Journal of Electronical & Cyber Defence
Vol. 9, No. 3, 2021, Serial No. 35

Malicious Domain Detection Using DNS Records

F. Bagheri[1] M. Rezvani [2*], M. Fateh[3], E. Tahanian[4]

*Shahrood Univercity of technical

(Received: 13/12/2020, Accepted: 08/02/2021)

ABSTRACT

One of the most important security challenges with the advance of technology in cyberspace is phishing attacks. Phishing is a type of cyber-attack that always tries to obtain information such as username, password, bank account information, and the like by forging a website, email address and convincing the user to enter this information. Due to the increasing growth of these attacks and the increasing complexity of the type of attack, current phishing detection systems often cannot adapt to new attacks and have low detection accuracy. Graph-based methods are one of the techniques for identifying malicious domains that use the connections between the domain and IP to identify. In this paper, a graph-based phishing detection system using deep learning is presented. The main steps in the proposed method include extracting IP from the domain, defining the relationship between the domains, determining the weights, and converting the data to a vector by the Node2vec algorithm. Then, using CNN and DENSE deep learning models, the classification and identification operations are performed. The experimental results over three different datasets show that the proposed method provides an accuracy of about 99% in identifying malicious domains, which has an acceptable improvement compared to state of the art in this context.

Keywords: Malicious Domain Detection, DNS Data, Phishing, Deep Learning

* Corresponding Author Email: mrezvani@shahroodut.ac.ir

نشریه علمی "پدافند الکترونیکی و سایبری"
سال نهم، شماره ۳، پاییز ۱۴۰۰، ص ۹۷-۸۳

علمی– پژوهشی

تشخیص هوشمند دامنه‌های مشکوک از داده‌های DNS

فهیمه باقری[۱]، محسن رضوانی[۲]*، منصور فاتح[۳]، اسماعیل طحانیان[٤]
۱- کارشناسی ارشد، ۲- دانشیار، ۳ و ٤- استادیار، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود
(دریافت: ۱۳۹۹/۰۹/۰۲، پذیرش: ۱۳۹۹/۱۰/۲۰)

چکیده

یکی از مهم‌ترین چالش‌های امنیتی با پیشرفت فناوری در فضای مجازی حملات فیشینگ یا تله‌گذاری است. تله‌گذاری نوعی حمله سایبری است که همواره در تلاش برای به‌دست آوردن اطلاعاتی مانند نام کاربری، گذرواژه، اطلاعات حساب بانکی و مانند آن‌ها از طریق جعل یک وب‌سایت، آدرس ایمیل و متقاعد کردن کاربر به منظور وارد کردن این اطلاعات می‌باشد. با توجه به رشد صعودی این حملات و پیچیده‌تر شدن نوع حمله، سیستم‌های تشخیص تله‌گذاری فعلی اغلب نمی‌توانند خود را با حملات جدید تطبیق دهند و دارای دقت پایین در شناسایی هستند. روش‌های مبتنی بر گراف یکی از روش‌های شناسایی دامنه‌های مشکوک است که از ارتباطات بین دامنه و IP برای شناسایی استفاده می‌کند. در این مقاله سیستم تشخیص تله‌گذاری مبتنی بر گراف با استفاده از یادگیری عمیق ارائه شده است. مراحل کار شامل استخراج IP از دامنه، تعریف ارتباط بین دامنه‌ها، تعیین وزن‌ها و همچنین تبدیل داده‌ها به بردار توسط الگوریتم Node2vec است. در ادامه با استفاده از نمونه‌های یادگیری عمیق CNN و DENSE عمل طبقه‌بندی و شناسایی انجام می‌شود. نتایج نشان می‌دهند که روش ارائه شده در این مقاله دقتی در حدود ۹۹ درصد در شناسایی دامنه‌های مشکوک دارد که در مقایسه با روش‌های قبل بهبود قابل قبول داشته است.

کلید واژه‌ها: تشخیص دامنه مشکوک، داده‌های DNS، تله‌گذاری، یادگیری عمیق

۱- مقدمه

فضای مجازی[۱] زندگی انسان را به‌طور معنی‌داری تغییر داده و با افزایش تصاعدی برنامه‌های متنوع نرم‌افزاری در طیف گسترده‌ای از حوزه‌ها مانند امور مالی، تجارت الکترونیکی، شبکه‌های اجتماعی، سیستم‌های اتوماسیون و سایر موارد همراه بوده است. از همین رو برقراری امنیت در این فضا به‌طور چشم‌گیری مورد توجه قرار گرفته است [۱]. دسترسی تعداد زیادی از مردم جهان به فضای مجازی و گسترش ارتباطات الکترونیکی بین افراد و سازمان‌های مختلف از طریق دنیای مجازی بستری مناسب را برای برقراری مراودات تجاری و اقتصادی فراهم کرده است [۲]. یکی از چالش‌های اصلی در فناوری اطلاعات حوزه امنیت آن است [۳]. هر روزه سارقان فضای مجازی راه‌های جدیدی را برای به‌دست آوردن هویت شخصی افراد و دست‌یابی به اطلاعات شخصی و اطلاعات مالی آن‌ها به‌کار می‌برند. یکی از روش‌هایی که اخیرا بسیار مورد توجه آن‌ها قرار گرفته و البته کمی هم پیچیده می‌باشد، فیشینگ یا تله‌گذاری[۲] نام دارد [۴]. فرآیند تله‌گذاری وب‌سایت را می‌توان یکی از چالش‌های امنیتی مهم برای ارتباطات برخط دانست. لذا، کشف و دسترسی به این حملات و مسدود کردن آن‌ها مهم است. معمولا طراحان این نوع حملات به روش‌های مختلف تلاش می‌کنند به اطلاعات شخصی افراد از طریق روش‌های متنوع مهندسی اجتماعی مانند ایمیل، تماس تلفنی، صفحات جعلی پرداخت، پیامک، انواع نمونه‌های ربات‌های تلگرام و انواع روش‌های جدیدی که انتظار آن نمی‌رود دست‌یابند، در این‌گونه حملات، حمله‌کننده با ارسال یک ایمیل، خود را به جای فرد یا شرکت معتبر و حتی بانک‌های معتبر جا می‌زند و با روش‌های گول‌زننده سعی می‌کند تا اطلاعات حساس را از قربانی بگیرد. روش‌های متنوعی برای شناسایی دامنه‌های مشکوک با استفاده از انواع مختلف شبکه محلی و اطلاعات میزبان[۳] پیشنهاد شده است [۵، ۳، ۶ و ۷].

توافقی در مورد نرم‌افزارهای ضد تله‌گذاری برای رسیدن به یک استاندارد مشترک وجود ندارد. از طرفی با وجود نکات امنیتی بازهم راه‌ها و روش‌هایی برای جلوگیری از تشخیص تله‌گذاری وجود دارد. امروزه مهاجمان سایبری نسبت به قبل آگاه‌تر و از روش‌های جدید و پیچیده‌ای برای گریز از شناسایی و مسدود شدن استفاده می‌کنند [۸].

۱-۱- انواع حملات تله‌گذاری

همان‌طور که قبل‌تر اشاره شد، حمله تله‌گذاری در واقع نوعی تلاش برای به‌دست آوردن اطلاعات از طریق شبیه‌سازی یک وب‌سایت، است. این نوع حملات دارای انواع مختلفی هستند که پایه اصلی آن‌ها جعل است.

*رایانامه نویسنده مسئول: mrezvani@shahroodut.ac.ir

[1] Internet
[2] Phishing
[3] Host

حملات تله‌گذاری بسیار متنوع بوده و با راه‌کارهای مختلفی قابل انجام است. همچنین بسته به خلاقیت فردی که این حملات را طراحی می‌کند به حالت‌های گوناگونی صورت می‌پذیرد. انواع حملات تله‌گذاری به‌شرح زیر است:

۱. تله‌گذاری تلفنی.
۲. تله‌گذاری پیامکی.
۳. تله‌گذاری توسط ربات‌ها و بدافزارها.
۴. تله‌گذاری در وب شامل جعل وب‌سایت، ایمیل‌های جعلی، درگاه‌های پرداخت و... [۹].

