ریسک سنج: ابزاری برای سنجش دقیق میزان ریسک امنیتی برنامه‌ها در دستگاه‌های همراه

محمود دی پیر

دانشکده رایانه و فناوری اطلاعات، دانشگاه علوم و فنون هوایی شهید ستاری، تهران، ایران

چکیده

حفظ امنیت دستگاه‌های همراه به دلیل نگهداری اطلاعات شخصی و کاری برای کاربران آنها بسیار حائز اهمیت است. نصب برنامه‌های جدید و ناشناخته روی این دستگاه‌ها ممکن است، منجر به آسیب‌های امنیتی شود؛ بنابراین محاسبه ریسک امنیتی برنامه‌ها در انجام تصمیم‌گیری درست در انتخاب نرم‌افزار، به کاربران می‌تواند کمک کند. در برخی از سیستم عامل‌های دستگاه‌های همراه، ریسک امنیتی برنامه‌ها از طریق مجوزهایی که درخواست می‌کنند قابل اندازه‌گیری است. در این مقاله، ابزار نرم‌افزاری جدیدی به‌منظور سنجش میزان ریسک امنیتی برنامه‌ها در دستگاه‌های همراه طراحی و پیاده‌سازی شده است. این ابزار از یک معیار جدید به‌منظور اندازه‌گیری ریسک بهره می‌برد. ما به‌منظور ارائه این معیار، مجوزهای درخواستی توسط ده‌ها بدافزار و صدها برنامه تلفن همراه را بررسی و تحلیل کرده‌ایم. علاوه‌بر این، به‌منظور ارزیابی دقیق‌تر، مجموعه داده‌های جدیدی از برنامه‌های ارائه‌شده در فروشگاه‌های داخلی و بدافزارهای جدید را گردآوری کرده‌ایم. آزمایش‌های صورت‌گرفته بر روی بدافزارها و نرم‌افزارهای بی‌خطر شناخته‌شده، نشان‌دهندهٔ دقت روش ارائه‌شده نسبت به معیارهای ارائه‌شدهٔ قبلی از نظر تخصیص ریسک امنیتی بالا به بدافزارها و ریسک پایین به نرم‌افزارهای بی‌خطر است.

واژگان کلیدی: امنیت تلفن همراه، ریسک امنیتی، بدافزار، مجوزهای امنیتی، ریسک‌سنج.

RiskMeter: A Tool for Measuring Precise Security Risk Values of Mobile Device Applications

Mahmood Deypir

Faculty of Computer and Information Technology, Shahid Sattari Aeronautical University of Science and Technology,
Tehran, Iran.

Abstract

Nowadays smartphones and tablets are widely used due to their various capabilities and features for end users. In these devices, accessing a wide range of services and sensitive information including private personal data, contact list, geolocation, sending and receiving messages, accessing social networks and etc. are provided via numerous application programs. These types of accessibilities, functionalities, and facilities make privacy and security issues more critical. Therefore, traditional security mechanism including biometric authentication, data encryption, access control, and etc. are not adequate. Therefore, danger of installing and using malwares must be taken into account in order to provide practical security for end users. Installing new and unknown applications on these devices might lead to security threats. Recently, smartphones and tablets utilize powerful operating system in which security of application is provided by application permissions. Android and BlackBerry are two examples of operating systems which reduce attack surface by using application permissions. In these operating systems, in order to perform malicious activities, an attacker must deceive users to install a malicious app since other ways of intrusion are almost closed. Recent statistics show that Android is the most popular operating system. For installing an app,

Android requires the user to grant privileges through the requested permissions. There is a large number of applications (Apps) developed for this operating system which require various permissions based on their functionalities and provided services. Therefore, measuring security risks of applications can help us to make better decision regarding to apps installation and removal. There exists some research regarding to enhance the Android security model and its security risk communication mechanism. In this mobile operating system, security risk values of applications can be computed using their requested permissions. In this study, a new software tool is designed and implemented to measure security risk values of mobile applications. This tool benefits from a new metric to compute the risk values. This risk metric exploits statistics of permission usages in known malwares and goodwares. However, they can be simply extended to other features of Android apps including static and dynamic ones. Moreover, we have attempted to give a better definition of permission criticality to aim users for making best decision in new apps installation or previously installed ones removal. In fact, we have designated a new formulation to assign higher risk values to permissions with a higher usage in malwares and very lower usage in benign apps. The idea is quite simple but produces interesting results. That is, the security risk of a permission is directly related to the difference of its usage in malicious and non-malicious apps. Given risk values of permissions, one can compute risk of an Android app based on its permission list. Since the proposed measurement compute the risk values of permissions according to simple statistics of known malwares and useful Android apps, they have good explainability. Users can be informed regarding to danger about approving risky permissions and they can make reasonable decisions based on total risk score of an app which can be simply computed using security risks of its requested permissions. In order to purpose the metric, we have analyzed requested permissions of large number of malicious and ordinary applications. Moreover, for realistic evaluations, we have constructed two new datasets of applications belonging to an Iranian market and new malwares. Experimental evaluations on real known malwares and benign apps reveal the superiority of the proposed criterion with respect to previously proposed method in terms of assigning higher risk values to malwares and lower risk values to the benign applications.

Keywords: Security of mobile devices, Security risk, Malwares, Permissions, RiskMeter.

۱- مقدمه

امروزه دستگاه‌های تلفن همراه قابلیت‌های متنوعی دارند و در محیط‌های مختلف قابل استفاده هستند. این دستگاه‌ها برای کاربران عادی و سازمانی مصارف متعددی دارند. بنابراین وجود داده‌های حساس، شخصی و سازمانی بر روی تلفن‌های همراه مسئلهٔ نشت اطلاعات و امنیت را برای کاربران آنها حساس‌تر کرده است. شرکت‌های سازنده این دستگاه‌ها، سیستم عامل‌های مختلفی مانند iOS، اندروید و بلک بری را توسعه داده‌اند. از میان سیستم عامل‌های توسعه داده شده برای تلفن همراه و دستگاه‌های هوشمند قابل حمل، سیستم عامل اندروید فراگیرترین آنهاست و سیستم عامل بلک بری نیز از نظر امنیتی قابلیت‌های خوبی دارد. برای این سیستم عامل‌ها تاکنون نرم‌افزارهای زیادی توسعه داده شده‌اند. بسیاری از این نرم‌افزارها توسط افراد ناشناخته و توسعه‌دهندگان گمنام ارائه شده‌اند و استفاده از آنها با خطرات امنیتی بالقوه‌ای همراه خواهد بود. بنابراین برای حفظ امنیت دستگاه‌های همراه، به‌کارگیری روش‌هایی چون احراز هویت[1]، رمزنگاری داده‌ها یا استفاده از گذرواژه‌های قوی به‌تنهایی کافی نیست و می‌بایست خطر برنامه‌های مخرب را نیز جدی گرفت. برنامه‌های مخرب به چند طریق بر امنیت و حریم خصوصی افراد می‌توانند تأثیرگذار باشند. نخست، برنامه‌های مخرب داده‌های شخصی و خصوصی افراد را در اختیار نفوذگر می‌توانند قرار دهند. دوم، برخی از برنامه‌های مخرب به هدف دزدی از کارت‌های اعتباری افراد طراحی می‌شوند؛ زیرا بسیاری از کاربران، کارهای بانکی خود را با استفاده از دستگاه‌های همراه انجام می‌دهند. سوم، برنامه‌های مخربی وجود دارند که از امکانات پیام کوتاه یا تماس تلفنی به‌منظور تبلیغات و اهداف مالی طراحی می‌شوند. چهارم، دسته‌ای از برنامه‌ها که موسوم به باج‌افزارها هستند، می‌توانند اطلاعات مهم شخص را رمزنگاری کرده و به‌منظور رمزگشایی آنها، از کاربران پول طلب کنند. پنجم، برخی از برنامه‌ها با هدردادن منابع سخت افزاری سبب کندشدن خدمات‌دهی می‌شوند. موارد ذکرشده مهمترین انواع برنامه‌های مخرب دستگاه‌های همراه هستند.

