ISSN: 2745-3673
http://www.Rcsj.ir

مجله علمی پژوهش در علوم رایانه (سال ششم)
شماره ۲۱ / بهار ۱۴۰۰ / ص ۲۸- ۱۴

شناسایی و طبقه بندی بدافزارها با استفاده از الگوریتم های یادگیری ماشین (بدون نظارت)

معصومه مرادی [۱]، مجتبی صالحی [۲]

[۱] گروه مهندسی کامپیوتر، واحد خرم آباد، دانشگاه آزاد اسلامی، خرم آباد، ایران.

[۲] گروه مهندسی کامپیوتر، واحد خرم آباد، دانشگاه آزاد اسلامی، خرم آباد، ایران.

نام نویسنده مسئول:
مجتبی صالحی

تاریخ دریافت: ۱۴۰۰/۱/۴
تاریخ پذیرش: ۱۴۰۰/۳/۱۳

چکیده

ما در این مقاله به تشخیص و طبقه بندی (بدون نظارت) بدافزارها که امری مهم در امنیت سیستم‌های کامپیوتری است، پرداخته ایم. در تحقیق حاضر از داده‌های موجود در پایگاه داده Malimg Dataset استفاده شد. این دیتاها شامل تصاویر بدافزار است که در این مقاله هشت خانواده از بدافزارها بررسی شد. از هر تصویر هیستوگرام و ماتریس هم وقوعی (در چهار زاویه صفر، ۴۵، ۹۰ و ۱۳۵ درجه) استخراج شد سپس عملیات استخراج ویژگی انجام شد. در مجموع از هر تصویر بدافزار ۲۰ ویژگی از ماتریس هم وقوعی و ۵ ویژگی از هیستوگرام استخراج شد که از این ویژگی‌های جهت مدلسازی روشهای تششخیص نوع بدافزار استفاده شد. در این مقاله از دو مدل جهت خوشه بندی نوع بدافزار استفاده شد. دقت تشخیص نوع بدافزار توسط روش‌های شبکه عصبی مصنوعی و الگوریتم k-means به ترتیب۹۶.۴۵ ٪ و ۸۷.۳۵ ٪ به دست آمد که بهترین مدل برای تشخیص نوع بدافزار مربوط به الگوریتم **k-means** بود.

واژگان کلیدی: بدافزار، تشخیص، یادگیری ماشین، پردازش تصویر.

مجلـه‌ی علمی پژوهش در علوم رایانه ، شماره ۲۱، بهار ۱۴۰۰، ص ۲۸- ۱۴

ISSN: 2745-3673
http://www.Rcsj.ir

مقدمه

امروزه، شناسایی نرم‌افزارهای مخرب عمدتاً با الگوریتم ابتکاری و روش‌های مبتنی بر امضا انجام می‌شود که تلاش می‌کنند خود را همراه با تکامل بدافزارها به روز نگه دارند. روش‌های مبتنی بر امضا برای نرم‌افزارهای آنتی ویروس برای دهه‌ها زیاد مورد استفاده قرار نگرفته‌اند. امضای بدافزار، یک الگوریتم یا هَش است که به طور منحصر به فرد یک ویروس خاص را شناسایی می‌کند. درحالی که شناسایی یک ویروس خاص مفید است، شناسایی یک خانواده ویروس از طریق یک امضای گروهی سریع‌تر است. (Ronen et al., 2018) بدافزار[1]، نرم افزاری است که نیت خرابکارانه و یا اثراتی تخریبی دارد. این نرم افزارها طیف وسیعی از خطرات و تهدیدات کامپیوتری، از قبیل ویروس‌ها، کرم‌ها، تروجان‌ها و نرم افزارهای جاسوسی را در بر می‌گیرند(Kramer and Bradfield, 2010). بدافزارهای با اهداف مختلفی توسعه پیدا کرده‌اند از جاسوسی اینترنتی گرفته تا سرقت اطلاعات کاربران از جمله اهداف توسعه آنها است(Saxe Berlin, 2015). یکی از اصلی‌ترین میزبانان بدافزارها فایل‌های اجرایی هستند، به همین منظور شناسایی و تشریح بدافزارها از فایل اجرایی، در مباحث امنیت کامپیوتری امری بسیار حیاتی و حائز اهمیت است (شیرازی و فرشچی، ۱۳۹۳). رشد سریع بدافزار باعث ایجاد تهدیدهای بسیاری در حوزه امنیت اطلاعات شده است؛ بنابراین، مراکز دفاع سایبری اهمیت زیادی در بسیاری از کشورها دارد. (رنجی و پارسا، ۱۳۹۷). بدافزار می‌تواند به دسته‌های آگهی افزار[2]، جاسوس افزار[3]، ویروس، کرم، تروجان، روت کیت[4]، درِ پشتی[5]، باج افزار[6] و بات‌ها تقسیم شود که با توجه هدف‌شان، می‌توانند به صورت تکی یا در کنار هم وارد سیستم شوند.

پیشینه تحقیق

تحلیل تکاملی بدافزار بر اساس سوء استفاده‌های تزریق کد توسط Ma et al., (2006) بررسی شده است. این کار به طور انحصاری شل کد[7] استخراج شده از نمونه‌های بدافزار را بررسی می‌کند. شل کد با استفاده از تکنیک‌های خوشه‌بندی برای تعیین روابط بین نمونه‌ها تحلیل می‌شود. این کار یک فیلوژنی از چنین سوء استفاده‌هایی ارائه می‌کند، که مقدار قابل توجهی اشتراک کد را نشان می‌دهد. به علاوه، مؤلفان قادر به شناسایی تغییرات ظریف در خانواده‌ها هستند. یکی از محدودیت‌های این کار این است که آن فقط شل کد را بررسی می‌کند، که می‌تواند به عنوان فاز اولیه یک حمله تزریق کد در نظر گرفته شود. در نتیجه، تکامل جنبه‌های دیگر حملات در نظر گرفته نمی‌شوند.

Gupta et al., (2009) ویژگی‌های ارثی بدافزا را بررسی می‌کنند، ویژگی‌هایی که با استفاده از یک تکنیک هرس گراف به دست آمدند. یکی از نقاط قوت این کار این است که آن بر اساس یک مجموعه داده بدافزار بزرگ و متنوع در مدت زمان دو دهه است. این تحقیق ادعا می‌کند که خانواده‌های زیادی با ویژگی‌های خاص «به ارث برده شده» از دیگر خانواده‌ها را آشکار می‌کند. با این وجود، کاملاً آشکار نیست که تمام این ویژگی‌های «به ارث برده شده» در واقع به ارث برده شدند، و ممکن است از منابع دیگر نشئت گرفته باشند یا به طور مستقل توسعه یافته باشند. و بر خلاف تکنیک بررسی شده در این مقاله، تحلیل مبتنی بر گراف توسط Gupta et al., (2009) به «بررسی غیر خودکار وسیع» نیاز دارد.

Mercaldo et al., (2018) تغییرات در کیفیت بدافزار اندرویدی را در مقایسه با نیک افزار اندرویدی بررسی کردند. مؤلفان ویژگی‌های متنوعی را استخراج کردند و روندها را بر اساس معیارهای استاندارد کیفیت نرم‌افزار تعیین کردند. مؤلفان متوجه شدند که روندها در بدافزار و نیک افزار مشابه هستند، که نشان می‌دهد برنامه نویسان بدافزار تلاش می‌کنند نرم‌افزار خود را به اندازه برنامه نویسان نیک افزار بهبود دهند.