لذا برای مقابله با حمله تله‌گذاری و عدم گرفتار آمدن در دام افراد کلاه‌بردار، شناسایی این حملات امری مهم و ضروری است.

۱-۲- دامنه‌های مشکوک

عامل اصلی انتشار حملات فضای مجازی از طریق دامنه‌هایی ایجاد می‌شود که دارای اهداف مضر هستند که به عنوان دامنه‌های مشکوک است. دامنه‌های مشکوک اجزای اصلی بسیاری از حملات سایبری مانند تله‌گذاری، بات‌نت و هرزنامه هستند که می‌تواند حریم خصوصی کاربر را به خطر بیندازد یا باعث خسارت مالی یا نصب بدافزار بر روی سیستم کاربر شود. همچنین دامنه‌های مشکوک برای تولید آدرس‌های مخرب استفاده می‌شود، که تهدیدی بسیار رایج و جدی برای امنیت فضای مجازی است [۱].

همان‌طور که در شکل (۱) نشان داده شده است نام دامنه یک آدرس منحصر بفرد بر روی فضای مجازی می‌باشد که عموما از سه بخش اصلی پیشوند، نام و پسوند تشکیل شده است.

شکل (۱): نمونه آدرس اینترنتی

۱-۳- سیستم نام دامنه(DNS)

سیستم نام دامنه (DNS)[1] یکی از مجموعه اصلی پروتکل‌های فضای مجازی است که وظیفه هدایت درخواست‌های منابع فضای مجازی به یک رایانه میزبان را بر عهده دارد و این منابع در فضای مجازی توسط آدرس‌هایی که از نام دامنه‌ها تشکیل شده‌اند شناسایی می‌شوند [۱۰]. در واقع DNS یک سیستم نام‌گذاری سلسله مراتبی برای منابع متصل به فضای مجازی است که نام‌های دامنه را به آدرس پروتکل اینترنت (IP)[2] مربوطه ترجمه می‌کند، بنابراین DNS یک سرویس ساده برای یافتن و تبدیل یک آدرس به آدرس IP و بالعکس است [۱۱]. داده‌های DNS یکی از بارزترین منابع اطلاعاتی مورد استفاده برای شناسایی دامنه‌های مشکوک است [۱۵- ۱۲].

رشد سریع اینترنت شناسایی دامنه‌های مشکوک را بسیار مهم و ضروری کرده است. رویکردهای متداول مورد استفاده برای شناسایی دامنه مشکوک، شامل لیست سیاه معمول نام دامنه‌ها [۱۶]، تجزیه و تحلیل ترافیک شبکه [۱۷]، تشریح محتوای صفحه وب [۱۸]، تجزیه و تحلیل ترافیک DNS [۱۹] و تجزیه و تحلیل ویژگی‌های برجسته واژگانی [۲۰] از دقت کافی برخوردار نیستند.

در این مقاله از یک روش منحصر به فرد برای شناسایی دامنه‌های مشکوک با استفاده از نمونه طبقه‌بندی یادگیری عمیق با روش مبتنی بر تحلیل گراف و نام دامنه از جمله داده‌های DNS، ایجاد ارتباطات دامنه -IP و تعیین وزن بین دامنه‌ها استفاده شده است. از سه مجموعه دادگان جداگانه که شامل دامنه‌های سالم و مشکوک است، ویژگی‌های ذکر شده را از نام آن دامنه‌ها استخراج کرده و سپس به نمونه طبقه‌بندی یادگیری عمیق DENSE و CNN[3] داده می‌شود، در نهایت با لیست جدیدی از داده‌ها آزمایش اجرا می‌شود تا دامنه‌ها به دو دسته مشکوک یا سالم طبقه‌بندی شوند.

در ادامه سازماندهی این مقاله به صورت زیر ارائه گردیده است. در بخش دوم مروری بر روش‌های پیشین خواهد شد. در بخش سوم روش پیشنهادی ارائه می‌شود. در بخش چهارم، نتایج آزمایش‌ها مورد بررسی قرار گرفته است و نهایتا در بخش پنجم، نتیجه‌گیری و جمع‌بندی کلی از بحث ارائه شده است.

۲- کارهای مرتبط

در این بخش، به بررسی جدیدترین کارهای انجام شده در زمینه تحلیل و شناسایی حملات پرداخته شده است. مطالعات موجود در مورد شناسایی دامنه مشکوک را می‌توان به‌طور گسترده‌ای بر اساس روش‌های اساسی به سه دسته راه‌حل‌های مبتنی بر طبقه‌بندی، مبتنی بر خوشه‌بندی و مبتنی بر نمودار دسته‌بندی کرد [۲۱]. در ادامه به تفصیل هریک از روش‌ها معرفی شده است.

[1] Domain Name System
[2] Internet Protocol
[3] Convolution Neural Network

۲-۱- راه‌حل‌های مبتنی بر طبقه‌بندی

در این روش‌ها از الگوریتم‌های طبقه‌بندی نظارت شده مانند ماشین‌های بردار پشتیبان و درخت تصمیم‌گیری برای ساخت طبقه‌بندی‌کننده برای تمایز دامنه‌های سالم و مشکوک استفاده می‌کنند [۷، ۱۹، ۲۲]. از آنجا که DNS به شدت توسط مهاجمان مورد سوء استفاده قرار گرفته است، در مرجع [۱۹] به توسعه مکانیسم‌های شناسایی بدافزار مبتنی بر DNS پرداخته شده است. این سیستم پیشنهادی EXPOSURE نام دارد. هدف این سیستم شناسایی دامنه‌های مخربی است که به عنوان بخشی از عملیات مخرب در اینترنت استفاده می‌شود. سیستم EXPOSURE با استفاده از ۱۵ ویژگی منحصر به فرد که در ۴ دسته، مبتنی بر زمان، مبتنی بر TTL[1]، مبتنی بر نام دامنه و مبتنی بر DNS Answer گروه‌بندی شده‌اند، برای شناسایی چنین دامنه‌هایی طراحی شده است. همچنین برای ارزیابی عملکرد سیستم از درخت تصمیم و الگوریتم J48 برای طبقه‌بندی دامنه‌ها استفاده می‌شود. دقت دامنه‌های شناسایی شده توسط این روش بیش از ۵۰ درصد می‌باشد.

در مرجع [۲۳]، یک روش مبتنی بر یادگیری ماشین برای شناسایی نام‌های دامنه مخرب با استفاده از ELM[2] پیشنهاد شده است. سیستم ELM یک شبکه عصبی مدرن با دقت بالا و سرعت یادگیری سریع است. از ELM برای طبقه‌بندی نام دامنه بر اساس ویژگی‌های استخراج‌شده استفاده کرده است. این روش با طبقه‌بندی‌کننده‌های LR (رگرسیون لجستیک)، CART، BPNN (شبکه عصبی برگشتی) و SVM مقایسه شده است. اگرچه کارایی محاسباتی LR و CART عالی است، ولی دقت حدود ۰/۹۱ است. این درحالی است که BPNN و SVM می‌توانند به دقت بالاتری با مقادیر ۰/۹۵ و ۰/۹۴ دست یابند ولی پیچیدگی زمان آموزش بسیار بالا است. دقت شناسایی به‌دست آمده توسط ELM بیش از ۹۶٪ می‌باشد.

در [۲۴] به منظور تشخیص وب‌سایت فیشینگ از وب‌سایت قانونی با استفاده از نرم افزار WEKA، اقدام به کاهش ویژگی‌ها با استفاده از الگوریتم‌های انتخاب ویژگی نموده است. سپس ویژگی‌های انتخاب شده با استفاده از الگوریتم‌های طبقه‌بندی J48، Random Forest، JRIP، SMO و PART مورد ارزیابی قرار گرفته است. نتایج تجربی نشان می‌دهد انتخاب ویژگی با روش پیشنهادی زمانی است که از جنگل تصادفی با حداقل ویژگی، دقتی برابر ۹۷/۲۲ جهت تشخیص فیشینگ به‌دست آمده است.

یکی ازمهم‌ترین روش‌های مبتنی بر طبقه‌بندی روش‌های یادگیری عمیق می‌باشد که در ادامه به آن پرداخته شده است.