ساختار و مدل امنیتی سیستم عامل‌های دستگاه‌های همراه تأثیر زیادی در محافظت از آنها در مقابل انواع برنامه‌های مخرب دارد. مدل امنیتی نرم‌افزارها در سیستم عامل اندروید و بلک بری بر اساس مجوز[1]هاست. این مجوزها، برنامه‌ها را در دسترسی به منابع سخت افزاری و

[1] Permission

نرم‌افزاری دستگاه محدود می‌کنند. در سیستم عامل اندروید، این مجوزها فقط یکبار و در ابتدای نصب هر نرم‌افزار از کاربر پرسیده می‌شوند و پس از آن، کاربر چندان دخالتی در امنیت نرم‌افزار نصب شده ندارد. به‌طورمعمول خود کاربران هم وقت زیادی به‌منظور مطالعه فهرست مجوزها در صفحهٔ ابتدایی نصب نرم‌افزار، اختصاص نمی‌دهند. علاوه‌براین، کاربران عادی به‌طورمعمول دانش فنی برای شناخت تأثیر استفاده از هر مجوز در نرم‌افزار مورد استفاده را ندارند. بنابراین این مدل امنیتی کارایی چندانی در ارتقای امنیتی کاربران در حفظ داده‌های شخصی و حریم خصوصی آنها ندارد. نرم‌افزارهای مخرب مانند تروجان‌ها، جاسوس افزار[1]ها و تبلیغ افزار[2]ها قادرند با فریب کاربران، خود را در قالب نرم‌افزار مفید و بی‌خطر[3] نشان داده و اطلاعات حساس شرکت‌ها و مراکز نظامی را به سرقت برند. این‌گونه بدافزار[4]ها همچنین می‌توانند با سرقت داده‌های شخصی افراد و انتشار آنها، حریم خصوصی آنها را نقض کنند. تاکنون پژوهش‌هایی به‌منظور آگاه‌سازی بهتر کاربران در زمینه امنیت نرم‌افزار در اندروید صورت گرفته است [15]. استفاده از عناوین مناسب‌تر برای مجوزها، دسته‌بندی مجوزها، کاهش تعداد عنوان مجوزها، استفاده از نظرات کاربران علاوه‌بر مجوزها، نمونه‌هایی از راه‌کارهایی ارائه‌شده در این پژوهش‌ها هستند. علاوه‌براین، تاکنون معیارهای مختلفی نیز برای سنجش ریسک یک نرم‌افزار اندرویدی ارائه شده است. تعداد مجوزهای حساس درخواستی، تعداد جفت مجوزهای حساس درخواستی نمونه‌هایی از این معیارها هستند [16]. براساس این معیارها و داشتن یک حدآستانه، می‌توان پس از سنجش ریسک یک نرم‌افزار مشکوک و در صورت بالابودن ریسک آن، هشدار امنیتی صادر کرد. در این مقاله معیار جدیدی به‌منظور اندازه‌گیری ریسک یک نرم‌افزار اندرویدی ارائه شده که نسبت به معیارهای ارائه شدهٔ قبلی کارایی بهتری دارد. بر اساس معیار ارائه‌شده، ابزار جدیدی به نام ریسک‌سنج در محیط اندروید توسعه داده شده است که به کاربران در بررسی و مقایسه برنامه‌ها از نظر ریسک امنیتی کمک می‌کند.

در بخش بعد به معرفی برخی از کارهای پژوهشی انجام‌شده مرتبط با امنیت اندروید می‌پردازیم. در بخش سوم مسئله را به‌طور رسمی بیان می‌کنیم؛ سپس در بخش چهارم معیار جدید پیشنهادی را بر اساس مفاهیم پایه‌ای ارائه داده و نحوهٔ محاسبه آن را تشریح می‌کنیم. در این بخش همچنین ابزار ریسک‌سنج را معرفی می‌کنیم. در بخش پنجم پس از معرفی مجموعه داده‌ها، آزمایش‌هایی را به‌منظور ارزیابی و مقایسه معیار پیشنهادی با معیارهای ارائه‌شده قبلی، ارائه می‌دهیم. درنهایت مقاله را در بخش ششم جمع‌بندی و نتیجه‌گیری می‌کنیم.

۲- مروری بر پژوهش‌های گذشته

با توجه به گسترش روز افزون دستگاه‌های تلفن همراه و کاربردهای آنها، در همین اواخر توجه ویژه‌ای به امنیت این دستگاه‌ها شده و پژوهش‌های زیادی در این حوزه انجام گرفته است. برخی از این پژوهش‌ها بر مبنای سیستم عامل‌های خاصی از دستگاه‌های همراه انجام شده و برخی دیگر جامعیت بیشتری دارند. شبارو و همکاران در [24] مدل جدیدی از کنترل دسترسی در دستگاه‌های همراه ارائه داده‌اند که خاص این نوع سامانه‌هاست. در این مدل، دسترسی‌ها با توجه به محیط کاری دارنده دستگاه تعیین می‌شود. به‌عنوان مثال اگر شخص در محل کار خود و در اتاق جلسه حضور داشته باشد، به‌منظور جلوگیری از افشای گفتگوهای کاری، دسترسی‌های مربوط به میکروفن قطع می‌شود. در پژوهشی دیگر ژو و همکاران در [28] به بررسی تخلفات در رتبه‌بندی برنامه‌ها پرداخته و سعی کرده‌اند، منابع مختلف مربوط به تشخیص این تخلفات را گردآوری و یکپارچه‌سازی کنند. نظر به اینکه دستگاه‌های همراه در سیستم‌های حساس مورد استفاده می‌توانند قرار گیرند، چگونگی سوء استفاده بدافزارها از آسیب‌پذیرهای ناشناخته دستگاه‌های تلفن همراه و تأثیر آن در امنیت مراکز حساس در [2] مورد بررسی قرار گرفته است. همچنین چارچوبی به‌منظور تشخیص و مقابله با دسترسی مستقیم بدافزارها به توابع حساس هسته سیستم‌عامل دستگاه‌های همراه ارائه شده است. یک دستگاه تلفن همراه ممکن است حاوی اطلاعات حساس خصوصی و یا کاری باشد. چگونگی برخورد با ریسک‌های مربوط به امنیت اطلاعات در زمانی که این دستگاه‌ها دزدیده و یا گم می‌شوند توسط توو و همکاران در [25] بررسی شده است. اگرچه سیستم عامل‌های متنوعی برای دستگاه‌های تلفن همراه ارائه شده ، اما بررسی‌های آماری نشان می‌دهد که سیستم عامل اندروید نسبت به سایر سیستم عامل‌ها بین کاربران محبوب‌تر و پرکاربردتر است. با

[1] Spyware
[2] Adware
[3] Benign App
[4] Malware

توجه به گستردگی استفاده و معماری امنیتی خاص اندروید و محدودیت‌های آن، پژوهش‌های مختلفی در این حوزه انجام گرفته است. مطالعات نشان می‌دهند که کاربران اغلب از بررسی مجوزهای درخواستی نرم‌افزارها در اندروید صرف‌نظر می‌کنند. در برخی از پژوهش‌های انجام‌شدهٔ اخیر، تلاش شده است بر این مشکل چیره شوند [17]،[14]،[13]،[8]. از جمله این تلاش‌ها روش‌هایی مانند تغییر دسته‌بندی مجوزها، تأکید بر مفهوم ریسک و چگونگی اختصاص مجوزها هستند. در [18] اطلاعاتی سطح بالا شامل موارد حفظ حریم خصوصی مانند داده‌های شخصی، مکانی و فهرست دفترچه تلفن، به جای فهرست مجوزها در صفحه معرفی نرم‌افزار پیشنهاد شد. به‌منظور کاهش فضای لازم برای نمایش اینگونه اطلاعات و کمک به کاربر برای تصمیم‌گیری بهتر در انتخاب و نصب، در [15] عوامل ریسک و ایمنی[1] که حاصل آنها مقادیر کمی بوده و از روی مجوزها قابل محاسبه هستند، ارائه شدند. با بررسی کاربران مشخص شد که این معیارها تأثیر بیشتری نسبت به اطلاعات متنی دارند. پنگ و همکارانش روشی اساسی بر مبنای مدل احتمالی را به‌منظور رتبه‌بندی نرم‌افزارهای اندروید بر اساس مجوزهای درخواستی، ارائه دادند. این روش قادر است یک نرم‌افزار را در میان سایر نرم‌افزارهای موجود در یک فروشگاه نرم‌افزاری مانند فروشگاه گوگل، رتبه‌بندی کند [19]. چنین رتبه‌بندی‌هایی به انتخاب نرم‌افزارهای با امنیت بیشتر توسط کاربران می‌تواند کمک کنند. در مقابل پژوهش‌های مرور شدهٔ بالا، دسته دیگری از پژوهش‌ها وجود دارند که روش‌هایی را به‌منظور دسته‌بندی نرم‌افزارها و بدافزارهای اندرویدی ارائه داده‌اند. برخی از این پژوهش‌ها با استفاده از مجوزهای درخواستی نرم‌افزارها، به تشخیص نرم‌افزارهای مخرب پرداخته‌اند [21]،[12]،[6]. برخی نیز از تحلیل ایستای کد نرم‌افزار، توابع برنامه نویسی مورد استفاده و مطابقت آن با برخی الگوهای موجود بدافزارها، به‌منظور تشخیص بدافزارهای جدید استفاده کرده‌اند [27]،[22]،[10]. تعدادی از پژوهش‌گران نیز روش‌هایی را ارائه داده‌اند که با تحلیل رفتاری نرم‌افزار درحال اجرا، سعی در تشخیص نرم‌افزارهای مخرب اندرویدی دارند [23]،[20]،[9]،[5].

باررا و همکارانش روشی را برای ارزیابی عملی مدل‌های امنیتی بر اساس مجوز به کمک نقشه‌های خود سازمانده[2] ارائه داده‌اند [4]. آنها روش خود را به‌منظور تحلیل توزیع مجوزها بر روی هزار برنامه اعمال کرده و نشان دادند که چگونه استفاده از مجوز با دسته‌بندی برنامه‌ها ارتباط پیدا می‌کند [5]. در [11] تلاش شده است که با دیکامپایل کردن و تحلیل کد به‌دست‌آمده نرم‌افزارها نشت داده را تشخیص دهند. همچنین سامانه‌ای را توسعه دادند که ترکیب مجوزهای خطرناک را به‌منظور چگونگی برآورده‌کردن سیاست‌های امنیتی به کار می‌برد [12]. در این سیاست‌ها، به‌صورت دستی ترکیب مجوزهای خطرناک مانند FINE_LOCATION و INTERNET در نظر گرفته شده است. این ترکیبات خود با تحلیل بدافزارهای شناخته‌شده به‌دست می‌آیند. ابزاری به‌نام MAST [7] توسعه داده شده که نرم‌افزارهایی را که به‌احتمال زیاد بدافزار هستند، بر اساس تحلیل کد و تحلیل مجوزها تشخیص می‌دهد. ابزار PScout [3] با استفاده از تحلیل کد ایستا روی اندروید چگونگی نگاشت مجوز به تابع را در این سیستم عامل بررسی می‌کند. این ابزار نشان داد که سامانهٔ مجوزهای اندروید کمینه افزونگی را داشته و این مسئله با توسعه اندروید و ارائه نسخه های جدید نیز پایدار باقی مانده است. ما در این پژوهش به دنبال معیاری هستیم که ریسک امنیتی نرم‌افزارهای ناشناخته را در اندروید به‌خوبی محاسبه کند.