---

[1] Malware
[2] Adware
[3] Spyware
[4] rootkit
[5] backdoor
[6] ransomware
[7] shellcode

ISSN: 2745-3673
http://www.Rcsj.ir

Ouellette et al., (2013) مسئله شناسایی گونه‌های مختلف بدافزاری را بررسی کردند. توانایی شناسایی چنین گونه‌های جدید و پیچیده‌ای، یک مسئله تشخیص تکامل است. این کار بر مجموعه ویژگی گسترده متکی است و از یادگیری نیمه نظارتی استفاده می‌کند. برعکس، روش ما یک تکنیک بدون نظارت است و می‌توانیم اصلاحات عمومی کد را شناسایی کنیم.

سه روش اصلی برای تشخیص بد افزار وجود دارد که عبارتند از روش‌های مبتنی بر ناهنجاری، امضاء و ویژگی. چیزی که باعث عدم تشخیص بد افزار از نرم افزار قانونی می‌شود استفاده از ترفندهای مبهم سازی در بدافزارها است(Idika and Mathur, 2007). تشخیص بدافزار براساس امضاء، به بررسی توالی بایت‌ها می‌پردازد. به عبارتی امضاء را می‌توان توالی بایت‌های برنامه دانست که منحصر به فرد بوده و در داخل فایل اجرایی بدافزار موجود است. توالی بایت‌های نرم افزارهای قانونی در پایگاه‌های داده موجود است و در روش تشخیص بدافزار مبتنی بر امضاء، این توالی بایت‌ها برای برنامه‌های در حال اجرا با توالی بایت در پایگاه داده مقایسه شده و در صورت وجود هر گونه تضاد بین آنها، برنامه را به عنوان بدافزار شناسایی می‌کند(Elhadi et al., 2012). در روش‌های تشخیص بدافزار براساس ناهنجاری، از دانش خود برای بررسی رفتار عادی در یک برنامه استفاده کرده و نوع آن را تشخیص می‌دهد. در این روش معمولا تشخیص ناهنجاری در دو فاز آموزش و تشخیص انجام می‌شود. در طی فاز آموزشی رفتارهای عادی به برنامه آموزش داده شده و پس از آموزش، هر رفتاری مغایر با رفتار عادی را به عنوان ناهنجاری تشخیص می‌دهد(Tang et al., 2014). تشخیص براساس ویژگی از زیر مجموعه‌های تشخیص مبتنی بر ناهنجاری است که برای کاهش خطای نرخ تشخیص بدافزار ارائه شده است. این روش به جای بررسی رفتار عادی، به بررسی و استخراج ویژگی‌های موثر از برنامه‌های عادی می‌پردازد. این تکنیک در زمان یادگیری، به فراگیری و استخراج ویژگی‌های رفتار عادی سیستم تحت حفاظت می‌پردازد و سپس این رفتارهای عادی را با رفتار برنامه‌های تحت بررسی در زمان اجرا مقایسه می‌کند(Landage and Wankhade, 2013).

Forrest et al., (1996) یک روش پیشنهاد کردند که توالی فراخوان‌های سیستم را به منظور کشف کد مخرب، نظارت می‌کرد. در این روش ابتدا، پروفایل‌های نشان دهنده رفتار عادی سرویس‌های سیستم باید توسعه یابد. در این روش، از فاصله همینگ برای تعیین میزان شباهت توالی فراخوان‌های سیستم با دیگر فراخوان‌ها استفاده شد. در این مدل به طور معمول فرآیندهای نشان دهنده مقادیر فاصله همینگ بزرگ غیرعادی تلقی می‌شدند.این روش قادر به یافتن نفوذهایی، که سعی در بهره برداری از برنامه‌های مختلف دارند، می‌باشد. نقطه ضعف روش ارائه شده نادیده گرفتن نفوذ براساس سایر پارامترها است که در این صورت این گونه حملات قادر به گریز از تشخیص خواهند بود.

Wang and Stolfo, (2004) پایل را به عنوان ابزاری که بار مورد انتظار برای هر سرویس بر روی یک سیستم را محاسبه می‌کند، معرفی کردند. توزیع فراوانی بار ایجاد شده، اجازه می‌دهد یک مدل مرکزی برای سرویس‌ها توسعه یابد، این مدل مرکزی در طی فاز یادگیری ایجاد می‌شود. شناسایی کننده، بار دریافتی را با مدل مرکزی مقایسه می‌کند و در آنوبیس فاصله میان این دو را مقایسه می‌کند. فاصله آنوبیس[8] نه تنها مقدار میانگین بردار ویژگی بلکه شباهت را توسط ابزارهای آماری قوی واریانس و کوواریانس محاسبه می کند. اگر بار دریافتی بیش از حد از مدل مرکزی فاصله داشته باشد، بار دریافتی مخرب در نظر گرفته می شود. داده‌های آنها شامل ۲۱ روز داده‌های آموزشی و ۱۴ روز داده‌های آزمایش بود. طبق نتایج به دست آمده، از ۲۰۱ حمله در آزمایشگاه داده لینکلن، ۹۷ حمله توسط این روش شناسایی شدند. نویسندگان نرخ تشخیص صحیح برای روش خود را ۶۰٪ و نرخ مثبت نادرست را ۱٪ و کمتر از آن تخمین زدند.

[8] Anubis Distance

مجلـه‌ی علمی پژوهش در علوم رایانه ، شماره ۲۱، بهار ۱۴۰۰، ص ۲۸- ۱۴

ISSN: 2745-3673
http://www.Rcsj.ir

جدول (۱): مروری بر تحقیقات انجام شده در زمینه کاربرد روش‌های مبتنی بر هوش مصنوعی در زمینه تشخیص نوع بد افزار

محقق(ین)	نوع بد افزار	الگوریتم تشخیص بد افزار	دقت طبقه بندی
مهدی زاده (۱۳۹۷)	Lolyda.AA3, Dropper, VB.AT, Adware.Elex.pjI, Patched.N	شبکه عصبی مصنوعی	٪ ۸۶
Kalash et al., (2018)	VB.AT, Fakerean, Allaple A C2LOP.P, Lolyda. AA3, Alueron.gen!j	ماشین بردار پشتیبان شبکه‌های عصبی یادگیری عمیق	٪ ۹۸/۵۲ ٪ ۹۹/۹۷
Gibert, (2016)	Ramnit, Lollipop, Kelihos_ver3 Vundo, Simda, Tracur, Gatak Kelihos_ver, Obfuscator.ACY	شبکه های عصبی یادگیری عمیق	٪ ۹۸/۵۶
Makandar and Patrot, (2015)	SWIZZOR, VAPSUP, IKING_DLL, VIKING_DZ VIRUT, WOIKOINER ZHELATIN,	شبکه عصبی مصنوعی	٪ ۹۶/۳۵
Burnap et al., (2018)	بد افزار، نرم افزار قانونی	ماشین بردار پشتیبان شبکه عصبی مصنوعی	٪ ۶۸/۰۸ ٪ ۷۹/۴۰

شرح روش پیشنهادی

در شکل (۱) مراحل انجام تحقیق حاضر به صورت شماتیکی و به صورت فلوچارت نشان داده شده است. در اولین گام فایل تصاویر بد افزار در برنامه متلب فراخوانی شد. در مرحله پردازش تصاویر، از هر تصویر ماتریس هم وقوعی در زوایای مختلف استخراج شد. سپس از ماتریس های هم وقوعی ویژگی های آماری مرتبه دوم استخراج و سپس در بین ویژگی های استخراجی، ویژگی های بهتر و بهینه برای تشخیص نوع بدافزار انتخاب شدند. در اخر نیز از طریق پارامترهای ارزیابی دقت روشهای تشخیص بدافزار با هم مقایسه شد.