۲-۱-۱- روش‌های یادگیری عمیق

یادگیری عمیق شاخه‌ای از بحث یادگیری ماشین و هوش مصنوعی[3] و مجموعه‌ای از الگوریتم‌هایی است که تلاش می‌کنند مفاهیم انتزاعی سطح بالا را با استفاده از یادگیری در سطوح و لایه‌های مختلف نمونه کنند. مطالعات بالینی نشان می‌دهند که ساختار مغز پستانداران از معماری شبکه‌های عصبی عمیق[4] بهره می‌برند. ایده یادگیری عمیق با الهام از ساختار طبیعی مغز انسان و به کمک امکانات و فن‌آوری‌های جدید، توانسته است در بسیاری از حوزه‌های مربوط به هوش مصنوعی و یادگیری ماشین، موفقیت‌های چشم‌گیری را کسب کند [۲۸-۲۵].

مهم‌ترین مزایای یادگیری عمیق عبارت‌اند از :

۱. یادگیری خودکار ویژگی‌ها.
۲. یادگیری چند لایه ویژگی‌ها.
۳. دقت بالا در نتایج.
۴. قدرت تعمیم بالا و شناسایی داده‌های جدید.
۵. پشتیبانی گسترده سخت‌افزاری و نرم‌افزاری [۲۶-۲۵].

در مرجع [۲۹]، رویدادنامه‌های DNS از دستگاه‌های مشتری در شبکه محلی (LAN) جمع شده است و آن را در یک سرور ذخیره می‌کنند. برای یافتن نام دامنه به عنوان خوش‌خیم یا مخرب، یک رویکرد مبتنی بر یادگیری عمیق پیشنهاد شده است. برای مقایسه، اثربخشی روش‌های مختلف یادگیری عمیق مانند شبکه عصبی RNN، حافظه کوتاه مدت بلند LSTM و سایر طبقه‌بندی‌کننده‌های یادگیری سنتی را ارزیابی کرده است. رویکردهای مبتنی بر یادگیری عمیق نسبت به سایر طبقه‌بندی‌کننده‌های یادگیری ماشین کلاسیک عملکرد خوبی داشته‌اند. دلیل اصلی این است که الگوریتم‌های یادگیری عمیق توانایی به‌دست آوردن ویژگی‌های مناسب را بطور ضمنی دارند. علاوه بر این، LSTM با دقت ۰/۹۷ و RNN دقتی برابر ۰/۹۶ در مجموعه داده اول به‌دست آورده است.

در مرجع [۳۰] به شناسایی نام‌های مخرب با استفاده از نمونه‌های مبتنی بر شبکه عصبی عمیق پرداخته است. شبکه LSTM دو جهته روی مجموعه داده آموزش توسعه داده می‌شود. دو کار روی شبکه LSTM مورد آزمایش قرار گرفته است. اولین کار شناسایی نام دامنه مخرب و کار دوم شناسایی کلاس نام دامنه است. روش پیشنهادی قادر است عملکرد شناسایی نام دامنه مخرب را با دقت ۹۸/۹ درصد و شناسایی کلاس مربوط به نام دامنه را با دقت فقط ۶۹/۷ درصد مورد بررسی قرار دهد.

[1] Time To Live
[2] Extreme Learning Machine
[3] Artificial intelligence
[4] Deep Neural Network

۲-۲- راه‌حل‌های مبتنی بر خوشه‌بندی

راه‌حل‌های مبتنی بر خوشه‌بندی [۳۳-۳۱]، شباهت ویژگی بین دامنه‌ها را استخراج می‌کنند و از مقیاس شباهت برای گروه‌بندی دامنه‌ها در خوشه‌های مشخص استفاده می‌کنند. در [۳۲]، شباهت بین دامنه‌های غیر موجود را در سطح معتبر محاسبه می‌کند و سپس خوشه‌بندی سلسله مراتبی را برای شناسایی گروه‌های دامنه تولید شده توسط الگوریتم‌های تولید دامنه اتخاذ می‌کند. دقت در این روش بیش از ۰/۹۷ است. به‌طور مشابه در مرجع [۳۱]، همبستگی زمانی بین پرس و جو داده‌های DNS را بررسی می‌کند و سپس از الگوریتم‌های خوشه‌بندی XMeans با چند دامنه مخرب برای شناسایی تعداد زیادی از گروه‌های دامنه مخرب استفاده می‌کند، دقت به‌دست آمده در این روش بیش از ۰/۹۶ است.

۲-۳- راه‌حل‌های مبتنی بر نمودار

در راه‌حل‌های مبتنی بر نمودار [۳۵- ۳۴]، ابتدا رفتار دامنه‌ها از طریق گراف‌های مختلف نمونه می‌شوند، سپس از روش‌های استخراج نمودار برای شناسایی دامنه‌های مخرب استفاده می‌شود. در [۳۴]، نمودارهایی برای شکست DNS ایجاد می‌کند تا از تعامل بین میزبان نهایی و نام دامنه استفاده کند و با موفقیت گروه‌های منسجمی را از نمودارهای DNS با یک روش فاکتورسازی ماتریس سه‌گانه غیرمنفی آماری استخراج کند، با تجزیه و تحلیل در گراف‌های شکست DNS در یک ردیابی ۳ ماهه انواع فعالیت‌های غیرعادی مانند باتنت‌ها حدود ۲۵,۲ و ۲۸/۱ تروجان‌ها قابل تشخیص است. در واقع این روش ناهنجاری‌های جدید را شناسایی می‌کند. در حالی که در [۳۴]، از الگوریتم استنباط انتشار باور (BP)[1] [۳۸- ۳۶] نمودار ارتباط دامنه میزبان برای شناسایی دامنه‌های مخرب استفاده می‌شود. در این روش در طول ۷ روز آزمایش میزان نرخ مثبت واقعی ۹۰/۲ است.

با توجه به محدودیت اطلاعات ارائه شده توسط داده‌های DNS، طراحی یک طرح ارتباط با دقت بالا و پوشش خوب تبدیل به یک چالش شده است. روش‌های مبتنی بر استنباط [۳۹]، یکی از روش‌های عمده برای تحلیل داده‌های DNS و شناسایی دامنه‌های مخرب می‌باشد. ایده کلیدی روش‌های استنباط، تعریف ارتباط بین دامنه‌ها براساس ویژگی‌های استخراج‌شده از داده‌های DNS است. از الگوریتم استنباطی مبتنی بر مسیر، به‌طور خاص برای تحلیل داده‌های DNS استفاده شده است. برای نشان دادن بیشتر قدرت طرح ارتباط دامنه‌ها و همچنین بهبود کارایی استنتاج، از الگوریتم انتشار باور [۳۶]، استفاده شده است. برای ساخت گراف، دو نوع ارتباط بین دامنه‌ها براساس نتایج طبقه‌بندی‌کننده IP تعریف می‌شود. گراف حاصل از اولین نوع ارتباط‌ها G-New و گراف حاصل از نوع دوم ارتباط G-Relaxed نامیده می‌شود. دو ارتباط تعریف شده با توجه به گراف اولیه G-Baseline شکل می‌گیرد. رویکرد استفاده شده با استفاده از الگوریتم انتشار باور و مبتنی بر مسیر دارای دقت ۹۸٪ نرخ TP با کمتر از ۱٪ نرخ FP است که پیچیدگی زمانی برای اجرای الگوریتم مبتنی بر مسیر بسیار بالا و هزینه‌بر می‌باشد و استفاده از آن مقرون به صرفه نیست [۳۹].