۳- بیان مسئله

برنامه‌های مخرب همواره تلاش دارند که با روش‌های متنوع کاربر را فریب داده و او را وادار به نصب خود کنند. اگر یک کاربر میزان خطر امنیتی برنامه‌ها را بداند انتخاب‌های بهتری می‌تواند داشته باشد و برنامه‌های با ریسک پایین را انتخاب و نصب کند. سیستم‌عامل‌هایی مانند اندروید و بلک بری دارای مجموعه‌ی مشخصی از مجوزها هستند. هر برنامه برای نصب خود نیاز به زیرمجموعه‌ای از این مجوزها دارد. توجه و مطالعه مجوزها قبل از نصب، اگرچه تا حدی از این مشکلات می‌کاهد، اما کاربران به‌طورمعمول به‌دلیل دانش تخصصی پایین یا عدم صرف وقت کافی از این قابلیت چندان استفاده نمی‌کنند. علاوه‌بر این، خطر امنیتی هر مجوز را نمی‌دانند؛ زیرا از سابقه سوء استفاده از مجوزها در بدافزارهای قبلی مطلع نبوده و از طرفی نمی‌دانند که با تأیید یک مجوز چه راه‌هایی را برای نفوذگر باز می‌کنند. شکل (۱) نشان می‌دهد که مجوزها تأثیر مختلفی در امنیت یک دستگاه تلفن همراه دارند. به‌عنوان مثال مجوزهای WRITE_SMS و INSTALL_PACKAGES بزرگتر نشان داده شده‌اند چون

[1] Safety
[2] Self-Organizing Map

تأثیر این‌ها نسبت به مجوزهای دیگر در امنیت یک دستگاه همراه بیشتر است.

(شکل-۱): مجوزها تأثیرات متفاوتی بر امنیت دستگاه‌های همراه دارند.

(Figure-1): Permissions have various impacts on the security of mobile devices.

هدف فرعی ما در اینجا محاسبه میزان ریسک امنیتی مجوزهاست. چون هر برنامه تلفن همراه برای اجرای خود نیاز به مجوزهایی دارد. با استفاده از ریسک امنیتی مجوزها، ریسک امنیتی برنامه‌ها را نیز می‌توان محاسبه کرد. به بیان دقیق‌تر فرض کنید که مجموعهٔ P شامل N مجوز یک سیستم عامل همراه است. این مجموعه به‌صورت $P = \{x_1,x_2,x_3,...,x_N\}$ تعریف می‌شود. یک برنامه قابل اجرا به‌نام A در این سیستم عامل، برای اجرای خود یک زیرمجموعه از P را درخواست می‌کند. می‌توان متغیری دودویی x_{iA} را به معنی وضعیت درخواست مجوز i ام در نرم‌افزار A تعریف کرد. مقدار این متغییر می‌تواند صفر باشد؛ یعنی این مجوز درخواست نشده است و یا یک باشد یعنی این مجوز در زیرمجموعه مجوزهای درخواستی A وجود دارد. هدف اصلی ما در اینجا محاسبه ریسک امنیتی نرم‌افزار A باتوجه به زیرمجموعهٔ مجوزهای درخواستی آن است. به این منظور نیاز است که برای هریک از عناصر مجموعهٔ P، میزان خطر را محاسبه کرده باشیم. ما این مقادیر را باتوجه به سوابق استفاده از هر مجوز در برنامه‌های مفید و مخرب محاسبه خواهیم کرد. بنابراین مقادیر، خطر امنیتی هر برنامه را می‌توانیم محاسبه کرده و در صورت لزوم به کاربر در مورد انتخاب و نصب آن هشدار دهیم.

۴- معرفی معیار پیشنهادی و ابزار ریسک‌سنج

در پژوهش‌های گذشته تعاریف مشابهی برای مفهوم ریسک امنیتی یک نرم‌افزار اندروید با توجه به مجوزهایی که استفاده می‌کند، ارائه شده است. در [16] سه معیار آماری مختلف اندازه‌گیری ریسک با توجه به مجوزهای درخواستی نرم‌افزار بررسی شده است. برخی از این معیارها در پژوهش‌های گذشته نیز مورد توجه قرار گرفته بود [12]. هر سه معیار بر اساس مفهومی به نام مجوز بحرانی عمل می‌کنند. یک مجوز بحرانی، مجوزی است که قبلاً در بدافزارهای اندرویدی شناخته‌شده، استفاده شده باشد. معیار نخست تعداد مجوزهای بحرانی درخواستی است که RCP[1] نامیده می‌شود. معیار دوم تعداد جفت مجوزهای بحرانی درخواستی است که RPCP[2] نام دارد و معیار سوم هم که با RCP+wPRCP نشان داده می‌شود با ترکیب این دو معیار و انتساب وزن دلخواه w به جفت مجوزهای بحرانی، به‌دست می‌آید [16]. از این معیارها می‌توان به‌منظور اعلام هشدار برای نرم‌افزارهای مشکوک و یا شناسایی بدافزارهای ناشناخته جدید، استفاده کرد. ما به‌دنبال معیاری برای ریسک امنیتی هستیم که ضمن ساده‌بودن، توصیف دقیق‌تری را از ریسک امنیتی نرم‌افزار در اندروید نشان دهد. علاوه بر آن بتواند به نرم‌افزارهای مفید ریسک امنیتی پایین و به نرم‌افزارهای مخرب نسبت به نرم‌افزارهای مفید ریسک امنیتی بالایی تخصیص دهد. ما برای به‌دست‌آوردن این معیارها، مجوزهای درخواستی ۸۰۸ بدافزار و ۱۳۶۵۳۴ نرم‌افزار شناخته‌شده را تحلیل کرده‌ایم. ابتدا یک معیار اولیه پیشنهاد داده و با توسعهٔ آن به معیار دومی می‌رسیم که کارایی بیشتری دارد. ابتدا معیار نخست را معرفی می‌کنیم. در این معیار وزن هر مجوز را در نظر می‌گیریم. این وزن با توجه به فراوانی استفاده از هر مجوز در تروجان‌ها، تبلیغ افزارها و درکل بدافزارهای شناخته شدهٔ اندروید، محاسبه می‌شود. ما کلیه مجوزهای بحرانی اندروید را از یک تا N شماره‌گذاری کرده‌ایم. به این ترتیب، برای هر مجوز x_i، وزن w_i را به‌صورت زیر تعریف می‌کنیم:

$$W(x_i)=\frac{\sum_{j=1}^{NM} x_{ij}}{NM} \qquad x_{ij}\in\{0,1\},\forall i\in\{1,2,...,N\} \qquad (۱)$$

[1] Rare Critical Permission
[2] Rare Pair Critical Permission

در فرمول بالا x_{ij} نشان‌دهنده وجود یا عدم وجود مجوز بحرانی iام در بدافزار jام است. با تقسیم این عدد بر NM[1] یعنی تعداد کل بدافزارهای مورد بررسی، عددی نرمال بین صفر و یک به‌دست می‌آید که نشان‌دهنده اهمیت مجوز iام در تشخیص بدافزارهاست. به عبارت دیگر هرچه یک مجوز مانند i در بدافزارهای بیشتری استفاده شده باشد، w_i به یک نزدیک‌تر خواهد شد. بدیهی است که این وزن‌ها ثابت نبوده و با گذشت زمان تغییر می‌کنند و می‌بایست در هر دوره زمانی دوباره محاسبه شوند؛ زیرا همواره بدافزارهای جدیدی معرفی می‌شوند و الگوی بدافزار نویسی برحسب استفاده از مجوزها تغییر می‌کند. علاوه‌براین، تعداد مجوزهای بحرانی نیز ممکن است تغییر کند. با توجه به این وزن‌ها، به‌منظور سنجش ریسک امنیتی یک نرم‌افزار اندرویدی، معیار ریسک امنیتی *RValue* را با فرمول زیر ارائه می‌دهیم:

(۲)

$$RValue(A)=\begin{cases}\dfrac{\sum_{i=1}^{N}(W(x_i)\times x_i)}{\sum_{i=1}^{N}x_i} & \sum_{i=1}^{N}x_i>0 \quad \forall x_i\in\{0,1\}\\ 0 & \sum_{i=1}^{N}x_i=0\end{cases}$$

در این فرمول منظور از A نرم‌افزاری است که می‌خواهیم ریسک امنیتی را برای آن محاسبه کنیم. x_i و w_i به‌ترتیب وضعیت درخواست مجوز iام توسط نرم‌افزار مورد تحلیل و وزن مجوز iام را نشان می‌دهند. متغیر x_i دودویی بوده و w_i هم مطابق فرمول (۱) درصدی از نرم‌افزارهای مخرب اندروید را نشان می‌دهد که از مجوز iام را استفاده کرده‌اند. N در همهٔ فرمول‌ها تعداد کل مجوزهاست. مخرج کسر مجموع مجوزهای بحرانی درخواست‌شده را نشان داده و به‌منظور نرمال‌سازی ریسک اضافه شده است. نرمال‌سازی به‌دلیل ارائه‌ی مقادیر به شکل درصد به کاربر انجام شده است. ما به‌منظور مقایسه معیارها در بخش ارزیابی، سایر معیارهای ارائه‌شدهٔ قبلی را نیز به‌صورت مشابه، نرمال‌سازی کرده‌ایم. در همهٔ این فرمول‌ها، درصورتی‌که نرم‌افزار مورد بررسی هیچ مجوزی را درخواست نکند، میزان ریسک امنیتی آن صفر خواهد بود.