جدول (۲): نوع و تعداد بدافزارهای مورد تحلیل در تحقیق حاضر

ردیف	نام بدافزار
۱	Instantaccess
۲	Yuner
۳	Obfuscator
۴	Skintrim
۵	Fakerean
۶	Wintrim.BX
۷	VB.AT
۸	Allaple.A

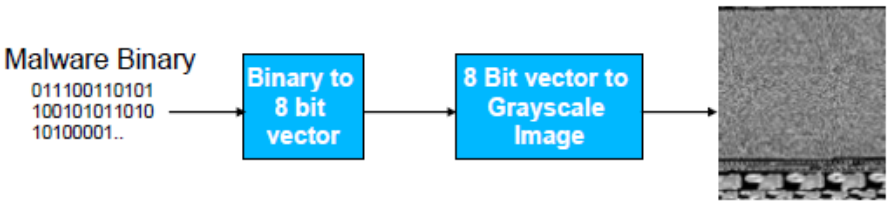
شکل (۱): فلوچارت انجام تحقیق

تبدیل تصویر به بدافزار

در شکل (۲) فرایند تبدیل کدباینری بدافزار به تصویر نشان داده شده است. لازم به ذکر است در این تحقیق از داده های دیتابیس که از نوع تصویر بودند استفاده شد و در ادامه فقط روش تبدیل هش های بد افزار به تصویر بیان شده است. برای تبدیل بدافزار به تصویر ابتدا فایل باینری بدافزار استخراج شد، سپس تمام (صفر و یک‌ها) به بردارهای هشت بیتی تقسیم شد یعنی یک آرایه‌ای ایجاد شد و هر درایه آرایه هشت بیت در نظر گرفته شد. سپس ماتریسی مطابق جدول (۳) ایجاد گردید

ISSN: 2745-3673
http://www.Rcsj.ir

(ستون ماتریس بسته به حجم فایل و ارتفاع ماتریس با توجه به اندازه فایل متفاوت خواهد بود) و مقادیر آرایه درون این ماتریس قرار داده شد و تمام درایه‌های ماتریس که بین صفر تا ۲۵۵ هست به بایت تبدیل شدند. در واقع این اعداد پیکسل‌های تصویر را تشکیل دادند. سپس از طریق کد رنگ‌ها از صفر تا ۲۵۵ هر عضو ماتریس به صورت پیکسل رنگی سیاه و سفید در آمد که صفر نشانگر رنگ سیاه و ۲۵۵ رنگ سفید بود و مابین آن، حالات خاکستری رنگ را دارد و در آخر ماتریس به عنوان تصویر ذخیره شد که تصویر تولید شده تصویر بدافزار می‌باشد.



شکل (۲): روش تبدیل کد باینری بدافزار به تصویر خاکستری(Nataraj et al., 2011)

جدول (۳): ارتباط بین حجم فایل باینری بدافزار و اندازه تصویر آن

محدوده حجم فایل	پهنای تصویر
<10 kB	32
10 kB – 30 kB	64
30 kB – 60 kB	128
60 kB – 100 kB	256
100 kB – 200 kB	384
200 kB – 500 kB	512
500 kB – 1000 kB	768
>1000 kB	1024

شبکه های عصبی مصنوعی

در شبکه عصبی مصنوعی نورون‌های موجود در لایه ورودی داده‌ها را دریافت کرده و از طریق اتصال‌های وزن دار آنها را به نورون‌های موجود در اولین لایه پنهان منتقل می‌کنند. داده‌ها با استفاده از روابط ریاضی پردازش می‌شوند و نتیجه را به نورون‌های موجود در لایه بعدی منتقل می‌کنند. خروجی شبکه توسط نورون‌های موجود در لایه آخر فراهم می‌شود. نورون j-th در یک لایه پنهان داده ورودی (xi) را از طریق محاسبه مجموع وزن‌ها (i) و افزودن یک « ترم بایاس » (θj) مطابق فرمول زیر پردازش می‌کند(Filippo Amato, 2013):

$$net_j = \sum_{i=1}^{m} x_i * w_{ij} + \theta_j \qquad j = (1,2,3,n) \qquad (۱)$$

لایه‌های شبکه عصبی

۱. لایه بُعدی (پرسپترون تک لایه) تابعی از متغیرهای حقیقی N مطابق با روش زیر هستند:

$$f(x_1, \ldots\ldots, x_N) = sgn(\sum_{i=1}^{N} w_i x_i - \theta) \qquad (۲)$$

در اینجا xi متغیرهای حقیقی هستند، $(x_1, \ldots\ldots, x_N)$ در برخی دامنه‌های متعلق به R، مقدار میگیرند، w_i پارامترهای حقیقی‌اند (وزنهای نورون)، _ آستانه فعال شدن نورون است، تابع sgm(x)برای $x \geq 0$ برابر با ۱ و برای $x < 0$ مساوی با صفر است. ما متغیر هم وزن شده نورون بالا را نیز در نظر گرفته و به جای استفاده از تابع sgm برای آن، از تابع افزایش یکنواخت هموار sgmکه از صفر تا ۱ متغیر است، استفاده می‌کنیم. ما بویژه نورون را مطابق فرمول زیر در نظر گرفتیم (Zakaria et al., 2014):

ISSN: 2745-3673
http://www.Rcsj.ir

$$f(x_1,\dots\dots,x_N)=sgn\left(\sum_{i=1}^{N}w_ix_i-\theta\right),sgm(x)=\frac{1}{1+e^{-x}} \quad (۳)$$

۲. پرسپترون چند لایه (MLP) : این نوع پرسپترون از سه لایه متوالی تشکیل شده است: یک لایه ورودی، یک لایه پنهان، و یک لایه خروجی. هر سیستم اساسا سه لایه است که این لایه ها عبارتند از لایه ورودی (آشکار)، لایه پنهان و لایه خروجی. لایه ورودی دارای نورون‌های ورودی است که داده‌ها را از طریق سیناپس‌ها به لایه پنهان منتقل می‌کنند و لایه پنهان نیز این داده‌ها را از طریق سیناپس‌های بیشتر به لایه خروجی منتقل می‌کند. سیناپس‌ها مقادیری به نام وزن را ذخیره می‌کنند که به آنها کمک می‌کند تا ورودی و خروجی را به لایه‌های مختلف هدایت کنند(Zakaria et al., 2014).

خوشه بندیk-means

یکی از رو ش‌های معتبر خوشه بندی k-means است که در این پژوهش از این الگوریتم برای خوشه بندی نوع بدافزار استفاده شده است. این روش، علیرغم سادگی آن، یک روش پایه برای بسیاری از رو ش‌های خوشه‌بندی دیگر (مانند خوشه‌بندی فازی) محسوب می‌شود. این روش، روشی انحصاری و مسطح محسوب می‌شود. برای این الگوریتم شکل‌های مختلفی بیان شده است، ولی همهٔ آ نها دارای روالی تکراری هستند که برای تعدادی ثابت از خوشه‌ها سعی در تخمین موارد زیر دارند (Zha et al., 2002):

الف-به دست آوردن نقاطی به عنوان مراکز خوشه ها. این نقاط در واقع، همان میانگین نقاط متعلق به هر خوشه هستند.