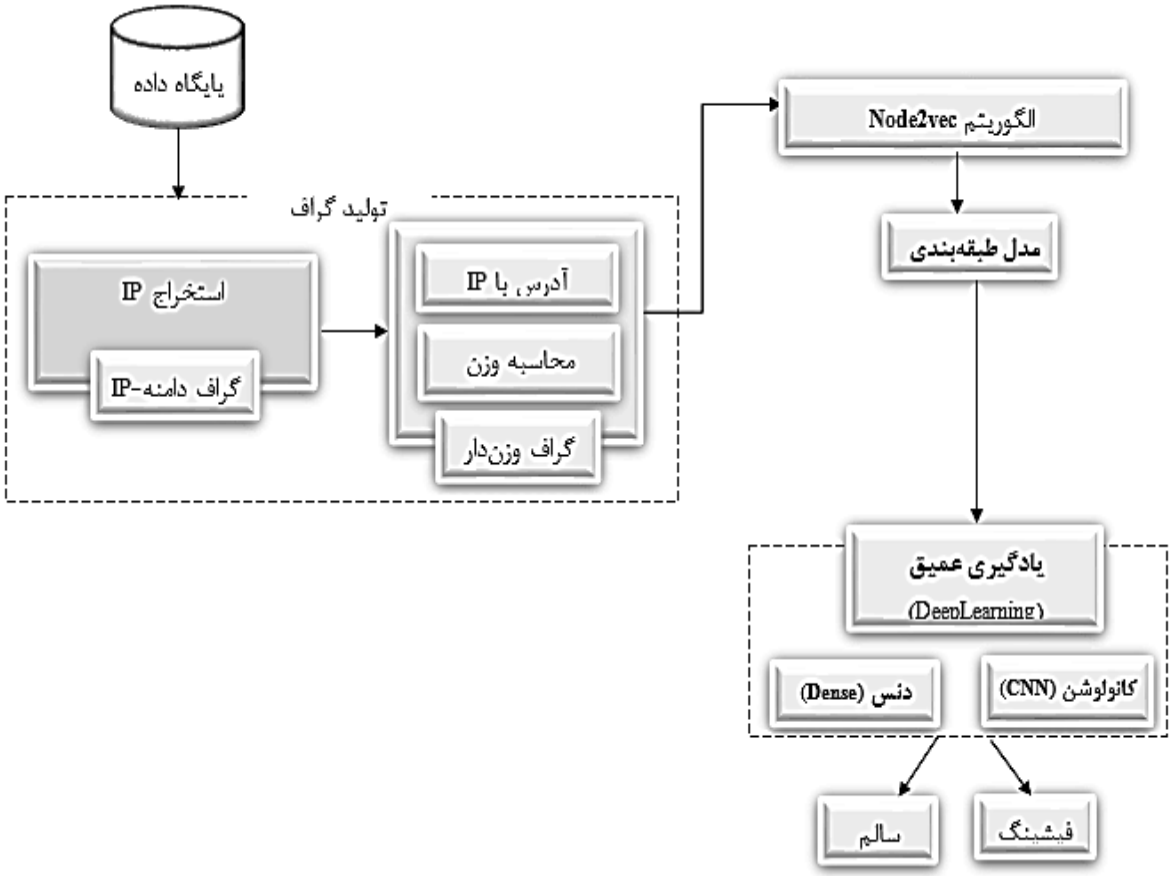
اشکال مهم روش‌های پیشین دقت پایین، استفاده از استخراج ویژگی به صورت دستی، محاسبات گران و هزینه‌بر در مقیاس داده‌های بزرگ است. لذا جهت کنترل موارد ذکر شده از یادگیری عمیق که مشکل استفاده از داده‌های زیاد و استخراج ویژگی دستی را مرتفع کرده است، با روش‌های مبتنی بر تحلیل گراف به منظور شناسایی دامنه‌های مشکوک استفاده شده است.

۳- روش پیشنهادی

در این بخش جزئیات مربوط به روش پیشنهادی برای تشخیص دامنه‌های مشکوک شرح داده می‌شود. روش پیشنهادی مورد استفاده در این مقاله یک روش مبتنی بر تحلیل گراف با استفاده از دو روش یادگیری عمیق شامل DENSE و CNN است. علت انتخاب دو نمونه ذکر شده در یادگیری عمیق یک بعدی بودن ورودی در روش پیشنهادی است، بنابراین دو نمونه DENSE و CNN با ورودی‌های تک بعدی جهت دسته‌بندی گره‌های گراف انتخاب شده است.

طرح کلی روش پیشنهادی در شکل (۲) نشان داده شده است؛ که از چهار قسمت آماده‌سازی داده، تولید گراف، الگوریتم تبدیل داده به بردار و نمونه دسته‌بندی تشکیل یافته است. در ادامه هرکدام از بخش‌ها به تفصیل توضیح داده شده است.

[1] Belief Propagation



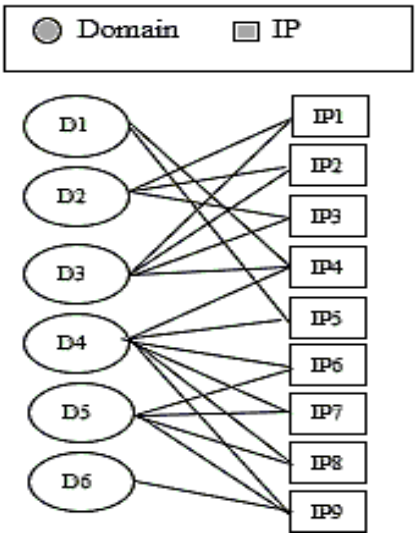
شکل (۲): طرح کلی روش پیشنهادی.

۳-۱- آماده‌سازی داده

آماده‌سازی بر روی داده‌ها همیشه یک موضوع چالشی بوده است. جهت آماده‌سازی داده، ابتدا داده‌ها (آدرس دامنه) همراه با برچسب سالم و مشکوک، به IP تبدیل می‌شوند. علت استفاده از IP، وجود ارتباطی است که بین دامنه و IP وجود دارد؛ همچنین کلاهبرداران اینترنتی عموما از IPهای عمومی برای حملات تله‌گذاری استفاده می‌کنند، بنابراین از IP به‌عنوان یک ویژگی در تشخیص دامنه‌های سالم از مشکوک همراه با دامنه استفاده می‌شود. سپس از IPهای به‌دست آمده همراه با دامنه‌های مربوط به هر IP، جهت تشکیل گراف دو بخشی دامنه -IP استفاده می‌شود.

۳-۲- گراف دو بخشی دامنه -IP

از گراف‌ها برای حل مسایل زیادی در ریاضیات و علوم رایانه استفاده می‌شود. ساختارهای زیادی را می‌توان به کمک گراف‌ها به نمایش درآورد [۴۰]. یک گراف از مجموعه‌ای گره تشکیل شده، که آن را با V نشان می‌دهیم و مجموعه‌ای نیز شامل یال‌ها است که گره‌ها را به هم وصل می‌کنند و با E نمایش داده می‌شود. این گراف با $G = (V, E)$ نشان داده می‌شود. بعد از آماده‌سازی داده، اکنون با استفاده از دامنه‌ها و IPهای ترجمه شده وارد مرحله ساخت گراف دو بخشی دامنه -IP می‌شویم. درگراف دو بخشی دامنه -IP یک گره شامل IP و گره دیگر شامل دامنه است. در این گراف هر کدام از دامنه‌ها به IPهایی که ترجمه شده است اشاره دارند. نمونه‌ای از الگو ساخت گراف دو بخشی دامنه-IP در شکل (۳) نمایش داده شده است.



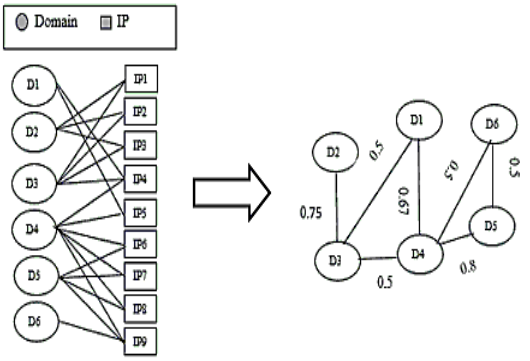
شکل (۳): گراف دو بخشی دامنه-IP.

۳-۳- گراف یک بخشی وزن‌دار

بدیهی است با تبدیل گراف دو بخشی دامنه -IP به گراف وزن‌دار، بخشی از اطلاعات موجود در گراف اولیه از بین می‌رود. در گراف دو بخشی ارتباط بین دامنه و IP مشخص است اما در گراف وزن‌دار ارتباطات یک تعداد از دامنه‌ها با IP از دست می‌رود، اطلاعات حذف شده تنها با ایجاد یال بین دامنه‌ها به‌دست نمی‌آید. لذا از مولفه وزن به منظور مشخص شدن تعداد یال استفاده می‌شود. در ادامه ساخت گراف وزن‌دار توضیح داده شده است.

