ میلی‌ثانیه

آزمایش‌شده و دقتی در حدود ۹۵٪ را نشان داده‌است.

کار دیگری که در همان سال توسط Slivinsky، Bleha و Hussien انجام شده‌است [6] بر

اساس رساله دکترای آقای Bleha در سال ۱۹۸۸ بوده است. این تحقیق با استفاده از روشهای آماری و

براساس یک رشته ثابت با طول زیاد و با پیش فرض اولیه توزیع نرمال برای داده‌ها انجام می‌شود و در آن

یک احتمال گاوسی چندمتغیره بر روی ویژگی‌های زمانهای بین کلیدی محاسبه می‌شود. بهترین نتایج این

کار دقتی در حدود ۹۶٪ و ۹۷٪ را نشان می‌دهد.

در سال ۱۹۹۷ آقایان De Ru و Ellof کلاس‌بندی را پیشنهاد کرده‌اند که در آن بطور کامل از روش فازی استفاده شده است. [7] در این کلاس‌بند از ویژگی‌های زمانهای بین‌کلیدی به‌همراه ویژگی جدیدی به نام سختی کلیدزنی استفاده شده است. برای ویژگی‌های زمانهای بین‌کلیدی چهار مجموعه فازی و برای ویژگی سختی کلیدزنی یک مجموعه فازی در نظر گرفته شده است. در این روش خروجی، دسته‌بندی زوج کلید متوالی است و برای آن چهار مجموعه فازی در نظر گرفته شده است. پس از دسته‌بندی اولیه، کلیه خروجی‌ها با یک قانون فازی با یکدیگر ترکیب شده و یک خروجی تک مقداری بدست می‌آید. در آزمایش این کلاس‌بند داده‌ها شامل تعدادی نام کاربری متفاوت بوده و دقت ۹۲/۵٪ و ۹۷٪ بدست آمده است.

در همین سال Obaidat و Sadoun تعدادی از شبکه‌های عصبی را برای ایجاد کلاس‌بند مناسب بکار گرفته‌اند [8]. این تحقیق در ادامه کار قبلی آقای Obaidat در سال ۱۹۹۴ است که با استفاده از شبکه عصبی MLP صورت گرفته است [28]. در این تحقیق بجز ویژگی‌های زمانهای بین‌کلیدی از زمانهای فشار کلیدها (Durations) نیز استفاده شده است. براساس این دودسته ویژگی تعدادی از شبکه‌های عصبی آزمایش شده‌اند. برای آزمایش‌های عملی نیز تعداد زیادی نمونه در مدت‌زمان حدود دوماه بطور مداوم جمع‌آوری شده است. پس از انجام آزمایش‌هایی بهترین نتایج با شبکه‌های Fuzzy ARTMAP , RBFN, LVQ و با دقت بالای ۹۹٪ گزارش شده است.

آخرین تحقیق گزارش شده توسط Robinson و Liang در سال ۱۹۹۸ بوده است [9]. که با استفاده از روشهای آماری و بکاربردن یک کلاس‌بند یادگیرنده استنتاجی^۳ بر روی حجم نسبتاً کم داده‌ها صورت گرفته است. در اینکار نیز از دو دسته ویژگی‌های زمانهای بین‌کلیدی و زمانهای فشار کلیدها

³ Inductive Learning Classifier