ب-نسبت دادن هر نمونهٔ داده به یک خوشه که آن داده کمترین فاصله تا مرکز آن خوشه را دارا باشد.

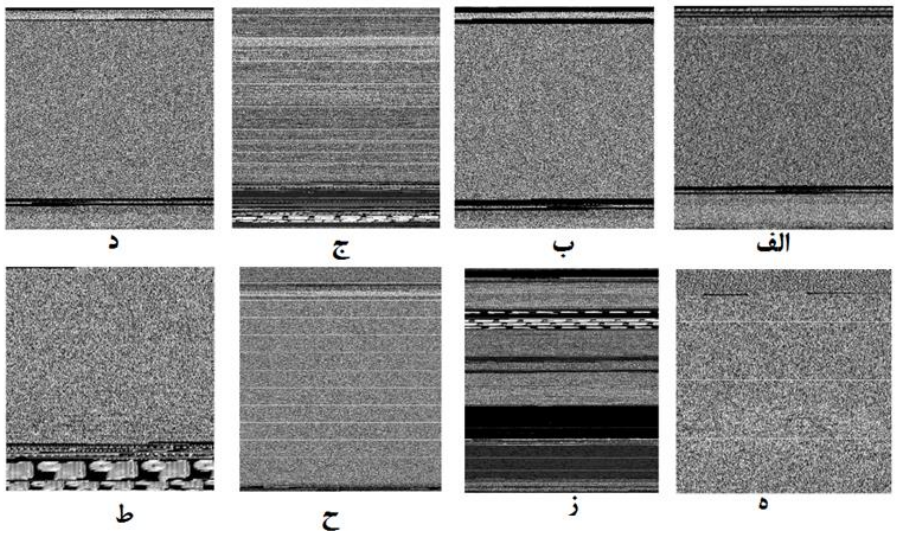
پایگاه داده مورد استفاده

در این تحقیق از داده های موجود در پایگاه داده Malimg Dataset استفاده شد. این پایگاه داده به صورت رایگان در رفرنس WebSite, (2020) قابل دسترسی و دانلود است. این پایگاه داده شامل ۲۵ خانواده از بدافزارهای مختلف است که در این تحقیق هشت خانواده از بدافزارها بررسی شد. در جدول (۲) این هشت خانواده برآورد شده است.

تصاویر بدافزار

ابتدا کدهای هش هر بدافزار تبدیل به کدهای باینری و سپس هر کد باینری تبدیل به یک تصویر سطح خاکستری شد. در شکل (۳) شکل سطح خاکستری تصاویر مربوط به هشت نوع بدافزار نشان داده شده است. همانطور که در شکل مشخص است اختلاف بین بافت تصاویر برای انواع مختلف بدافزار کاملا مشخص است. بررسی ظاهری بدافزارهای مختلف نشان می‌دهد که بیشترین اختلاف بین بدافزارهای خانواده Obfuscator.AD، Wintrim.BX و Allaple.A است که کاملا مشهود و مشخص است اما اختلاف بین سایر خانواده‌های مختلف بدافزارها جزئی بوده و با تحلیل شکل ظاهری به راحتی نمی‌توان این نوع بدافزارها را تشخیص داد. بیشترین اختلافی که بین تصاویر بدافزار وجود دارد مربوط به بافت تصاویر است تا رنگ تصاویر هر چند اختلاف رنگ نیز بین خانواده‌های Obfuscator.AD و Wintrim.BX با سایر بدافزارها بیشتر است اما عمده تفاوت در تصاویر بدافزارها مربوط به بافت تصاویر است. برای تحلیل بیشتر تصاویر و یافتن اختلاف بین تصویر هر بدافزار با سایر بدافزارها عملیات استخراج ویژگی از هیستوگرام تصاویر و ماتریس هم وقوعی انجام شد که نتایج آن در بخش‌های بعدی ارائه شده است.

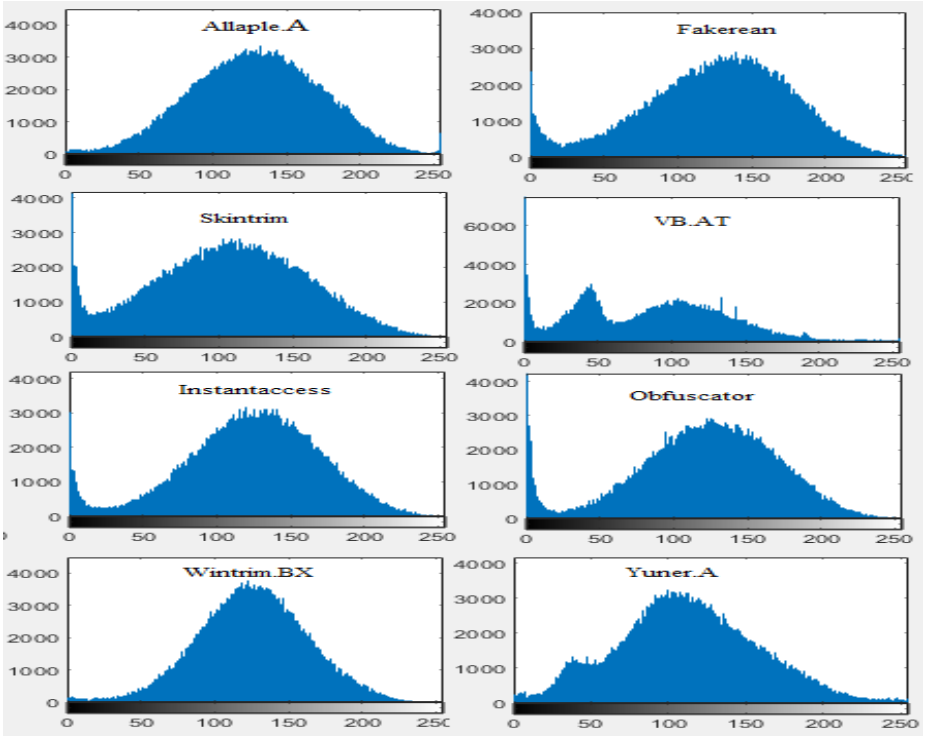
ISSN: 2745-3673
http://www.Rcsj.ir



شکل (۳): تصاویر سطح خاکستری بدافزارهای مختلف(الف: Instantaccess. ب:Yuner.A . ج:Obfuscator.AD . د: Skintrim. ه: Fakerean. ز: Wintrim.BX. ح:VB.AT . ط.Allaple.A)

هیستوگرام تصاویر

چون تصاویر بدافزار از نوع سطح خاکستری بودند لذا فقط ماتریس سطح خاکستری برای هر تصویر استخراج گردید. در شکل (۴) هیستورگرام تصاویر مربوط به بدافزارهای مختلف نشان داده شده است. اختلاف زیادی بین هیستوگرام‌های مختلف وجود ندارد و فقط هیستوگرام بدافزار خانواده VB-AT با بقیه خانواده‌ها زیاد است به طوری که توزیع هیستوگرام این نوع بدافزار بیشتر به سمت سطح تیره و رنگ تیره تمایل دارد این در حالی است که هیستوگرام سایر بدافزارها دارای توزیعی نرمال بوده که بیشترین نرخ رنگ تصاویر ببین شدت ۸۰ تا ۱۸۰ است. هیستوگرام بدافزار Yuner-A نیز داریا توزیع غیرنرمال بوده و نمودار هیستوگرام آن به سمت رنگ تیره تمایل دارد و یا به اصطلاح به سمت چپ چوله است.