برای تبدیل گراف دو بخشی دامنه -IP به گراف وزن‌دار دامنه‌هایی در نظر گرفته می‌شود که حداقل یک IP مشترک بین آن‌ها وجود دارد. در واقع برای نمایش چگونگی رابطه دامنه‌ها به یکدیگر می‌توان هر دامنه را به یک گره در گراف تبدیل کرده و در صورتی‌که در این دامنه ارتباطی با دامنه دیگر وجود داشت، یک یال از این دامنه به گره‌ای که دامنه دیگر را نمایش می‌دهد وصل می‌شود. به عبارت دیگر می‌توان گفت بین دامنه‌هایی یال رسم می‌شود که IP مشترک بین دو آدرس وجود دارد. بعد از اینکه گراف ساخته شد، گام بعدی محاسبه وزن بین دامنه‌ها است. هر چه تعداد IP بین دامنه‌ها بیشتر باشد وزن بین آن‌ها نیز بیشتر است. برای پیاده‌سازی این مفهوم از معادله (۱) که در مرجع [۱۵] برای محاسبه وزن بین دامنه‌ها استفاده کرده است، استفاده می‌شود. همان‌طور که قبل‌تر به آن اشاره شد، اطلاعات در گراف دامنه-IP مبتنی بردامنه و IP است، در تبدیل گراف دامنه-IP به گراف وزن‌دار اطلاعات IP ازبین می‌رود و عملا مشخص نیست دامنه‌ها با چه IPهایی در ارتباط بوده‌اند. برای برگرداندن این اطلاعات از وزن استفاده شده است. لذا باید رابطه‌ای استفاده شود که وقتی تعداد IPهای مشترک بین دو دامنه بیشتر است وزن بیشتری نسبت به زمانی که تعداد IP مشترک کمتری وجود دارد، اختصاص داده شود. رابطه ریاضی (معادله شماره ۱) کاملا مفهوم مورد نظر را برای این مهم پیاده‌سازی می‌نماید، لذا از این رابطه استفاده شده است. لازم به ذکر است روابط ریاضی دیگری هم برای محاسبه وزن موجود است که مفهوم موردنظر را در این مقاله برآورده نمی‌کرد بنابراین از معادله (۱) برای محاسبه وزن گراف استفاده شده است. نمونه‌ای از گراف وزن‌دار در شکل (۴) نمایش داده شده است.

$$w(d1,d2)=\begin{cases}1-\dfrac{1}{1+ip(d1)\cap ip(d2)} & \text{if } d1\neq d2\\ 1 & \text{otherwise}\end{cases} \qquad (۱)$$



شکل (۴): تبدیل گراف دوبخشی دامنه - IP به گراف وزن‌دار

۳-۴- الگوریتم Node2vec

مشکل بسیار بزرگ در گراف این است که نمی‌توان گره‌ها و یال‌ها را با عدد معرفی نموده و عملیات ریاضی را روی آن‌ها اجرا نمود [۴۲-۴۱]. برای رفع این مشکل در این مقاله از الگوریتم Node2vec استفاده شده است که هر گره باید به عددی بر روی یک فضا تبدیل شود که بتوان بر روی آن محاسبات انجام داد. بنابراین ضرورت ایجاب می‌کند تا اطلاعات به‌دست آمده در گراف وزن‌دار برای استفاده در زمینه‌های بعدی به بردار تبدیل شود [۴۲].

در الگوریتم Node2vec از قدم‌های تصادفی که بتواند دید محلی و جهانی شبکه را حفظ کند استفاده می‌شود. این الگوریتم از دو مولفه p و q که مشخص‌کننده رفتار الگوریتم هستند، استفاده می‌کند. این دو مولفه‌های مسیر قدم‌زنی، در محدوده گره اولیه باشد یا به سمت گره‌های بیرونی حرکت بکند و در هر قدم دورتر بشود را تعیین می‌کنند. همچنین در این الگوریتم از جستجوی اول سطح (BFS) و جستجوی اول عمق (DFS) برای تعریف همسایگان یک گره استفاده می‌شود [۴۲].

به وسیله این الگوریتم می‌توان داده‌های گراف را به وسیله بردار نشان داد و هر گره که دارای نشانی عددی هست عملیات‌های کلاسیک یادگیری ماشین را بر روی آن انجام داد. داده‌های تولید شده در گراف وزن‌دار که شامل گره مبدا و مقصد و وزن بین گره‌ها بود توسط این الگوریتم به بردار تبدیل می‌شود. در واقع به ازای هر گره، الگوریتم Node2vec یک بردار در خروجی ایجاد می‌کند.

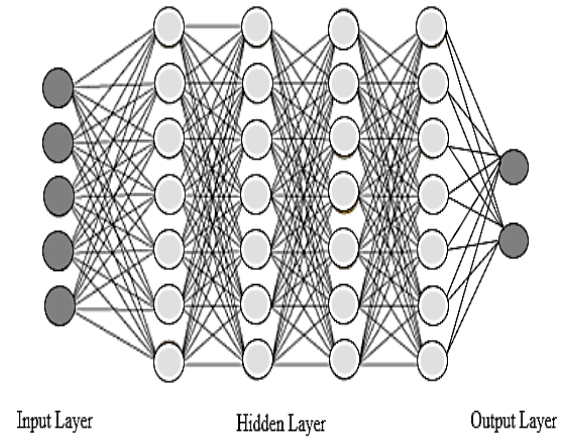
۳-۵- نمونه طبقه‌بندی

بعد از اجرای مراحل قبل، ورودی الگوریتم برای اجرای عملیات دسته‌بندی آماده شده است. در این مرحله روش‌های یادگیری عمیق شبکه DENSE و شبکه CNN برای طبقه‌بندی گره‌های گراف استفاده می‌شود.

۳-۵-۱- لایه‌های DENSE

برای دسته‌بندی مسائل در یادگیری عمیق از معماری شبکه عصبی تماماً متصل استفاده می‌شود. به همین منظور برای اضافه کردن لایه به نمونه ترتیبی، از لایه‌های استاندارد Dense استفاده می‌شود. به این لایه‌ها Dense گفته می‌شود؛ چون تمامی گره‌ها در یک لایه به تمام گره‌ها در لایه بعدی متصل می‌شوند [۴۴- ۴۳].

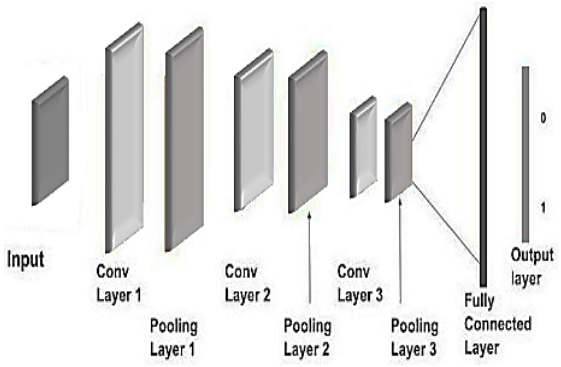
وقتی از لایه‌های Dense استفاده می‌کنیم ابتدا باید تعداد یاخته‌های عصبی در هر لایه را مشخص کنیم. تعداد این یاخته‌های عصبی یک ابر مولفه است که مقدار مناسب برای آن را می‌توان با سعی و خطا به دست آورد. به علاوه، باید توجه کرد که پیش از این که مقادیر خروجی شبکه از یک لایه به لایه دیگر منتقل شوند باید از یک تابع فعال‌سازی غیرخطی عبور کنند چون در غیر این صورت شبکه قادر به یادگیری الگوهای غیرخطی موجود در داده‌ها نخواهد بود [۴۴]. معماری DENSE را در شکل (۵) مشاهده می‌کنید.



شکل (۵): معماری DENSE [۴۴].