ما برمبنای ریسک نرم‌افزار، ریسک سیستم اندرویدی را معرفی می‌کنیم. با اعتماد به سیستم عامل نصب‌شده بر روی یک دستگاه تلفن همراه، می‌توان ریسک کلی یک دستگاه اندرویدی مانند S را با توجه به نرم‌افزارهای نصب‌شده بر روی آن به‌صورت زیر محاسبه کرد:

$$RValue(S)=\frac{\sum_{i=1}^{N}(w(x_i)\times n_i)}{NA\times\sum_{i=1}^{N}x_i} \qquad (۳)$$

در این فرمول n_i تعداد برنامهٔ نصب‌شده بر روی سیستم است که از مجوز بحرانی iام را استفاده می‌کنند. NA نیز تعداد کل برنامه های نصب شده بر روی سیستم عامل اندروید دستگاه را نشان می‌دهد. همانند فرمول‌های قبلی، w_i وزن مجوز بحرانی iام را نشان می‌دهد. فرمول (۳) هم با توجه به نرم‌افزارها، مقداری نرمال بین صفر و یک را برای ریسک امنیتی یک سیستم به ما می‌دهد.

اگرچه معیار پیشنهادی *RValue*، فراوانی مجوزها در بدافزارهای مختلف را در نظر گرفته و در محاسبه ریسک امنیتی تأثیر آنها را لحاظ می‌کند، اما این معیار هم، معیار کاملی نیست. یک معیار مناسب باید هم فراوانی استفاده از مجوزها در بدافزارها را در نظر گرفته و هم این فراوانی را در نرم‌افزارهای مفید در نظر بگیرد تا بتواند میزان ریسک بالا را برای بدافزارها و میزان ریسک پایین را برای نرم‌افزارها به‌دست آورد. درواقع ما نیاز به تعریف دقیق‌تری از مجوز بحرانی داریم که دو بعد فراوانی مجوز در بدافزار و نرم‌افزار مفید را در نظر بگیرد. به همین دلیل تعریف جدید و دقیق‌تری از مفهوم مجوز بحرانی را ارائه می‌دهیم. یک مجوز بحرانی، مجوزی است که تفاوت فراوانی درخواست آن توسط بدافزارهای اندرویدی نسبت به نرم‌افزارهای مفید و بی‌خطر زیاد باشد.

دو نکته در مورد این تعریف حائز اهمیت است. نخست اینکه استفاده از یک مجوز ممکن است، سبب دسترسی برنامه به داده‌های خصوصی شود؛ ولی تاکنون به هر دلیلی مورد استفاده زیاد از جانب بدافزارنویسان قرار نگرفته است. نکته دوم، یک مجوز برای اینکه متمایزکننده بدافزارها نسبت به نرم‌افزارها شود می‌بایست بیشتر توسط بدافزارها مورداستفاده قرار گیرد تا در نرم‌افزارهای بی‌خطر برای انجام عملیات مفید به کار گرفته شود. به‌کارگیری چنین مجوزی در محاسبه ریسک سبب می‌شود که ریسک بالای امنیتی برای بدافزارها و ریسک امنیتی پایین‌تر برای نرم‌افزارهای مفید محاسبه شود. ما برای تعیین بحرانی‌بودن

[1] Number of Malwares

یک مجوز، فرمول وزن تفاضلی را برای هر مجوز i به‌صورت زیر ارائه می‌دهیم:

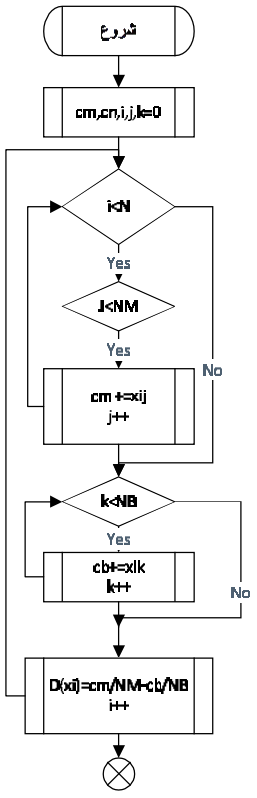
$$D(x_i)=\frac{\sum_{j=1}^{NM}x_{ij}}{NM}-\frac{\sum_{k=1}^{NB}x_{ik}}{NB}=W(x_i)-\frac{\sum_{k=1}^{NB}x_{ik}}{NB} \qquad (۴)$$
$$x_{ij},x_{ik}\in\{0,1\},\forall i\in\{1,2,...,N\}$$

علاوه‌بر نمادهای استفاده‌شدهٔ فرمول‌های (۱) و (۲)، در اینجا NB[1] تعداد نرم‌افزارهای مفید مورد بررسی و x_{ik} وضعیت استفاده از مجوز iام در نرم‌افزار بی‌خطر k ام را نشان می‌دهند. حاصل این فرمول یعنی $D(x_i)$ تفاوت وزنی برای مجوز iام را محاسبه می‌کند.

شکل (۲) روندنمای نحوه محاسبه وزن تفاضلی را برای N مجوز یک سیستم عامل فرضی نشان می‌دهد. در گام ابتدایی متغیرهای لازم مقداردهی اولیه می‌شوند. متغییرهای cn و cm به‌ترتیب به‌منظور شمارش رخداد یک مجوز در نرم‌افزارها و بدافزارها مورداستفاده قرار می‌گیرند. در این روندما رخداد N مجوز در NB برنامه مفید و NM برنامه مخرب شمارش می‌شوند. روندنما از یک حلقهٔ اصلی و دو حلقهٔ داخلی تشکیل شده است. در هر بار اجرای حلقه اصلی، وزن تفاضلی برای مجوز xi محاسبه می‌شود. در حلقه داخلی اول متغییر xij رخداد مجوز xi را در برنامه‌های مفید مشخص می‌کند و در حلقه داخلی دوم متغییر xik رخداد همین مجوز را در برنامه‌های مخرب مشخص می‌کند. همان‌طور که در گام آخر دیده می‌شود از این رخدادها به‌منظور محاسبه وزن تفاضلی هر مجوز طبق فرمول (۴) استفاده می‌شود. با داشتن وزن تفاضلی همه مجوزهای یک سیستم عامل می‌توان میزان ریسک را برای یک برنامه مشکوک ورودی محاسبه کرد.

ما در تحلیل اولیه حدود ۱۳۶ هزار نرم‌افزار مفید و ۸۰۸ بدافزار را بررسی کرده‌ایم. برای مجوزهای مختلف، با توجه به نرم‌افزارها و بدافزارهایی که تحلیل کرده‌ایم، مقادیر مختلفی را با استفاده از فرمول (۴) و روندنمای شکل (۲) به‌دست آورده‌ایم. علاوه‌براین، ما با توجه به وزن تفاضلی به‌دست‌آمده برای مجوزها، آنها را به‌صورت نزولی مرتب و سپس رتبه‌بندی کرده‌ایم. بنابراین هرچه میزان وزن تفاضلی یک مجوز بیشتر باشد، رتبه بالاتری خواهد داشت. از حدود ۱۳۰ مجوز مورداستفاده در اندروید، ۴۹ مجوز وزن بزرگ‌تر از صفر را دارند. ۱۰ مجوز تفاضل وزنی صفر و مابقی تفاضل وزنی منفی دارند. به‌علت محدودیت فضای مقاله، جدول (۱) رتبه، وزن و وزن تفاضلی ۲۵ مجوز نخست را با دقت سه رقم اعشار نشان می‌دهد.

در این جدول ستون‌های سوم و چهارم برای هر مجوز به‌ترتیب با استفاده از فرمول‌های (۱) و (۴) محاسبه شده‌اند. همان‌طورکه در جدول (۱) مشخص است یک مجوز ممکن است وزن W زیادی داشته باشد؛ ولی وزن تفاضلی D کمتری را نسبت به سایر مجوزها داشته باشد و همین سبب شود در رتبه پایین‌تری نسبت به سایر مجوزها قرار بگیرد. به‌عنوان مثال مجوزهای ACCESS_NETWORK_STATE و INTERNET نسبت به سایر مجوزها وزن بالایی را دارند؛ اما وزن تفاضلی آنها نسبت با مجوزهای با رتبهٔ بالاتر خود بسیار کمتر است. چنین مجوزهایی هم در بدافزارها و هم در نرم‌افزارهای بی‌خطر مورداستفاده قرار می‌گیرند؛ اما نسبت به مجوزهای با رتبه بالاتر باعث تمایز نرم‌افزارها از بدافزارها و بالعکس نمی شوند. ما در محاسبه ریسک امنیتی نرم‌افزارها باید به مجوزهایی که وزن تفاضلی بیشتری دارند بیشتر بها دهیم. بنابراین می‌توان از وزن تفاضلی کلیه مجوزهای مورد استفاده در یک نرم‌افزار استفاده کرده و فرمولی ارائه دهیم.