شکل (۴): هیستوگرام تصاویر سطح خاکستری بدافزارهای مختلف

ISSN: 2745-3673
http://www.Rcsj.ir

ماتریس هم وقوعی

در این بخش نتایج حاصل از استخراج ماتریس هم وقوعی برای خانواده های مختلف بدافزار و هم چنین در جها مختلف مثلثاتی ارائه شده است. در شکل (۵) ماتری هم وقوعی برای بدافزار Yuner-A در چهار جهت مختلف صفر، ۴۵، ۹۰ و ۱۳۵ درجه نشان داده شده است. بررسی اعداد موجود در درایه‌های این ماتریس‌ها نشان داد که راستای استخراج ماتریس هم وقوعی حتی برای یک نوع تصویر بدافزار نیز با هم متفاوت است. پس این موضوع نشان داد که بافت تصاویر ماتریس هم وقوعی همگن نیست به طوری که در بدافزار نوع Yuner-A تغییرات بافت در راستای افقی یا صفر درجه تقریبا همگن است اما شدت تغییرات بافت در راستای عمودی یا ۹۰ درجه خیلی بیشتر از سایر راستاها است و شدت تغییر در بافت بد افزار در راستای ۴۵و ۱۳۵ درجه نیز تقریبا مشابه هم بوده و حد وسط حالت صفر و ۹۰ درجه است. در شکل (۶) ماتریس هم وقوعی تصویر سطح خاکستری در زاویه صفر درجه برای هشت خانواده مختلف از بدافزارها نشان داده شده است. بررسی اعداد موجود در درایه‌های متناظر ماتریس‌ها با هم نشان داد که اختلاف زیادی بین ماتریس‌های هم وقوعی خانواده های مختلف وجود دارد اما با توجه به اینکه امکان بررسی درایه به درایه اعداد موجود در این ماتریسها برای هشت خانواده بدافزار به صورت مشاهده عینی امکان پذیر نیست لذا از این ماتریسها ویژگی‌های آمار استخراج شد که در بخش بعدی به بررسی و تحلیل نتایج حاصل از آن پرداخته شده است.

صفر درجه

	1	2	3	4	5	6	7	8
1	2132	1516	197	81	32	5	2	1
2	1494	6593	2718	1334	419	77	17	1
3	205	2655	9605	8148	2922	701	84	4
4	77	1291	8020	12557	7323	2275	355	23
5	28	436	2959	7154	7322	3593	748	39
6	13	109	690	2262	3528	3990	1449	122
7	0	13	88	350	739	1437	1309	220
8	0	1	8	17	45	116	218	833

٤٥ درجه

	1	2	3	4	5	6	7	8
1	1125	1601	569	285	154	172	43	17
2	1790	4733	2597	2013	899	495	117	9
3	581	2615	6684	7588	4311	1930	537	49
4	244	1989	7589	10448	7122	3361	946	89
5	111	1002	4279	6975	5530	3083	975	182
6	82	493	1853	3397	3095	2177	856	192
7	7	109	484	968	1048	826	473	241
8	4	25	64	144	153	149	235	464

۱۳۵ درجه

	1	2	3	4	5	6	7	8
1	1357	1539	550	215	115	155	23	12
2	1750	5008	2894	1816	741	362	78	6
3	543	2954	7535	7879	3566	1443	374	36
4	194	1959	7608	10929	7155	3194	763	66
5	71	697	3709	7138	6147	3320	996	178
6	45	366	1503	3037	3436	2594	1053	183
7	1	63	345	786	1067	1008	654	266
8	0	22	48	116	154	158	249	498

۹۰ درجه

	1	2	3	4	5	6	7	8
1	1145	1810	604	250	88	57	6	1
2	1588	4733	2707	2062	934	455	103	24
3	589	2473	6850	7771	4171	1777	461	65
4	254	2051	7690	10481	7079	3223	939	119
5	158	957	4104	6946	5716	3192	1027	162
6	182	475	1826	3285	3016	2305	905	168
7	21	105	437	898	997	981	481	236
8	12	10	38	78	186	186	260	468

شکل (۵): ماتریس هم وقویع بدافزار خانواده Yuner-A در چهار جهت مختلف مثلثاتی

ISSN: 2745-3673
http://www.Rcsj.ir

Yuner.A

	1	2	3	4	5	6	7	8
1	2132	1516	197	81	32	5	2	1
2	1494	6593	2718	1334	419	77	17	1
3	205	2655	9605	8148	2922	701	84	4
4	77	1291	8020	12557	7323	2275	355	23
5	28	436	2959	7154	7322	3593	748	39
6	13	109	690	2262	3528	3990	1449	122
7	0	13	88	350	739	1437	1309	220
8	0	1	8	17	45	116	218	833

Allaple.A

	1	2	3	4	5	6	7	8
1	675	231	242	227	121	38	8	0
2	234	1088	1868	1827	1110	415	74	3
3	301	1834	4403	5646	4431	1853	401	25
4	207	1850	5698	8679	8190	4459	1241	120
5	94	1136	4405	8340	9372	6210	2070	211
6	28	412	1877	4435	6280	5352	2341	338
7	4	79	406	1223	2115	2371	1382	247
8	0	10	55	170	313	423	328	69

Fakerean

	1	2	3	4	5	6	7	8
1	5196	777	340	205	108	33	4	1
2	723	1728	1702	1475	928	356	61	4
3	397	1640	3895	4334	3433	1638	336	31
4	205	1479	4381	7379	6738	3728	1057	90
5	112	973	3408	6747	9232	5920	1769	182
6	26	309	1593	3817	5863	6430	2369	261
7	21	78	381	1063	1888	2267	1795	319
8	1	14	43	112	248	282	320	482

Instantaccess

	1	2	3	4	5	6	7	8
1	5189	399	188	144	74	35	4	0
2	405	1086	1614	1615	937	302	51	5
3	207	1585	4166	5416	3964	1572	325	30
4	130	1633	5318	8971	8125	4080	992	75
5	65	954	4008	8043	9033	5664	1698	186
6	36	302	1621	4023	5563	4237	1711	243
7	2	57	318	1039	1741	1623	948	172
8	0	3	33	85	205	206	179	65

Obfuscator

	1	2	3	4	5	6	7	8
1	8610	357	208	159	84	28	2	0
2	361	1197	1662	1564	911	328	75	5
3	203	1605	4189	5066	3826	1602	360	23
4	156	1569	5086	8326	7454	3986	954	70
5	88	944	3763	7489	8002	5249	1592	163
6	22	356	1581	3934	5223	4052	1704	234
7	2	57	357	977	1674	1630	794	180
8	1	7	28	80	159	215	187	57