۳-۵-۲- شبکه عصبی کانولوشن (CNN)

شبکه‌های عصبی کانولوشن (CNN) یکی از مهم‌ترین روش‌های یادگیری عمیق هستند که در آن‌ها چندین لایه با روشی قدرتمند آموزش می‌بینند، این روش بسیار کارآمد بوده و یکی از رایج‌ترین روش‌ها در کاربردهای مختلف بینایی رایانه است. بطور کلی یک شبکه کانولوشنی از سه لایه اصلی تشکیل می‌شود که عبارتند از لایه کانولوشن، لایه pooling و لایه تماما متصل، لایه‌های مختلف وظایف مختلفی را انجام می‌دهند [۴۶-۴۵]. در شکل (۶) یک معماری کلی از شبکه عصبی کانولوشن برای دسته‌بندی بصورت لایه به لایه نمایش داده شده است.

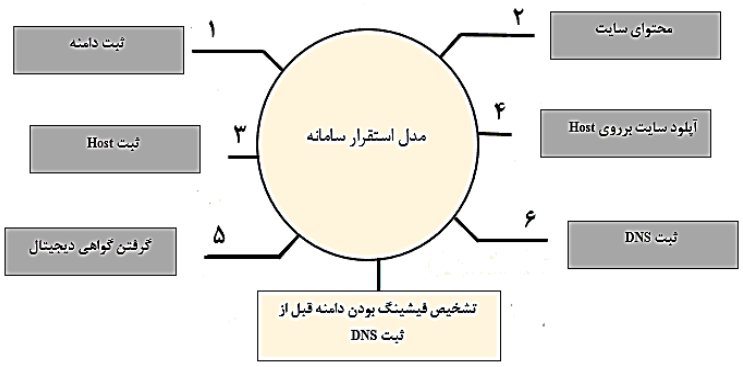


شکل (۶): معماری شبکه عصبی کانولوشن [۴۵].

۳-۶- نمونه استقرار سامانه

نمونه پیشنهادی با دریافت یک رکورد دامنه و آدرس IP آن، کلاس رکورد را تشخیص می‌دهد. برای راه‌اندازی سایت تله‌گذاری همان‌طور که در شکل (۷) نشان داده شده است، کلاهبردار اینترنتی ابتدا یک نام دامنه برای خود ثبت می‌کند. در مرحله بعد نوبت به نوشتن محتوای سایت و دریافت میزبان (Host) مربوط به این دامنه است. بعد از گرفتن میزبان در مرحله بعد کلاهبردار اینترنتی باید محتوای سایت نوشته شده را روی میزبان مورد نظر نصب کند؛ سپس به منظور راه‌اندازی این دامنه نیاز به گرفتن گواهی دیجیتال دارد تا بتواند دامنه‌های بانکی یا دامنه‌های خاصی را تله‌گذاری کند. بعد از گرفتن ثبت این گواهی از مراکز موجود، رکوردی از ثبت این گواهی در اینترنت قرار می‌گیرد و در نهایت با ثبت DNS، سایت تله‌گذاری برای کاربر قابل رویت و استفاده می‌شود. جهت تشخیص زودهنگام دامنه‌های فیشینگ، رکوردهای مربوط به ثبت گواهی دیجیتالی مورد رصد قرار می گیرد. لازم به ذکر است که این رکوردهای در اینترنت به صورت عمومی در دسترسی بوده و هر لحظه در حال به روز شدن می باشند.

رکوردهای ثبت گواهی‌های دیجیتال نیازمندی سیستم پیشنهادی در این مقاله است. در واقع تشخیص تله‌گذاری بودن رکورد موردنظر به محض دریافت گواهی دیجیتال انجام می‌پذیرد، یعنی قبل از در دسترس قرار گرفتن سایت فیشینگ، رکورد مورد نظر به سامانه دسته‌بندی داده می‌شود و توسط دسته بند ساخته شده، احتمال مشکوک یا سالم بودن آن بررسی می‌شود و در صورت تشخیص مشکوک بودن دامنه با دقت بالا امکان اعمال سیاست‌های امنیتی برقرار می‌شود. در واقع می‌توان گفت سامانه موردنظر تشخیص جرم قبل از وقوع آن را پیش‌بینی کرده است.



شکل (۷): نمونه استقرار سامانه

۴- نتایج و آزمایش‌ها

در این بخش به ارزیابی روش پیشنهادی با معیارهای رایج پرداخته شده است. در ادامه به مقایسه نتایج به‌دست آمده با سیستم پیشنهادی و دو روش موجود در کارهای انجام شده شامل الگوریتم‌های BP و SVM پرداخته می‌شود.

۴-۱- مجموعه دادگان

مجموعه دادگان شامل دامنه‌های سالم و مشکوک می‌باشد که در این مقاله از سه مجموعه دادگان با رکوردهای DNS استفاده شده است. هر مجموعه داده از تعدادی رکورد به صورت ماتریس با دو ویژگی تشکیل یافته است. این رکوردها شامل آدرس‌های دامنه و برچسب مربوط به هر دامنه می‌باشد. همان‌طور که پیش‌تر به آن اشاره شد، لازم بود به دلیل ارتباطی که بین دامنه و IP وجود دارد ابتدا دامنه‌ها به IPهای مربوطه ترجمه شوند. همچنین هر دامنه ممکن است از دو یا تعداد بیشتری IP تشکیل شده باشد. از اطلاعات به‌دست آمده در این قسمت، دامنه و IP برای به‌دست آوردن ارتباط بین دامنه‌ها با IP مشترک و محاسبه وزن استفاده می‌شود. اطلاعات هریک از سه مجموعه دادگان مورد استفاده در جدول (۱)، نمایش داده شده است. در این جدول تعداد دامنه های مشکوک، دامنه‌های سالم و تعداد کل دامنه‌ها در هریک از مجموعه دادگان بیان شده است. از آنجایی که در این زمینه کاری مجموعه دادگان استانداری وجود ندارد و به دلیل محرمانه بودن بیشتر مجموعه دادگان امکان در دسترس قرار دادن این داده‌ها وجود ندارد. لذا از سه مجموعه دادگان مجزا که هریک از این مجموعه دادگان شامل دامنه‌های سالم با برچسب صفر و دامنه‌های مشکوک با برچسب یک است، استفاده شده است. همچنین مجموعه دادگان اول شامل مجموعه دادگان ایرانی است و توسط بانک مرکزی و سازمان فناوری اطلاعات ایران گردآوری شده است در دسترس قرار گرفت. دامنه‌های مشکوک و سالم در مجموعه دادگان اول شامل سایت‌های داخلی است. مجموعه دادگان دوم نیز شامل سایت‌های خارجی است و توسط نویسنده مربوطه در دسترس قرارگرفت. مجموعه دادگان سوم برای انجام آزمایشات بیشتر با ترکیب دو مجموعه دادگان اول و دوم ساخته شده است.

همچنین مجموعه دادگان اول از دامنه‌های مشکوک ثبت شده توسط بانک مرکزی و دامنه‌های سالم از سازمان فناوری اطلاعات ایران [۴۷] گردآوری و در اختیار قرارگرفته است. مجموعه دادگان دوم از مرجع [۲۹] به‌دست آمده است. در نهایت مجموعه دادگان سوم شامل ترکیب مجموعه دادگان استفاده شده در مراجع [۴۸- ۲۹] می‌باشد.