(شکل-۲): روندنمای محاسبه وزن تفاضلی مجوزها
(Figure-2): Differential weight computation Flowchart.

[1] Number of Benign Apps

(جدول-۱): مجوزهای بحرانی اندروید و وزن محاسبه‌شده آنها.

(Table-1): Critical permissions of Android and their computed weights.

رتبه	عنوان مجوز	وزن (W)	وزن تفاضلی (D)
۱	READ_PHONE_STATE	۰/۹۳۱	۰/۶۶۱
۲	READ_SMS	۰/۶۷۹	۰/۶۶۰
۳	ACCESS_WIFI_STATE	۰/۶۷۱	۰/۵۶۵
۴	WRITE_SMS	۰/۵۶۲	۰/۵۴۸
۵	RECEIVE_BOOT_COMPLETED	۰/۵۶۶	۰/۴۸۲
۶	SEND_SMS	۰/۴۸۹	۰/۴۴۸
۷	RECEIVE_SMS	۰/۴۶۰	۰/۴۲۸
۸	WRITE_CONTACTS	۰/۴۱۷	۰/۳۸۱
۹	ACCESS_NETWORK_STATE	۰/۸۰۸	۰/۳۳۲
۱۰	CALL_PHONE	۰/۴۱۵	۰/۳۲۸
۱۱	RESTART_PACKAGES	۰/۳۴۸	۰/۳۲۲
۱۲	WRITE_APN_SETTINGS	۰/۳۲۴	۰/۳۲۱
۱۳	WRITE_EXTERNAL_STORAGE	۰/۶۴۴	۰/۳۰۶
۱۴	READ_CONTACTS	۰/۳۹۲	۰/۳۰۵
۱۵	DISABLE_KEYGUARD	۰/۲۹۰	۰/۲۷۴
۱۶	WAKE_LOCK	۰/۴۱۶	۰/۲۶۸
۱۷	VIBRATE	۰/۴۵۳	۰/۲۶۵
۱۸	READ_LOGS	۰/۲۶۹	۰/۲۴۹
۱۹	CHANGE_WIFI_STATE	۰/۲۶۲	۰/۲۳۵
۲۰	INTERNET	۰/۹۷۸	۰/۲۲۷
۲۱	INSTALL_PACKAGES	۰/۲۱۸	۰/۲۰۹
۲۲	ACCESS_COARSE_LOCATION	۰/۳۹۰	۰/۱۹۹
۲۳	ACCESS_FINE_LOCATION	۰/۳۵۰	۰/۱۳۵
۲۴	ACCESS_LOCATION_EXTRA_COMMANDS	۰/۱۲۶	۰/۰۹۸
۲۵	MOUNT_UNMOUNT_FILESYSTEMS	۰/۱۰۴	۰/۰۹۲

به‌منظور توجه بیشتر به مجوزهای با وزن بالاتر و همچنین کاهش تأثیر اختلاف وزن تفاضلی بین مجوزها، و درنهایت ارائهٔ معیاری مؤثرتر، از رتبه مجوزها استفاده کرده و معیار E-Risk را برای هر مجوز مانند i به‌صورت زیر ارائه می‌دهیم:

$$ER_i = \frac{1}{e^{a \times r_i}} \quad \forall i, r_i \in \{1,2,...,N\} \qquad (۵)$$

در اینجا r_i نشان‌دهندهٔ رتبه مجوز i و e هم عدد نپر است. عدد ثابت a هم طوری انتخاب می‌شود که مجوز با رتبهٔ یک، ریسکی امنیتی نزدیک به یک داشته باشد و همچنین مجوز با پایین‌ترین رتبه هم ریسکی در حدود صفر داشته باشد. بنابراین انتخاب مقدار مناسب برای a سبب می‌شود که فرمول بالا ریسکی هرچند ناچیز را برای مجوزهای با رتبه‌های بسیار پایین محاسبه کند. با توجه با داشتن رتبه‌های یک تا N برای مجوزها، مقدار ۰/۰۵ را برای ثابت a در نظر گرفته‌ایم. ما برای همهٔ N مجوز با استفاده از رتبهٔ آنها و فرمول (۵) مقدار ریسک را محاسبه کرده‌ایم. علت استفاده از همهٔ مجوزها در محاسبه ریسک برنامه‌ها این است که یک مجوز ممکن است در آینده به تنهایی و یا با ترکیب با سایر مجوزها، مورد توجه بدافزارنویسان قرار گیرد.

بنابراین می‌توان برای هر نرم‌افزار A مقدار ریسک امنیتی آن را با استفاده از فرمول (۶) به‌دست آورد:

$$E-RISK(A) = \begin{cases} \dfrac{\sum_{i=1}^{N}(1/e^{a \times r_i})}{\sum_{i=1}^{N} x_i} & \sum_{i=1}^{N} x > 0 \\ 0 & \sum_{i=1}^{N} x = 0 \end{cases} \qquad (۶)$$

$$x_i \in \{0,1\}, \forall i, r_i \in \{1,2,...,N\}, a = 0.05$$

در اینجا r_i رتبهٔ مجوز i درخواست‌شده توسط A و x_i نشان‌دهندهٔ درخواست یا عدم درخواست مجوز i توسط نرم‌افزار مورد بررسی A است. مشابه فرمول (۲) مخرج کسر به‌منظور نرمال‌سازی مقدار حاصل، اضافه شده است. بدیهی است که اگر نرم‌افزار A هیچ مجوزی را درخواست نکند میزان ریسک آن مطابق فرمول (۶) صفر خواهد بود. مشابه فرمول (۳) از وزن‌های به‌دست‌آمدهٔ نرم‌افزارهای نصب‌شده و نصب‌نشدهٔ یک سیستم با استفاده از فرمول (۶) می‌توان به‌منظور محاسبه ریسک کلی یک سیستم اندرویدی مانند S استفاده کرد. به‌منظور خوانایی بیشتر روابط ریاضی (۱) تا (۶)، نمادهای استفاده‌شده در این فرمول‌ها در جدول (۲) خلاصه شده اند.

(جدول-۲): نمادهای استفاده شده و مفهوم آنها

(Table-2): Used symbols and their meanings

نام	مفهوم
NB	تعداد برنامه مفید(Number of Benign apps)
NM	تعداد برنامه مخرب(Number of Malwares)
$W(x_i)$	وزن مجوز x_i در بدافزارها
$D(x_i)$	وزن تفاضلی مجوز x_i
x_{ij}	وضعیت درخواست مجوز i در برنامه مفید j
x_{ik}	وضعیت درخواست مجوز i در بدافزار k
E-Risk	ریسک دقیق (Exact Risk)
RValue	مقدار ریسک (Risk Value)
n_i	تعداد برنامه های استفاده کننده از مجوز i

داشتن ریسک لزوماً نشان‌دهندهٔ مخرب‌بودن یک نرم‌افزار نیست؛ اما ریسک بالای یک نرم‌افزار هشداری برای کاربر محسوب می‌شود و نشان می‌دهد که کاربر باید در استفاده از این نرم‌افزار تردید کرده و نسبت به استفاده از نرم‌افزارهای مشابه ولی با ریسک کمتر اقدام کند. ریسک بالای سیستم نیز نشان می‌دهد که تعداد نرم‌افزارهای خطرناک و مشکوک سیستم زیاد بوده و کاربر می‌بایست

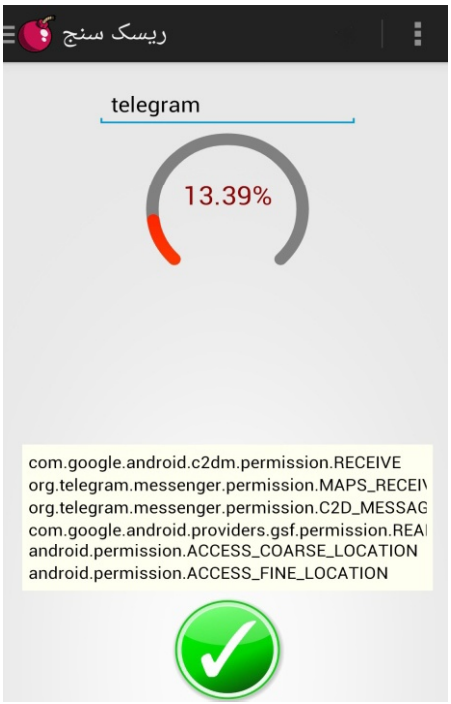
نسبت به حذف و یا جایگزینی برخی از نرم‌افزارهای با ریسک بالا که با استفاده از معیار فرمول (۶) قابل شناسایی هستند، اقدام کند. بنابراین، با استفاده از این معیار رؤیت مجوزها تنها مربوط به زمان نصب آنها نمی‌شود و کاربر می‌تواند پایش امنیتی سیستم خود را در هر زمان و به شکل ملموس‌تری انجام دهد. معیارهای ریسک امنیتی همچنین می‌توانند به‌عنوان یک پیش‌پردازش برای بررسی یک نرم‌افزار، مورداستفاده قرار گیرند. به این صورت که در مراحل بعدی توسط روش‌های تحلیل ایستای کد و/یا تحلیل‌های رفتاری پویا، مخرب یا مفیدبودن یک نرم‌افزار را مشخص کرد.