Skintrim

	1	2	3	4	5	6	7	8
1	7786	1411	1114	707	323	93	12	0
2	1435	3393	3741	3198	1768	610	96	5
3	1113	3805	6018	6075	4097	1797	391	36
4	688	3095	6122	7742	5986	3253	886	90
5	312	1784	4071	6065	5946	3592	1265	150
6	78	615	1834	3149	3674	2909	1196	177
7	12	108	410	874	1263	1213	929	144
8	1	7	30	74	160	170	155	75

VB.AT

	1	2	3	4	5	6	7	8
1	25295	2226	103	21	7	1	0	0
2	2227	15922	1642	480	142	35	5	0
3	103	1674	7171	4887	1386	229	19	2
4	20	488	4972	10056	4332	787	86	1
5	6	124	1349	4355	5736	1571	164	10
6	5	17	221	867	1558	1743	340	23
7	0	2	16	71	154	377	655	142
8	0	0	0	2	10	23	147	889

Wintrim.BX

	1	2	3	4	5	6	7	8
1	2132	1516	197	81	32	5	2	1
2	1494	6593	2718	1334	419	77	17	1
3	205	2655	9605	8148	2922	701	84	4
4	77	1291	8020	12557	7323	2275	355	23
5	28	436	2959	7154	7322	3593	748	39
6	13	109	690	2262	3528	3990	1449	122
7	0	13	88	350	739	1437	1309	220
8	0	1	8	17	45	116	218	833

شکل (۶): ماتریس هم وقوعی تصویر سطح خاکستری در زاویه صفر درجه

استخراج ویژگی

ویژگی‌های استخراجی به دو دسته ویژگی‌های رنگ و ویژگی‌های بافت دسته بندی شدند. پنج ویژگی رنگی از هیستوگرام سطح خاکستری و ۲۰ ویژگی از ماتریس‌های هم وقوعی سطح خاکستری هر تصویر بدافزار استخراج گردید. در مجموع هر تصویر بدافزار با ۲۵ مشخصه اماری معرفی شد که برای شناسایی این مشخصه ها برای هر کدام برچسب ویژگی تعریف گردید که این برچسب ها در جدول (۴) ارائه شده‌اند. به عنوان مثال ویژگی Hem-G45 یعنی ویژگی همگنی که از ماتریس هم وقوعی در راستای ۴۵ درجه گرفته شده است. برای ویژگی های استخراجی از هیستوگرام برچسب گذاری بر این اساس انجام شد که F1، F2، F3، F4 و F5 به ترتیب میانگین، انحراف معیار، واریانس، چولگی و میانگین هارمونیک شدت بودند به طوری که F1-Hist یعنی ویژگی میانگین شدت هیستوگرام و یا F3-Hist یعنی واریانس شدت پیکسل‌های موجود در هیستوگرام.

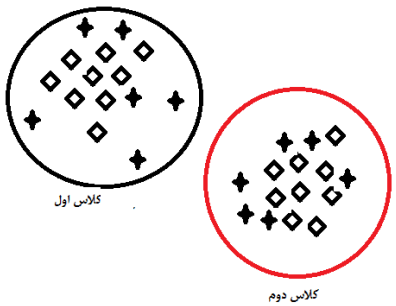
جدول (۴): برچسب گذاری ویژگی های استخراجی از تصاویر

نوع ویژگی	برچسب	سطح ویژگی	برچسب
همبستگی	Cor	ماتریس هم وقوعی-صفر درجه	G-0D
کنتراست	Con	ماتریس هم وقوعی-۴۵ درجه	G-45D
همگنی	Hem	ماتریس هم وقوعی-۹۰ درجه	G-90D
انرژی	Eng	ماتریس هم وقوعی-۱۳۵ درجه	G-135D
آنتروپی	Ent	هیستوگرام	Hist

ISSN: 2745-3673
http://www.Rcsj.ir

انتخاب ویژگی

در فرایند انتخاب ویژگی، ویژگی‌هایی که برای هر کلاس دارای همبستگی زیادی هستند ولی مقدار آنها برای کلاس‌های مختلف با هم همبستگی ندارد به عنوان ویژگی برتر و بهینه انتخاب شدند. در این تحقیق برای انتخاب ویژگی از روشی با عنوان CFS[9] استفاده شد. در شکل (۷) فرض کنید لوزی و ستاره دو ویژگی باشد که مبنای تفکیک دو کلاس اول و دوم از هم باشد. براساس روش همبستگی، همبستگی بین ویژگی لوزی برای کلاس اول و هم چنین برای کلاس دوم زیاد است اما مقادیر این دو ویژگی برای هر دو کلاس به طور همزمان همسبته نیست اما ویژگی ستاره در بین هر دو کلاس کاملا پراکنده است و برای هیچ کدام از دو کلاس اول و دوم دارای همبستگی نیست.



شکل (۷): مبنای انتخاب ویژگی براساس همبستگی

نتایج نشان داد که از بین ۲۵ ویژگی انتخابی (۵ ویژگی رنگ و ۲۰ ویژگی بافت) فقط ۱۳ ویژگی حاوی اطلاعات مربوط به بدافزار هستند و سایر ویژگی ها فاقد اطلاعات مناسب برای تشخیص نوع بدافزار می‌باشند. بررسی ویژگی‌های منتخب نشان داد که فقط یک ویژگی رنگی به عنوان ویژگی مفید انتخاب شده است. البته بررسی هیستوگرام تصاویر بدافزار هم نشان داد که بین رنگ سطح خاکستری تصاویر بدافزار اختلاف زیادی وجود ندارد. تنها ویژگی مناسب رنگی میانگین شدت پیکسل تصاویر سطح خاکستری بود.

بررسی ویژگی‌های استخراجی از ماتریس هم وقوعی نشان داد که تمام پنج ویژگی استخراجی به عنوان ویژگی مناسب انتخاب شده‌اند اما ماتریس هم وقوعی راستای ۹۰ درجه دارای بیشترین اطلاعات مربوط به نوع بدافزار است. هم چنین زوایای ۴۵ و ۱۳۵ درجه نیز به ترتیب در اولویت‌های بعدی از لحاظ مفید بودن جهت تشخیص نوع بدافزار هستند. در مجموع ویژگی‌های همبستگی در راستای ۴۵ و ۹۰ درجه، کنتراست در راستای ۴۵، ۹۰ و ۱۳۵ درجه، همگنی در راستای صفر، ۳۵، ۹۰ و ۱۳۵ درجه، انرژی در راستای ۴۵، ۹۰ و ۱۳۵ درجه، انتروپی در راستای ۹۰ درجه و در نهایت میانگین هیستوگرام تصاویر به عنوان ویژگی‌های مناسب و بهینه انتخاب شده اند.