جدول (۱): مجموعه دادگان

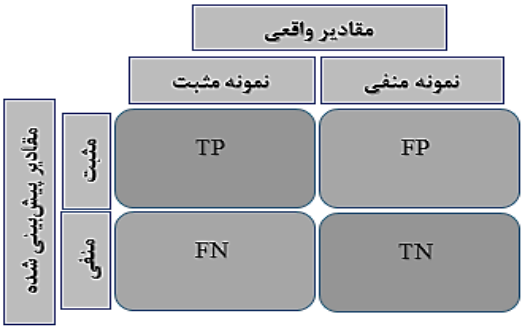
مجموعه دادگان	تعداد داده‌های سالم	تعداد داده‌های مشکوک	تعداد کل داده‌ها
مجموعه دادگان اول	۲۰،۰۱۲	۸،۱۲۱	۲۸،۱۳۳
مجموعه دادگان دوم	۱۵۵،۹۷۲	۱۰۸،۰۲۸	۲۶۴،۰۰۰
مجموعه دادگان سوم	۵۶۶،۶۲۸	۳۳۳،۳۷۲	۹۰۰،۰۰۰

۴-۲- معیارهای ارزیابی

یکی از مهم‌ترین مسائل پس از طراحی و ساخت یک نمونه یا یک روش‌های مبتنی بر دسته‌بندی، ارزیابی کارایی آن است. انتخاب معیار ارزیابی وابسته به مسئله است. برای مسائل دسته‌بندی از معیار صحت[1]، دقت[2]، بازخوانی[3] و میانگین هارمونی[4] استفاده شده است. ماتریس درهم‌ریختگی[5] مطابق شکل (۸) چگونگی عملکرد الگوریتم دسته‌بندی را با توجه به مجموعه

[1] Accuracy
[2] Pracision
[3] Recall
[4] F1-score
[5] Confusion Matrix

داده ورودی به تفکیک انواع دسته‌های مسئله دسته‌بندی نشان می‌دهد [۴۹]. معیارهای ارزیابی مورد استفاده در رابطه‌های (۵- ۲) نمایش داده شده است.



شکل (۸): ماتریس درهم‌ریختگی

$$Accuracy = (TP+TN) \div (TP+FN+FP+TN) \quad (۲)$$

$$Precision = TP \div (TP+FP) \quad (۳)$$

$$Recall = Sensitivity = (TPR) = TP \div (TP+FN) \quad (۴)$$

$$F\text{-measure} = 2 \times (Recall \times Precision) \div (Recall + Precision) \quad (۵)$$

۴-۳- محیط ارزیابی و پیاده‌سازی

زبان برنامه‌نویسی مورد استفاده Python و محیط پیاده‌سازی Spyder می‌باشد. در این پیاده‌سازی از کتابخانه‌های Keras و TensorFlow استفاده شده است[۱]. همچنین مشخصات سخت‌افزاری سیستم مورد استفاده به شرح زیر می‌باشد:

Windows 10: 64 bit

RAM: 32GB

CPU: Intel(R) COREi7 6700K 4.00GHz

HARD: 2TB

۴-۴- نتایج ارزیابی

برای ارزیابی عملکرد روش ارائه شده، دامنه‌های مشکوک را به عنوان نمونه‌های منفی و دامنه‌های سالم به عنوان مثبت قرار داده شد. همچنین صحت به عنوان معیار اصلی تعریف شده است. در ادامه به بررسی صحت نمونه پیشنهادی به کمک ماتریس درهم‌ریختگی پرداخته می‌شود. این ارزیابی با سه مجموعه دادگان که در بخش ۴-۱ به آن اشاره شد؛ انجام می‌شود. آزمایشات برروی هر مجموعه دادگان به‌طور جداگانه انجام گرفته و معیارهای ارزیابی برای هر یک از الگوریتم‌های مورد بررسی، محاسبه شده است. همچنین نمودار ROC[۲] برای هرکدام از الگوریتم‌ها به همراه مقدار AUC[۳] نیز محاسبه شده است.

جدول (۲): مجموعه دادگان اول

	Accuracy	Precision	Recall	F1-score	Auc
BP	۰/۶۵	۰/۶۳	۰/۶۴	۰/۶۳	۰/۶۷
SVM	۰/۹۷	۰/۹۷	۰/۹۶	۰/۹۶	۰/۹۴
Dense	۰/۹۹	۰/۹۹	۰/۹۹	۰/۹۹	۰/۹۹
CNN	۰/۹۹	۰/۹۹	۰/۹۹	۰/۹۹	۰/۹۹

جدول (۳): مجموعه دادگان دوم

	Accuracy	Precision	Recall	F1-score	Auc
BP	۰/۶۱	۰/۶۶	۰/۵۹	۰/۶۳	۰/۶۳
SVM	۰/۸۰	۰/۸۹	۰/۸۱	۰,۸۴	۰/۷۷
Dense	۰/۹۲	۰/۹۷	۰/۹۵	۰/۹۵	۰/۹۰
CNN	۰/۹۲	۰/۹۶	۰/۹۳	۰/۹۴	۰/۸۹

جدول (۴): مجموعه دادگان سوم

	Accuracy	Precision	Recall	F1-score	Auc
BP	۰/۳۴	۰/۱۳	۰/۱۲	۰/۱۲	۰/۴۹
SVM	۰/۴۶	۰/۵۱	۰/۴۶	۰/۴۸	۰/۵۶
Dense	۰/۸۶	۰/۳۶	۰/۵۹	۰/۴۴	۰/۸۳
CNN	۰/۸۴	۰/۴۷	۰/۵۰	۰/۴۹	۰/۸۳

نتایج به‌دست آمده در جداول (۴- ۲) و نیز مقادیر AUC نشان می‌دهد که الگوریتم BP در تشخیص و طبقه‌بندی دامنه‌های مشکوک از سالم در هر سه مجموعه دادگان به دقتی پایین‌تر از سایر نمونه‌ها دست یافته است. در مقابل SVM در مقایسه با BP بهتر عمل کرده است و صحت بالاتری دارد. نمونه پیشنهادی مبتنی بر یادگیری عمیق DENSE و CNN در مقایسه با دو الگوریتم ذکرشده در هر سه مجموعه دادگان به‌طور قابل‌توجهی صحت کلی تشخیص را بهبود بخشیده است. در توجیه عملکرد روش پیشنهادی در مقایسه با الگوریتم BP و SVM می‌توان گفت نمونه‌های یادگیری عمیق به صورت خودکار عمل استخراج ویژگی و انتخاب ویژگی‌های مناسب‌تر را انجام می‌دهند و همچنین برروی داده‌ها در مقیاس زیاد عملکرد قابل قبولی دارند که این می‌تواند در انتخاب ویژگی‌های موتر صحت را تا حد قابل قبولی بهبود بخشد. این در حالی است که دو الگوریتم BP و SVM هر چه که تعداد داده زیاد می‌شود عملکرد خوبی در انتخاب ویژگی مناسب و در نتیجه طبقه‌بندی ندارند.

لازم به ذکر است صحت به‌دست آمده روش پیشنهادی در مجموعه دادگان اول نسبت به دو مجموعه دادگان دیگر بهتر عمل کرده است، زیرا هر چه مجموعه داده گسترده شده است، مولفه‌ها مقدار کمتری نسبت به قبل پیدا کرده‌اند و صحت رو به کاهش است. همچنین در توجیه کاهش صحت می‌توان گفت؛ شناسایی ما مبتنی بر اطلاعات IP است، پس هر چقدر دسته‌بند توانسته است دسته‌بندی را صحیح انجام دهد؛ بر اساس تمایز IPها در مجموعه دادگان می‌باشد که با کاهش تمایز IP، صحت کاهش یافته است، که این به نوبه خود باعث افت شدید صحت و سایر مولفه‌های ارزیابی شده است و نمونه، قدرت تمایز و شناسایی صحیح دامنه‌ها را نداشته است.

[۱] تمامی سورس کدهای پیاده‌سازی روش پیشنهادی در لینک https://github.com/fb2021/phishingdeep منتشر شد.