ما با استفاده از معیار پیشنهادی، ابزاری به‌نام ریسک‌سنج را با زبان جاوا در محیط *Eclipse* توسعه داده‌ایم که می‌تواند ریسک امنیتی یک نرم‌افزار انتخابی و همچنین ریسک کلی سیستم را محاسبه کند. این نرم‌افزار دارای طبقه‌هایی به‌منظور محاسبه ریسک و پیش‌پردازش یک برنامه در سیستم عامل اندروید است. قبل از اجرای ابزار ریسک‌سنج مقدار ریسک مجوزها مطابق رابطه (۵) محاسبه می‌شود. از این مقادیر به‌منظور محاسبه ریسک هر برنامه انتخابی مطابق فرمول (۶) استفاده می‌شود. با توجه به اینکه همواره بدافزارها و نرم‌افزارهای جدیدی برای اندروید ارائه می‌شود، این ابزار می‌تواند به‌صورت دوره‌ای مطابق روندنمای شکل (۲) ریسک مجوزها را دوباره محاسبه و به‌روزرسانی کند. فهرست مجوزهای هر نرم‌افزار اندرویدی از طریق فایل AndroidManifest.xml در پکیج apk مربوط به آن نرم‌افزار قابل دسترسی است. شکل (۳) صفحه نخست این نرم‌افزار را در حالتی نشان می‌دهد که ریسک امنیتی را برای برنامه انتخابی تلگرام محاسبه کرده است.

همان‌طورکه در شکل (۳) دیده می‌شود، مقدار ریسک در اینجا به‌صورت درصد محاسبه شده است، به این معنی که برنامهٔ انتخابی با توجه به مجموع ریسک موجود در مجوزهای مورد استفاده، درصد مشخص‌شده ریسک را دارد. در بخش پایینی این فرم، مجوزهای مورد استفاده این نرم‌افزار فهرست شده‌اند. این ابزار همچنین قابلیت محاسبه ریسک برای کل نرم‌افزارهای موجود یا نصب‌شده در یک سیستم را نیز دارد.

۵- ارزیابی و مقایسه

به‌منظور استخراج وزن ها و ارزیابی معیارهای ارائه‌شده، از



(شکل-۳): شمای نرم‌افزار ریسک سنج در حال محاسبه ریسک یک برنامه

(Figure-3): A Picture of RiskMeter software during risk computation for an application

مجموعه داده‌های مختلفی استفاده کرده‌ایم. در این بخش ابتدا مجموعه داده‌های مورداستفاده در ارزیابی‌های خود را معرفی کرده و سپس به گزارش آزمایش‌های انجام‌شده و تحلیل نتایج به‌دست‌آمده می‌پردازیم.

۵-۱- مجموعه داده‌های مورداستفاده در ارزیابی

به‌منظور ارزیابی معیارهای پیشنهادی و مقایسه آن با معیارهای قبلی از چهار مجموعه داده استفاده کرده‌ایم. جدول (۳) خلاصه‌ای از اطلاعات مربوط به این مجموعه داده‌ها را نشان می‌دهد. دو مجموعه دادهٔ نخست جزء مجموعه داده‌های موجود بوده که در پژوهش‌های قبلی امنیتی مورداستفاده قرار گرفته‌اند. دو مجموعه دادهٔ دوم را گردآوری و پیش‌پردازش کرده‌ایم.

(جدول-۳): مشخصات مجموعه داده‌های مورداستفاده در ارزیابی‌ها

(Table-3): Properties of the used datasets in the evaluations

نام مجموعه داده	تعداد برنامه	تعداد مجوز
Market2012	۱۳۶۵۳۴	۱۲۲
Malwares	۸۰۸	۱۲۲
بازار	۱۹۲۶	۱۳۵
بدافزار جدید	۱۰۱۴	۱۳۵

به‌منظور ارزیابی معیار ارائه‌شده بر روی نرم‌افزارهای مفید، از مجموعه داده Market2012 ارائه‌شده در [16] استفاده کرده‌ایم. این مجموعه داده شامل حدود ۱۳۶ هزار نرم‌افزار مفید مربوط به سایت فروش نرم‌افزار اندروید شرکت گوگل در سال ۲۰۱۲ میلادی است. مجموعه داده دوم مربوط به بدافزارهای اندروید [26] است که از آن برای ارزیابی و مقایسه معیارها بر روی بدافزارها استفاده شده است. دو مجموعه دادهٔ بعدی جدید بوده و ما آنها را گردآوری و پیش‌پردازش کرده‌ایم. مجموعه دادهٔ بازار شامل ۱۹۲۶ برنامه مفید بر روی سایت بازار است که در سال‌های ۱۳۹۴ و ۱۳۹۵ به‌دست آمده‌اند. ما فرض کرده‌ایم که این برنامه‌ها همگی برنامه‌های مفید هستند؛ زیرا توسط یک فروشگاه معتبر داخلی ارائه شده‌اند. مجموعه دادهٔ بدافزار جدید شامل بدافزارهایی می‌شوند که در سال‌های ۲۰۱۴ تا ۲۰۱۶ میلادی گردآوری شده‌اند. ما داده‌های مربوط به دو مجموعه داده آخر را پس از گردآوری، به‌صورت جداگانه پیش‌پردازش کرده و فهرست مجوزهای مورداستفاده در آنها را استخراج کرده‌ایم.

ما معیارها را هم بر روی بدافزارها و هم بر روی نرم‌افزارهای مفید جدول (۳) محاسبه کرده‌ایم. همان‌طورکه درقبل گفته شد، یک معیار خوب سنجش ریسک می‌بایست برای بدافزارها عدد بالایی تولید کند و برای نرم‌افزارهای مفید و غیر مخرب عدد کمی به‌دست آورد. البته به‌دست‌آوردن عدد بالا برای یک نرم‌افزار لزوماً به معنای مخرب بودن آن نیست.

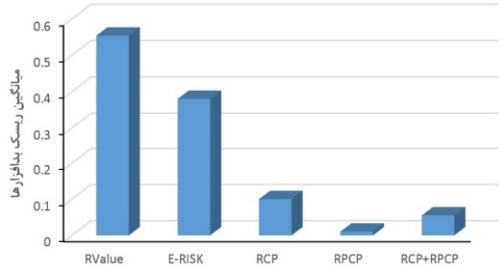
۲-۵- آزمایش‌ها

ما در ارزیابی‌های خود از مجموعه داده‌های جدول (۳) استفاده کرده‌ایم تا عمومیت ارزیابی‌های خود را بر روی برنامه‌های مفید و مخرب قدیمی و جدید، نشان دهیم. پس از تحلیل بدافزارهای موجود و محاسبه دقیق وزن‌ها، معیارهای پیشنهادی خود یعنی *RValue* و *E-Risk* را با سه معیار قبلی یعنی *RCP*، *RPCP* و ترکیب این دو (RCP&RPCP)، مقایسه کرده‌ایم. نتیجهٔ مقایسه برای بدافزارهای شناخته‌شده در مجموعه داده Malwares در شکل (۴) نشان داده شده است. در این شکل محور افقی شماره نرم‌افزار و محور عمودی نشان‌دهنده میزان ریسک محاسبه‌شده توسط معیارهای مختلف است. به‌منظور صاف‌کردن و قابل‌تحلیل‌بودن نمودار، در محور عمودی میزان میانگین حرکتی[1] ریسک با پنجره‌ای به اندازه بیست محاسبه و نشان داده شده است. همان‌طورکه در شکل پیداست مقادیر محاسبه‌شده ریسک توسط معیارهای پیشنهادی *RValue* و *E-Risk* در همهٔ موارد بالاتر از معیارهای قبلی است.

(شکل-۴): مقایسه مقدار ریسک به‌دست‌آمده توسط معیارهای پیشنهادی با سایر معیارها برای بدافزارهای شناخته‌شده در مجموعه داده Malwares

(Figure-4): Comparing computed risk for known malwares of Malwares dataset using various metrics.

علت این برتری، درنظر گرفتن وزن هر مجوز در محاسبه مقدار ریسک امنیتی است. همچنین در شکل (۴) دیده می‌شود که معیار پیشنهادی دوم یعنی *E-Risk* نسبت به *RValue* در رتبه دوم قرار گرفته است. معیار پیشنهادی *E-Risk* در محاسبات خود مجوزهایی که در بدافزارها با درصد بیشتری نسبت به نرم‌افزارهای مفید استفاده می‌شوند، مدنظر قرار می‌دهد.



(شکل-۵): مقایسه میانگین ریسک به‌دست‌آمده بر روی همهٔ‌بدافزارهای مجموعه داده Malwares توسط همه معیارها.

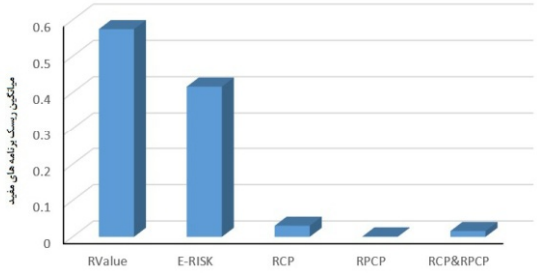
(Figure-5): Comparing average risk values of all malwares within Malware dataset computed by all metrics

در آزمایش بعدی میانگین کلی ریسک امنیتی همهٔ بدافزار را توسط هر پنج معیار به‌دست آوردیم. شکل (۵) نتیجه این آزمایش را نشان می‌دهد. در این شکل، هر ستون مربوط به یک معیار است و محور عمودی میانگین به‌دست‌آمده ریسک توسط هر معیار را نشان می‌دهد. همان‌طورکه از شکل پیداست، معیارهای پیشنهادی مقدار میانگین ریسک بسیار بیشتری را روی همهٔ بدافزارها به‌دست

[1] Moving Average

می‌دهند که نشان‌دهندهٔ عملکرد بهتر این معیارها در این زمینه است. به ویژه معیار پیشنهادی *RValue* بیشترین میانگین ریسک را برای بدافزارها محاسبه کرده است.