اجرای الگوریتم شبکه عصبی مصنوعی

ورودی‌های شبکه عصبی مصنوعی همان ۱۴ ویژگی برتر هستند. در واقع ماتریس ورودی به شبکه عصبی دارای ۴۰۰ سطر و ۱۴ ستون است که ستون‌های نشان دهنده تعداد ویژگی ها و سطرها نیز تعداد تکرارهای مربوط به خانواده های مختلف بدافزار هستند. نوع داده‌های ورودی به شبکه عصبی از نوع آرایه تک بعدی است. نتایج مدلسازی جهت تشخیص و طبقه بندی نوع بدافزار توسط شبکه عصبی مصنوعی و هم چنین پارامترهای ارزیابی شبکه در جدول (۵) ارائه شده است. بررسی نتایج نشان داد که دقت کلی شبکه عصبی مصنوعی برای طبقه بندی و تشخیص نوع بدافزار ۹۶.۴۵ ٪ بود. حساسیت شبکه عصبی در تشخیص نوع بدافزار برای هر خانواده متفاوت بود به طوری که برای تشخیص بدافزار خانواده Instantaccess، Yuner.A، Obfuscator.AD، Skintrim، Fakerean، Wintrim.BX، VB.AT و Allaple.A به ترتیب ۹۴، ۹۴، ۱۰۰، ۹۲، ۹۲، ۹۸، ۹۸ و ۸۴٪ به دست آمد که کمترین حساسیت طبقه بند مربوط به تشخیص بدافزارهای خانواده Allaple.A بود.

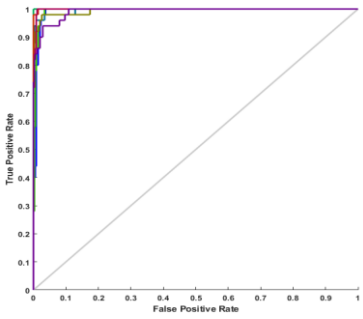
[9] Corrolation-base Feature Selection

ISSN: 2745-3673
http://www.Rcsj.ir

جدول (۵): ماتریس اغتشاش و پارامترهای ارزیابی شبکه عصبی مصنوعی مدل شده جهت تشخیص نوع بدافزار

		پیش بینی شده								
		In	YA	OA	Sk	Fa	WB	VA	AA	D
واقعی	Instantaccess(In)	۴۷	۰	۰	۰	۰	۰	۰	۳	۹۴
	Yuner.A(YA)	۰	۴۷	۰	۱	۰	۰	۲	۰	۹۴
	Obfuscator.AD(OA)	۰	۰	۵۰	۰	۰	۰	۰	۰	۱۰۰
	Skintrim(Sk)	۰	۱	۰	۴۶	۳	۰	۰	۰	۹۲
	Fakerean(Fa)	۰	۰	۰	۴	۴۶	۰	۰	۰	۹۲
	Wintrim.BX(WB)	۰	۰	۰	۰	۰	۴۹	۱	۰	۹۸
	VB.AT(VA)	۰	۰	۰	۰	۰	۱	۴۹	۰	۹۸
	Allaple.A(AA)	۸	۰	۰	۰	۰	۰	۰	۴۲	۸۴
دقت کلی طبقه بندی: ۹۶.۴۵ ٪										

در شکل (۸) منحنی مشخصه عملکرد شبکه عصبی مصنوعی در طبقه بندی نوع بدافزار نشان داده شده است. منحنی TPR ROC مشخص می‌کند که به چه نسبتی پیش‌بینی صحیح صورت گرفته است. در این منحنی هر چه نمودارها به سمت نقطه (۱و۰) تمایل داشته باشند نشان میدهد که آن طبقه بند دارای کمترین اشتباه در تصمیم گیری است. در مجموع نمودار مشخصه عملکرد یا ROC توسط خط y=x به دو بخش بالی خط و پایین خط دسته بندی می‌شود که چناچنه خطوط عملکردی مربوط به هر کلاس (در شکل ۴-۸ با رنگ های مختلف نشان داده شده است) رد بالای خط y=x باشد و به نقطه (۱و۰)نزیدک تر باشد نشان دهنده حساسیت بالا در تشخیص نوع بدافزار است.

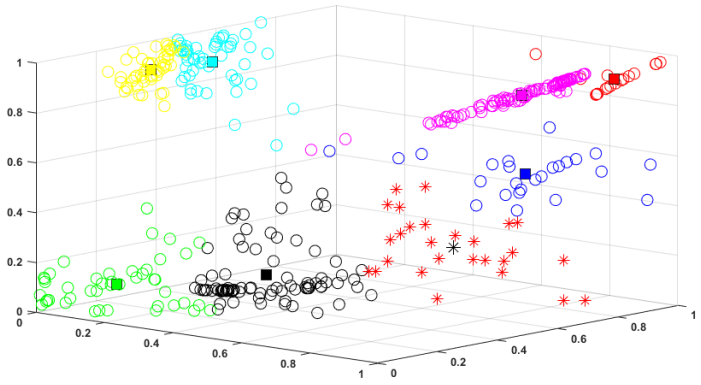


شکل(۸): منحنی مشخصه عملکرد شبکه عصبی مصنوعی در طبقه بندی نوع بدافزار

اجرای الگوریتم k-means

همانند سایر روشهای فوق ورودی از نوع آرایه یک بعدی و دارای ابعاد ۱۴×۴۰۰ بود. در این بخش نتایج خوشه بندی بدافزارها با استفاده از روش k-means ارائه شده است. در شکل (۹) خروجی کامینز جهت خوشه بندی بدافزارها نشان داده شده است. همانطور که در شکل مشخص است مرکز هر خوشه و نوع بدافزار برای هر خوشه تعیین شده است.

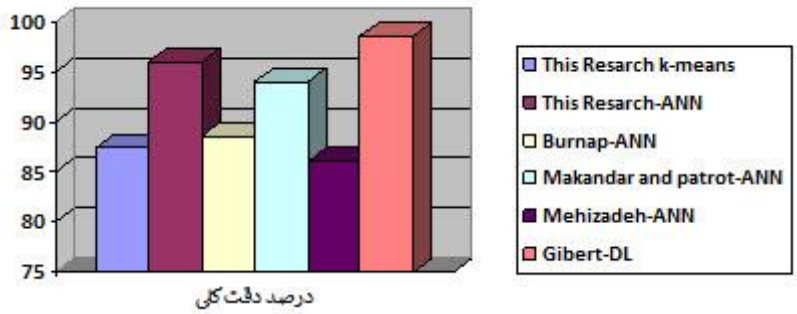
ISSN: 2745-3673
http://www.Rcsj.ir



شکل (۹): نتایج خوشه بندی بدافزارهای توسط روش K-means

مقایسه نتایج

در شکل (۱۰) مقایسه نتایج نشان داده شده است. بررسی تحقیقات مشابه نشان داد که مهدی زاده (۱۳۹۷) با استفاده از شبکه عصبی مصنوعی توانست با دقت ۸۶ ٪ پنج نوع بدافزار را تشخیص دهد. در تحقیق دیگری Burnap et al., (2018) با استفاده از شبکه‌های عصبی مصنوعی شش نوع بدافزار را با دقت ۸۸.۵۲ ٪ تشخیص دادند. Gibert, (2016) هشت نوع بدافزار را با استفاده از شبکه‌های عصبی یادگیری عمیق و Makandar and Patrot, (2015) هفت نوع بدافزار را با شبکه عصبی مصنوعی به ترتیب با دقت ۹۸.۵۶ ٪ و ۹۴ ٪ تشخیص دادند. در نهایت مقایسه نتایج نشان داد که دقت به دست آمده در این تحقیق نسبت به تحقیقات مشابه قابل قبول است و ذکر این نکته ضروری است که در روش‌های گذشته تعداد کلاس‌های بدافزار طبقه بندی کمتر از تعداد کلاس‌های تحقیق حاضر بوده است به همین دلیل در برخ موراد دقت تحقیقات مشابه نسبت به تحقیق حاضر بیشتر شده است.



