

## عنوان مقاله:

بهبود تشخیص ناهنجاری در سیستم های ارایه دهنده سرویس های ایمیل با استفاده از طبقه بندی مبتنی بر محتوا با ماشین بردار پشتیبان

## محل انتشار:

دوازدهمین سمپوزیوم پیشرفت های علوم و تکنولوژی کمیسیون دوم: سرزمین پایدار تازه های کامپیوتر و فناوری اطلاعات (سال: 1396)

تعداد صفحات اصل مقاله: 9

## نویسندگان:

ابراهیم فردهفتخوانی - کارشناسی ارشد فناوری اطلاعات، امنیت اطلاعات، موسسه آموزش عالی مهرآستان، گیلان

غلامحسین اکباتانی فرد - استادیار گروه کامپیوتر، واحد لاهیجان، دانشگاه آزاد اسلامی، لاهیجان، ایران

## خلاصه مقاله:

از جایی که ایمیلها دارای ویژگیهایی نظیر تعداد کلمات متن پیام، تعداد کلمات عنوان پیام و... میباشند هدف این پژوهش یافتن ویژگیهای استخراج شده از ایمیلها و انتخاب مناسبترین و بهترین آنها به منظور تشخیص ناهنجاریها در ایمیلهای ارسالی و برقراری ارتباطات است. روند کار پژوهش این گونه است که در گام اول ویژگیهای مختلف برای هر کاربر براساس ایمیلهای مبادله شده استخراج میشود و سپس براساس الگوریتم ژنتیک، بهترین ویژگیها برای تشخیص ناهنجاری توسط طبقه بند ماشین بردار پشتیبان انتخاب میشود. در این پژوهش برای بررسی کارایی سیستم پیشنهادی و تولید یک مدل رفتاری برای شخص عملکرد و ارتباطات یک کاربر در مجموعه داده به مدت یک ماهه بخشبندی و بررسی شد. به منظور بررسی کارایی سیستم پیشنهادی نتایج مطلوبی که از شبیه سازی حاصل شد به این صورت می باشد که روش پیشنهادی با دسته بندهای KNN و ANN مورد مقایسه قرار گرفت. در آزمایش خطای تشخیص ناهنجاری توسط دسته بند SVM بهوسیله ی الگوریتم ژنتیک 18 ویژگی اصلی در نظر گرفته شد و انتخاب 10 ویژگی به بالا میزان خطای سیستم حدود 5.2 درصد بدست آمد اما با انتخاب 8 ویژگی میزان هزینه نهایی به کمترین حالت خود رسید. همچنین در آزمایش خطای تشخیص ناهنجاری توسط دسته بند KNN به وسیله ی الگوریتم ژنتیک تعداد تکرار الگوریتم ژنتیک 100 مرتبه در نظر گرفته شد و با انتخاب 14 ویژگی به بالا میزان خطای سیستم به 6.8 درصد رسید که انتخاب 8 ویژگی سبب کاهش همزمان حجم محاسبات و خطای تشخیص میگردد. در نهایت در آزمایش خطای تشخیص ناهنجاری توسط دسته بند ANN به وسیله ی الگوریتم ژنتیک، با یک لایه و تعداد نرونها از 5 الی 10 بررسی شد و کمترین میزان خطا با 5 نرون نتیجه شد و حدود 6.5 درصد خطا با انتخاب 14 ویژگی بهدست آمد. در نتیجه میتوان گفت ماشین بردار پشتیبان هم از نظر میزان محاسبات و هم از نظر میزان خطای تشخیص ناهنجاری از دو دسته بند دیگر عملکرد بهتری دارد و طبق نتایج بهدست آمده دیگر میتوان گفت با استفاده از دسته بند ماشین بردار پشتیبان و انتخاب 70 درصد دادهها به عنوان دادههای آموزش میتوان به خوبی ناهنجاریهای یک سرویس ایمیل را تشخیص داد.

## کلمات کلیدی:

تشخیص ناهنجاری، سرویس ایمیل، طبقه بند، SVM، ماشین بردار پشتیبان و الگوریتم ژنتیک

## لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/725774>