در آزمایش بعدی همین میانگین ریسک را برای همهٔ نرم‌افزارهای مفید مجموعه Market2012 محاسبه کرده‌ایم. اگرچه اینها همگی نرم‌افزار مفید هستند؛ ولی برای ارائه‌دادن توانایی‌های خود و استفاده از قابلیت‌های مختلف دستگاه، نیاز به درخواست مجوز دارند. بنابراین برای نرم‌افزارهای مفید نیز میزان ریسک امنیتی با توجه به کم و کیف درخواست مجوزها قابل محاسبه است. همان‌طورکه درقبل گفته شده یکی از ویژگی‌های یک معیار خوب این است که توسط آن بتوان مقدار کمی برای ریسک یک نرم‌افزار بی‌خطر محاسبه کرد؛ البته به شرطی که همین معیار بتواند ریسک بالا را برای یک نرم‌افزار مخرب به‌دست آورد. میانگین ریسک محاسبه‌شده توسط هر معیار بر روی کل این نرم‌افزارها در شکل (۵) نشان داده شده است. همانند شکل قبلی محور افقی معیارها و محور عمودی میانگین ریسک امنیتی است.

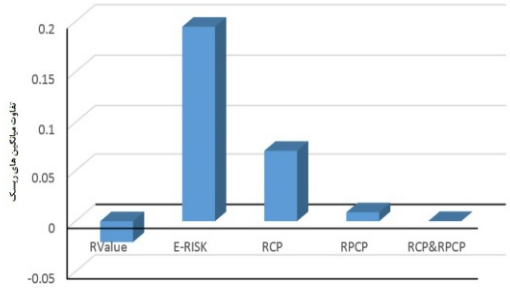


(شکل-۶): مقایسه میانگین ریسک به‌دست‌آمده بر روی همهٔ نرم‌افزارهای مفید مجموعه داده Market2012 توسط همهٔ معیارها.

(Figure-6): Comparing average risk values of all benign apps within Market2012 dataset computed by all metrics

معیار پیشنهادی *E-Risk* نسبت به *RValue* عملکرد بهتری داشته زیرا برای نرم‌افزارهای مفید مقدار ریسک امنیتی کمتری را به‌دست آورده است. همان‌طورکه در شکل (۶) دیده می‌شود، کم‌ترین میانگین مقدار ریسک توسط معیار *RPCP* به‌دست آمده است. توسط معیار دیگر *RCP* و ترکیب این دو یعنی *RCP&RPCP* نیز نسبت به معیارهای پیشنهادی ما مقادیر متوسط کم‌تری به‌دست آمده‌اند؛ اما این مسئله دلیل برتری این معیارها نسبت به معیار پیشنهادی ما نیست؛ چون همان‌طورکه در آزمایش‌های قبلی دیدیم، توسط این معیارها، برای بدافزارها هم مقادیر کم‌تری به‌دست آمد. یعنی این معیارها هم برای بدافزارها و هم برای نرم‌افزارها مقادیر کم ریسک را محاسبه کرده‌اند. از طرف دیگر برای تشخیص بدافزار از نرم‌افزار، معیاری کاربردی است که تفاوت بیشتری را تولید کند. به عبارت دیگر توسط این معیار باید بتوان برای بدافزارها نسبت به نرم‌افزارها مقادیر بیشتری را به‌دست آورد. به همین دلیل در شکل بعدی این تفاوت را نشان داده‌ایم.

در شکل (۷) تفاوت میانگین محاسبه‌شده برای بدافزارها و نرم‌افزارها توسط هر معیار به تصویر کشیده شده است. به عبارت دقیق‌تر، محور عمودی این شکل حاصل تفریق میزان میانگین ریسک محاسبه‌شده برای بدافزارها منهای مقدار متوسط محاسبه‌شده برای نرم‌افزارهای مفید را نشان می‌دهد.



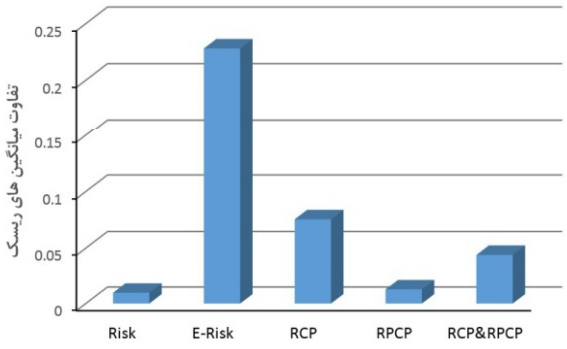
(شکل-۷): مقایسه تفاوت میانگین ریسک به‌دست‌آمده بر روی بدافزارها و نرم‌افزارها توسط همهٔ معیارها(مجموعه داده‌های Market2012 و Malwares).

(Figure-7): Comparing subtraction of computed risks on malwares and benign apps for all metrics (Market2012 and Malwares datasets)

همان‌طورکه در این شکل دیده می‌شود، معیار پیشنهادی دوم ما یعنی *E-Risk* بیشترین تفاوت را نشان می‌دهد. یعنی این معیار برای تمایز بدافزار از نرم‌افزار نسبت به سایر معیارها بهتر عمل می‌کند. معیار *RPC* در رده دوم است و معیار پیشنهادی نخست ما یعنی *RValue* در این زمینه بسیار بد عمل کرده است. علت این مسئله بالابردن وزن تعدادی از مجوزها بدون توجه به وضعیت استفاده از این مجوزها در نرم‌افزارهای مفید است؛ زیرا برخی از این معیارها، هم در نرم‌افزارها و هم در بدافزارها به وفور استفاده می‌شوند و بالابردن وزن آنها سبب می‌شود که ریسک محاسبه شده برای نرم‌افزارها نیز بسیار بیشتر محاسبه شود. درواقع مزیت معیار پیشنهادی *E-Risk* این است که سوابق آماری استفاده از مجوزها را هم در بدافزارها و هم در نرم‌افزارهای مفید مورد توجه قرار می‌دهد و سبب می‌شود که مقدار محاسبه‌شده ریسک نسبی برای بدافزارها نسبت به نرم‌افزارهای مفید به شکل قابل توجهی بیشتر شود. استفاده

از ایـن قابلیـت سـبب می‌شـود کـه در انتخـاب برنامـه‌های کم‌خطر مؤثرتر عمل کنیم.

ما همهٔ آزمایش‌های بالا را بر روی مجموعه داده‌هـای نرم‌افزار و بدافزار خود نیز تکرار کرده‌ایم. یعنی یک‌بار توسـط پنج معیار مورد مقایسه میانگین ریسک را روی مجموعه داده بازار و یک‌بار دیگر جداگانه روی مجموعه‌داده بـدافزار جدیـد مشابه آزمایش‌های بالا محاسبه کرده‌ایم. همان‌طورکـه گفتـه شد تفاوت میانگین‌های ریسک میـان بـدافزارها و نرم‌افزارهـا مشخص‌کننده قابلیت یک معیار در تخصیص ریسک بـالا بـه بدافزارها و ریسک پایین به نرم‌افزارها اسـت. بنـابراین مـا در اینجا تنهـا تفـاوت میانگین‌هـای ریسـک محاسبه‌شـده را در شکل (۸) نمایش داده‌ایم. همان‌طورکه در این شکل نیز دیده می‌شود، بازهم معیار پیشنهادی *E-Risk* بیشـترین کـارایی را دارد. درواقع داده‌های جدید دربرگیرندهٔ الگوهای متفاوتی در اسـتفاده از مجوزهـا در برنامـه‌های مفیـد و مخـرب اسـت. بنابراین معیار پیشنهادی عمومیت لازم را دارد یعنـی در هـر دو نوع مجموعه داده بهتر عمل می‌کند.



(شکل-۸): مقایسه تفاوت میانگین‌های ریسک به‌دست‌آمده روی بدافزارها و نرم‌افزارها توسط همهٔ معیارها(مجموعه داده‌های بازار و بدافزار جدید)

(Figure-8): Comparing subtraction of computed risks on malwares and benign apps for all metrics (Bazar and New Malwares datasets)

با توجه به آزمایش‌ها و بـا مجمـوع کـل مقایسـه‌های انجام‌شده می‌توان نتیجه گرفت که معیار پیشـنهادی دوم مـا یعنی *E-Risk* مطابق فرمول (۶) و روال محاسبهٔ مربوطـه، در مجموع آزمایش‌ها بهترین عملکرد را داشته و به‌عنوان معیـار پیشنهادی نهایی این پژوهش به دنیای اندروید ارائه می‌شود. درواقع این معیار نسبت به سایر معیارهای بررسی‌شده مقـدار ریسک امنیتی را دقیق‌تر و واقعی‌تر محاسبه می‌کند.