شکل (۱۰): مقایسه نتایج

نتیجه گیری

نتایج به دست آمده در این تحقیق به شرح زیر است:

۱-اختلاف بین بافت تصاویر برای انواع مختلف بدافزار کاملا مشخص بود. بررسی ظاهری بدافزارهای مختلف نشان دا که بیشترین اختلاف بین بدافزارهای خانواده Obfuscator.AD، Wintrim.BX و Allaple.A وجود دارد اما اختلاف بین سایر خانواده‌های مختلف بدافزارها جزئی است.

۲-بیشترین اختلافی که بین تصاویر بدافزار وجود دارد مربوط به بافت تصاویر است تا رنگ تصاویر .هر چند اختلاف رنگ نیز بین خانواده‌های Obfuscator.AD و Wintrim.BX با سایر بدافزارها بیشتر بود. در کل اختلاف زیادی بین هیستوگرام‌های مختلف وجود ندارد و فقط هیستوگرام بدافزار خانواده VB-AT با بقیه خانواده‌ها زیاد بود.

۳- از بین ۲۵ ویژگی انتخابی فقط ۱۳ ویژگی حاوی اطلاعات مربوط به بدافزار بودند که فقط یک ویژگی رنگی به عنوان ویژگی مفید انتخاب شد و سایر ویژگی‌ها شامل همبستگی در راستای ۴۵ و ۹۰ درجه، کنتراست در راستای ۴۵، ۹۰ و ۱۳۵

ISSN: 2745-3673
http://www.Rcsj.ir

درجه، همگنی در راستای صفر، ۳۵، ۹۰ و ۱۳۵ درجه، انرژی در راستای ۴۵، ۹۰ و ۱۳۵ درجه، انتروپی در راستای ۹۰ درجه بودند.

۴-دقت مدل‌های مورد بررسی در تشخیص نوع بدافزار به ترتیب برای، شبکه عصبی مصنوعی و الگوریتم k-means به ترتیب۹۶.۴۵ ٪، و ۸۷.۳۵ ٪ بود.

پیشنهادها و راهکارهای آتی

پیشنهاد می‌گردد: از روش‌های تلفیق در حد تصمیم جهت طبقه بندی نوع بدافزار استفاده گردد. تعداد خانواده‌های بدافزار مورد بررسی بیشتر از ۸ عدد گردد. این مدل به صورت یک برنامه کاربردی مدلسازی شده و بر روی رایانه نصب و کارایی آن ارزیابی شود.

ISSN: 2745-3673
http://www.Rcsj.ir

منابع و مراجع

[۱] شیرازی، حسین.، فرشچی، سیدمحمدرضا. ۱۳۹۳. ارائه یک روش جدید برای شناسایی بدافزارها در سطح مجازی ساز در ماشین های مجازی. پدافند الکترونیکی و سایبری: ۷(۲): ۲۳-۳۴.

[۲] رنجی، هادی، و پارسا، سعید. ۱۳۹۷. شناسائی بدافزارها با استفاده از تصویرسازی. *فصلنامه پدافند غیرعامل*, ۹ (۲): ۹۵-۱۰۱.

[3] Gibert, D., 2016. Convolutional neural networks for malware classification. *University Rovira i Virgili, Tarragona, Spain*.

[4] Gupta, A., Kuppili, P., Akella, A. and Barford, P., 2009, January. An empirical study of malware evolution. In *2009 First International Communication Systems and Networks and Workshops* (pp. 1-10). IEEE.

[5] Ma, J., Dunagan, J., Wang, H.J., Savage, S. and Voelker, G.M., 2006, October. Finding diversity in remote code injection exploits. In *Proceedings of the 6th ACM SIGCOMM conference on Internet measurement* (pp. 53-64).

[6] Mercaldo, F., Di Sorbo, A., Visaggio, C.A., Cimitile, A. and Martinelli, F., 2018. An exploratory study on the evolution of Android malware quality. *Journal of Software: Evolution and Process*, *30*(11), p.e1978.

[7] Ouellette, J., Pfeffer, A. and Lakhotia, A., 2013, October. Countering malware evolution using cloud-based learning. In *2013 8th International Conference on Malicious and Unwanted Software:" The Americas"(MALWARE)* (pp. 85-94). IEEE.

[8] Ronen, R., Radu, M., Feuerstein, C., Yom-Tov, E. and Ahmadi, M., 2018. Microsoft malware classification challenge. *arXiv preprint arXiv:1802.10135*.

[9] WebSite, 2020. Available at: https://sarvamblog.blogspot.com/2014/08/supervised-classification-with-k-fold.html

[10] Zakaria, M., Al-Shebany, M. and Sarhan, S., 2014. Artificial neural network: a brief overview. *Int J Eng Res Appl*, *4*, pp.7-12.

[11] Zha, H., He, X., Ding, C., Gu, M. and Simon, H.D., 2002. Spectral relaxation for k-means clustering. In Advances in neural information processing systems (pp. 1057-1064).

[12] Ronen, R., Radu, M., Feuerstein, C., Yom-Tov, E. and Ahmadi, M., 2018. Microsoft malware classification challenge. *arXiv preprint arXiv:1802.10135*.

[13] Saxe, J. and Berlin, K., 2015, October. Deep neural network based malware detection using two dimensional binary program features. In *2015 10th International Conference on Malicious and Unwanted Software (MALWARE)* (pp. 11-20). IEEE.

[14] Kramer, S. and Bradfield, J.C., 2010. A general definition of malware. *Journal in computer virology*, *6*(2), pp.105-114.

[15] Forrest, S., Hofmeyr, S.A., Somayaji, A. and Longstaff, T.A., 1996, May. A sense of self for unix processes. In *Proceedings 1996 IEEE Symposium on Security and Privacy* (pp. 120-128). IEEE.

[16] Wang, K. and Stolfo, S.J., 2004, September. Anomalous payload-based network intrusion detection. In *International workshop on recent advances in intrusion detection* (pp. 203-222). Springer, Berlin, Heidelberg.

[17] Kalash, M., Rochan, M., Mohammed, N., Bruce, N.D., Wang, Y. and Iqbal, F., 2018, February. Malware classification with deep convolutional neural networks. In *2018 9th IFIP international conference on new technologies, mobility and security (NTMS)* (pp. 1-5). IEEE.

[18] Makandar, A. and Patrot, A., 2015, December. Malware analysis and classification using artificial neural network. In *2015 International conference on trends in automation, communications and computing technology (I-TACT-15)* (pp. 1-6). IEEE.

[19] Idika, N. and Mathur, A.P., 2007. A survey of malware detection techniques. *Purdue University*, *48*, pp.2007-2.

ISSN: 2745-3673
http://www.Rcsj.ir

[20] Elhadi, A.A., Maarof, M.A. and Osman, A.H., 2012. Malware detection based on hybrid signature behaviour application programming interface call graph. *American Journal of Applied Sciences*, *9*(3), p.283.

[21] Tang, A., Sethumadhavan, S. and Stolfo, S.J., 2014, September. Unsupervised anomaly-based malware detection using hardware features. In *International Workshop on Recent Advances in Intrusion Detection* (pp. 109-129). Springer, Cham.

[22] Landage, J. and Wankhade, M.P., 2013. Malware and malware detection techniques: A survey. *International Journal of Engineering Research and Technology (IJERT)*, *2*(12), pp.2278-0181.