۶- جمع بندی و نتیجه گیری

شناسایی بدافزارهای دستگاه‌های همـراه نیازمنـد معیارهـای امنیتـی دقیـق‌تری اسـت. ایـن معیارهـا می‌تواننـد در آنتی‌ویروس‌های مربوط به سیستم عامل‌های ایـن دسـتگاه‌ها برای شناسایی اولیه نرم‌افزارهای مخـرب و اعـلام هشـدار در استفاده از آنها به کار روند؛ اما برای تشـخیص دقیـق‌تر یـک بدافزار لازم است از روش‌های مکملی ماننـد تحلیـل ایسـتا و پویای کد و همچنین روش‌های داده‌کاوی اسـتفاده کـرد. در این مقاله مـا معیـار جدیـدی را بـه‌منظور محاسـبه ریسـک نرم‌افزارهای مشکوک ارائه دادیم کـه نسـبت بـه سـه معیـار ارائه‌شده قبلی بهتر عمل می‌کند. آزمایش‌ها بر روی داده‌های واقعی نشان دادند کـه معیـار پیشـنهادی بـرای بـدافزارهای شناخته‌شده نسبت به نرم‌افزارهای مفید، مقدار ریسک بسیار بیشـتری را به‌دسـت می‌دهـد. معیـار پیشـنهادی در فروشگاه‌های داخلی برنامه‌های اندروید بـه‌منظور فیلترکـردن برنامه‌های مخرب ارسال شده توسط نفـوذگران، می‌توانـد بـه کارگرفته شود. علاوه‌براین، نرم‌افزار ریسک‌سـنج در محاسـبه ریسک دقیق برنامه‌های ناشناخته و مقایسه آنها بـه کـاربران می‌تواند کمک کند. ما در آینـده قصـد داریـم بـا اسـتفاده از روش‌های داده‌کاوی و تحلیل ایستای کد، دسته‌بند[1] دقیقـی را به‌منظور شناسایی بدافزارهای روز صفر در اندروید طراحی کنیم.

۷-مراجع 7-References

[1] لسانی، فاطمه سادات، فتوحی، قزوینی فرانک، دیانت، روح الله، "لب‌خوانی: روش جدید احراز هویت در برنامه‌های کاربردی گوشی‌های تلفن همراه اندروید"، *مجله پردازش علائم و داده‌ها*، جلد ۱۴، شماره ۱، صفحات ۳-۱۴، ۱۳۹۶.

[1] F. Sadat Lesani, F. Fotouhi Ghazvini, and R. Dianat, " Lip Reading: a New authentication method in Android mobile phone's applications," Journal of Signal and Data Processing, vol. 14, no. 1, pp. 3-14, 2017.

[2] A. Armando, A. Merlo, and L. Verderame, "Security considerations related to the use of mobile devices in the operation of critical infrastructures," *International Journal of Critical Infrastructure Protection*, vol. 7, no. 4, pp. 247–256, 2014.

[3] Y. Au, K. Wain, Y. F. Zhou, Z. Huang, and D. Lie, "PScout : Analyzing the Android Permission

[1] Classifier

comprehension, and behavior," in *Proc. of SOUPS*, 2012, pp. 1–14.

[15] C. S. Gates, J. Chen, N. Li, and R. W. Proctor, "Effective risk communication for Android apps," *IEEE Transactions on Dependable and Secure Computing*, vol. 11, no. 3, pp. 252–265, 2014.

[16] C. S. Gates, N. Li, H. Peng, B. Sarma, Y. Qi, R. Potharaju, C. Nita-Rotaru, and I. Molloy, "Generating summary risk scores for mobile applications," *IEEE Transactions on Dependable and Secure Computing*, vol. 11, no. 3, pp. 238–251, 2014.

[17] P. G. Kelley, S. Consolvo, L. F. Cranor, J. Jung, N. Sadeh, and D. Wetherall, "A conundrum of permissions: Installing applications on an android smartphone," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2012, vol. 7398 LNCS, pp. 68–79.

[18] P. G. Kelley, L. F. Cranor, and N. Sadeh, "Privacy as part of the app decision-making process," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, 2013, p. 3393.

[19] H. Peng, C. Gates, B. Sarma, N. Li, Y. Qi, R. Potharaju, C. Nita-Rotaru, I. Molloy, and I. Molloy, "Using Probabilistic Generative Models for Ranking Risks of Android Apps," *Proceedings of the 2012 ACM Conference on Computer and Communications Security*, pp. 241–252, 2012.

[20] K.Rieck, T. Holz, C. Willems, P. Dussel, and P. Laskov, " Learning and classification of malware behavior, " in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2008, vol. 5137 LNCS, pp. 108-125.

[21] B. Sarma, N. Li, C. Gates, R. Potharaju, C. Nita-rotaru, and I. Molloy, "Android Permissions: A Perspective Combining Risks and Benefits," in *Symposium on Access Control Models and Technologies (SACMAT)*, 2012, pp. 13–22.

[22] A. D. Schmidt, R. Bye, H. G. Schmidt, J. Clausen, O. Kiraz, K. A. Yüksel, S. A. Camtepe, and S. Albayrak, "Static analysis of executables for collaborative malware detection on andr-oid," in *IEEE International Conference on Communications*, 2009.

[23] A. Shabtai and Y. Elovici, "Applying behavioral detection on android-based devices," in *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*, 2010, vol. 48 LNICST, pp. 235–249.

Specification," in *CCS '12 Proceedings of the 2012 ACM conference on Computer and communications security*, 2012, pp. 217–228.

[4] D. Barrera, H. G. üne ş Kayacik, P. C. van Oorschot, and A. Somayaji, "A methodology for empirical analysis of permission-based security models and its application to android," in *Proceedings of the 17th ACM conference on Computer and communications security - CCS '10*, 2010, p. 73.

[5] I. Burguera, U. Zurutuza, and S. Nadjm-Tehrani, "Crowdroid: Behavior-Based Malware Detection System for Android," in *Proceedings of the 1st ACM workshop on Security and privacy in smartphones and mobile devices - SPSM '11*, 2011, p. 15.

[6] L. Cen, C. S. Gates, L. Si, and N. Li, "A Probabilistic Discriminative Model for Android Malware Detection with Decompiled Source Code," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 4, pp. 400–412, 2015.

[7] S. Chakradeo, B. Reaves, and W. Enck, "MAST: Triage for Market-scale Mobile Malware Analysis," in *ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec)*, 2013, pp. 13–24.

[8] E. Chin, A. P. Felt, V. Sekar, and D. Wagner, "Measuring user confidence in smartphone security and privacy," in *Proceedings of the Eighth Symposium on Usable Privacy and Security - SOUPS '12*, 2012, p. 1.

[9] M. Christodorescu, S. Jha, and C. Kruegel, "Mining specifications of malicious behavior," in *Proceedings of the 1st conference on India software engineering conference - ISEC '08*, 2008, p. 5.

[10] A. Desnos, "Android: Static analysis using similarity distance," in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2011, no. X, pp. 5394–5403.

[11] W. Enck, D. Octeau, P. McDaniel, and S. Chaudhuri, "A Study of Android Application Security.," in *USENIX Security*, 2011, vol. 39, no. August, pp. 21–21.

[12] W. Enck, M. Ongtang, and P. McDaniel, "On lightweight mobile phone application certification," in *Proceedings of the 16th ACM conference on Computer and communications security - CCS '09*, 2009, p. 235.

[13] A. Felt, K. Greenwood, and D. Wagner, "The effectiveness of application permissions," in *WebApps '11: 2nd USENIX Conference on Web Application Development*, 2011, pp. 75–86.

[14] A. Felt, E. Ha, S. Egelman, and A. Haney, "Android permissions: User attention,

[24] B. Shebaro, O. Oluwatimi, and E. Bertino, "Context-based access control systems for mobile devices," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 2, pp. 150–163, 2015.

[25] Z. Tu, O. Turel, Y. Yuan, and N. Archer, "Learning to cope with information security risks regarding mobile device loss or theft: An empirical examination," *Information and Management*, vol. 52, no. 4, pp. 506–517, 2015.

[26] Y. Zhou and X. Jiang, "Dissecting Android malware: Characterization and evolution," in *Proceedings - IEEE Symposium on Security and Privacy*, 2012, pp. 95–109.

[27] Y. Zhou, Z. Wang, W. Zhou, and X. Jiang, "Hey, You, Get Off of My Market: Detecting Malicious Apps in Official and Alternative Android Markets," in *Proceedings of the 19th Annual Network and Distributed System Security Symposium*, 2012, no. 2, pp. 5–8.

[28] H. Zhu, H. Xiong, Y. Ge, and E. Chen, "Discovery of ranking fraud for mobile apps," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 1, pp. 74–87, 2015.

محمود دی پیر مدرک دکترای خود را در رشتهٔ کامپیوتر-سامانه‌های نرم‌افزاری و مدرک کارشناسی ارشد خود را در رشتهٔ کامپیوتر نرم‌افزار هر دو از دانشگاه شیراز و مقطع کارشناسی خود را نیز در همین رشته از دانشگاه هوایی شهید ستاری دریافت کرده است. زمینه‌های پژوهشی ایشان شامل داده‌کاوی و امنیت فضای سایبر و دارای مقالات متعددی در مجلات و کنفرانس‌های معتبر ملی و بین‌المللی است. ایشان در پروژه‌های پژوهشی و صنعتی متعددی در زمینه نرم‌افزار، امنیت شبکه و داده‌کاوی به‌عنوان مشاور، مجری و همکار مشارکت داشته است.

نشانی رایانامه ایشان عبارت است از:

mdeypir@ssau.ac.ir