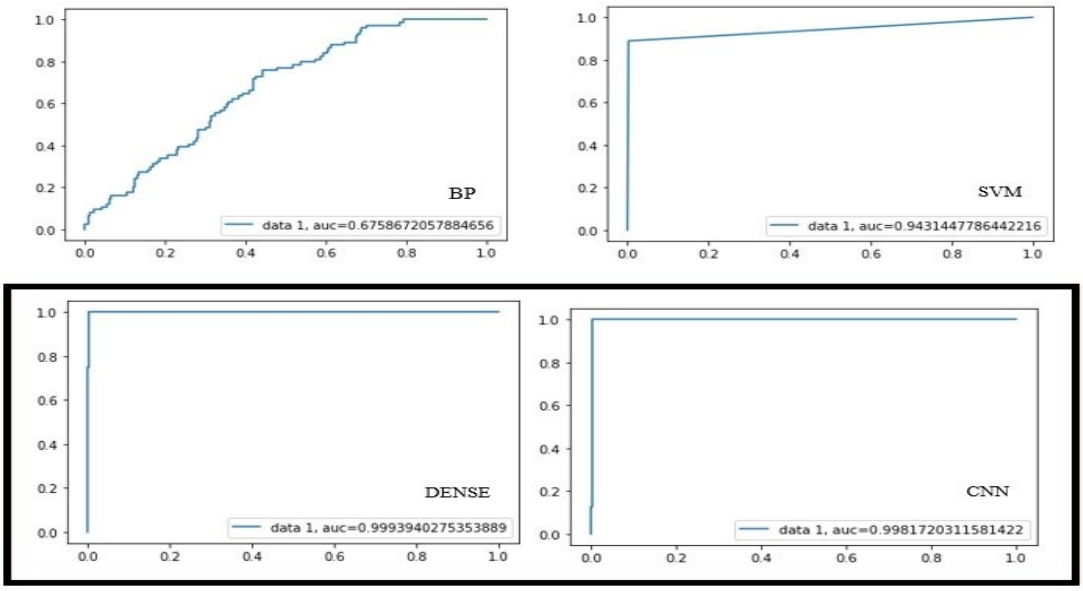
[2] Receiver Operating Characteristic

[3] Area Under The Curve

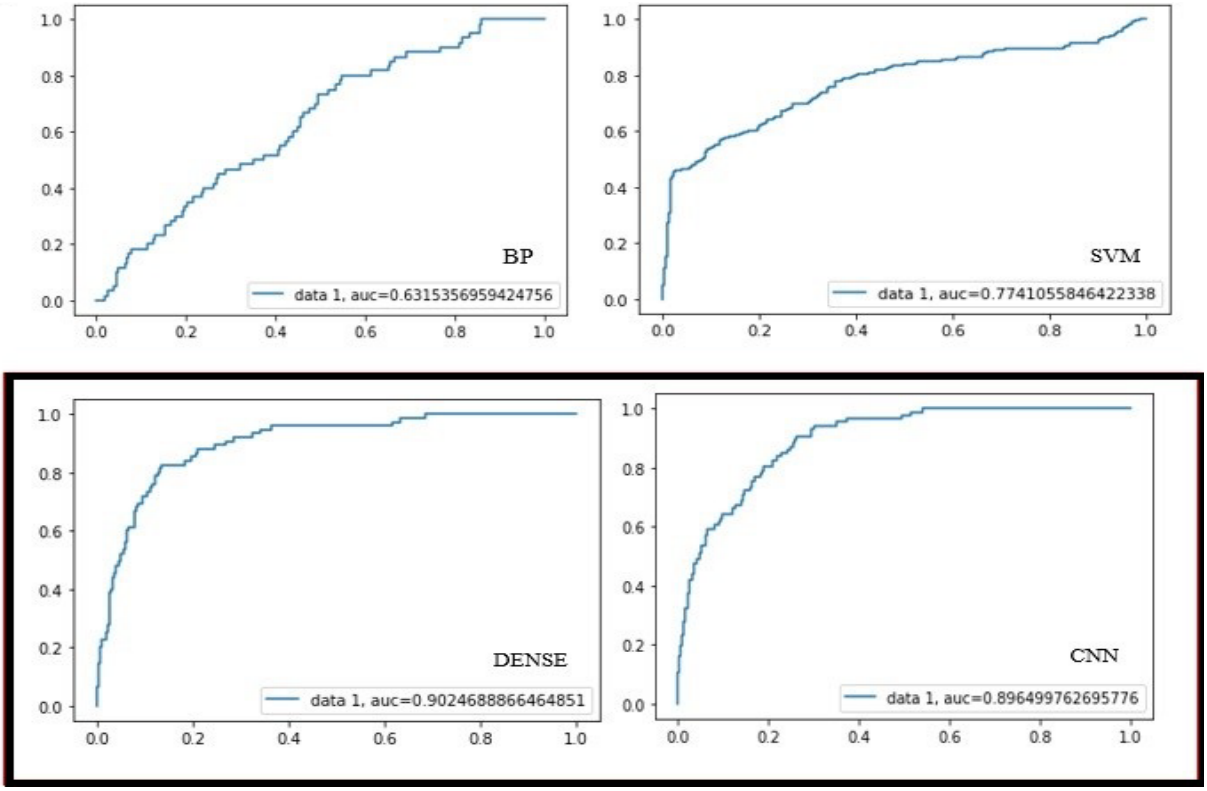
نمودارهای مربوط به ROC هر یک از مجموعه دادگان در شکل (۹)، شکل (۱۰) و شکل (۱۱) نمایش داده شده است. در مجموعه دادگان اول و مجموعه دادگان دوم نمودارهای ROC برای دو نمونه DENSE و CNN با مقادیر AUC، ۰/۹۹، ۰/۹۹، ۰/۹۲و ۰/۸۹ بهتر از دو الگوریتم SVM و BP بوده است. در این نمودارها قسمت افقی FPR نرخ تولید خطا و قسمت عمودی نشان‌دهنده TPR نرخ تولید داده‌های درست است. در این نمودار نقاطی که بالای خط نیمساز قرار گرفته‌اند مقدار حساسیت آن‌ها نسبت به FPR بیشتر است. این به آن معناست که در این بخش، نرخ مثبت صحیح بیشتر از نرخ مثبت کاذب است. قرار گرفتن نقاط در این محیط مطلوب خواهد بود. در دو مجموعه دادگان اول و دوم نمودارهای ROC بالای خط نیمساز قرار گرفته‌اند که این نشان‌دهنده کارایی روش پیشنهادی در دسته‌بندی صحیح دامنه‌ها است. در مجموعه دادگان سوم نمودار ROC دو نمونه DENSE و CNN با مقادیر AUC، ۰/۸۳ و ۰/۸۳ نیز بالای خط نیمساز قرار گرفته و توانسته است دامنه‌ها را به خوبی دسته‌بندی نماید، اما نمودار ROC برای دو الگوریتم BP و SVM با مقادیر AUC، ۰/۴۹ و ۰/۵۶ بسیار ضعیف در دسته‌بندی عمل کرده است. در BP خط منحنی زیر خط نیمساز قرار گرفته است و به این معناست که نرخ مثبت صحیح کمتر از نرخ مثبت کاذب است. قرار گرفتن نقاط در این بخش اصلاً مطلوب نیست. در مقابل خط منحنی در الگوریتم SVM تقریبا بروی خط نیمساز قرار گرفته است که مقدار عددی نرخ مثبت صحیح و نرخ مثبت کاذب با یکدیگر برابر شده است، این نقاط چندان مطلوب نیستند. همچنین در توضیح مقدار عددی AUC (سطح زیر منحنی نمودار ROC) می‌توان گفت،مقدار AUC عددی بین صفر تا یک است و نشان می‌دهد قدرت تشخیص یک آزمون چقدر است. اگر این عدد به یک نزدیک باشد به معنای آن است که داده‌ها عموماً در بالای خط نیمساز قرار گرفته‌اند و میزان نرخ مثبت صحیح بالا است در هر سه مجموعه دادگان مقادیر AUC برای روش پیشنهادی DENSE و CNN به مقدار یک نزدیک است.

در مجموعه دادگان سوم مقدار AUC در الگوریتم SVM نزدیک به ۰/۵ است که همان برابری نرخ مثبت صحیح و نرخ مثبت کاذب را نشان می‌دهد و در الگوریتم BP مقدار AUC کمتر از ۰/۵، که بیانگر بالاتر بودن نرخ مثبت کاذب در مقایسه با نرخ مثبت صحیح است. در این بخش نقاطی قرار گرفته‌اند که مقدار حساسیت آن‌ها نسبت به FPR کمتر است. در واقع هر چه نمودار به یک نزدیکتر باشد، دقت بالای نمونه را در تشخیص نشان می‌دهد[۵۳- ۵۰]. با توجه به نتایج و نمودار سطح زیر منحنی در مجموعه دادگان اول، مجموعه دادگان دوم و مجموعه دادگان سوم، روش پیشنهادی قادر به شناسایی موثر و دقیق دامنه‌های مشکوک از سالم است.

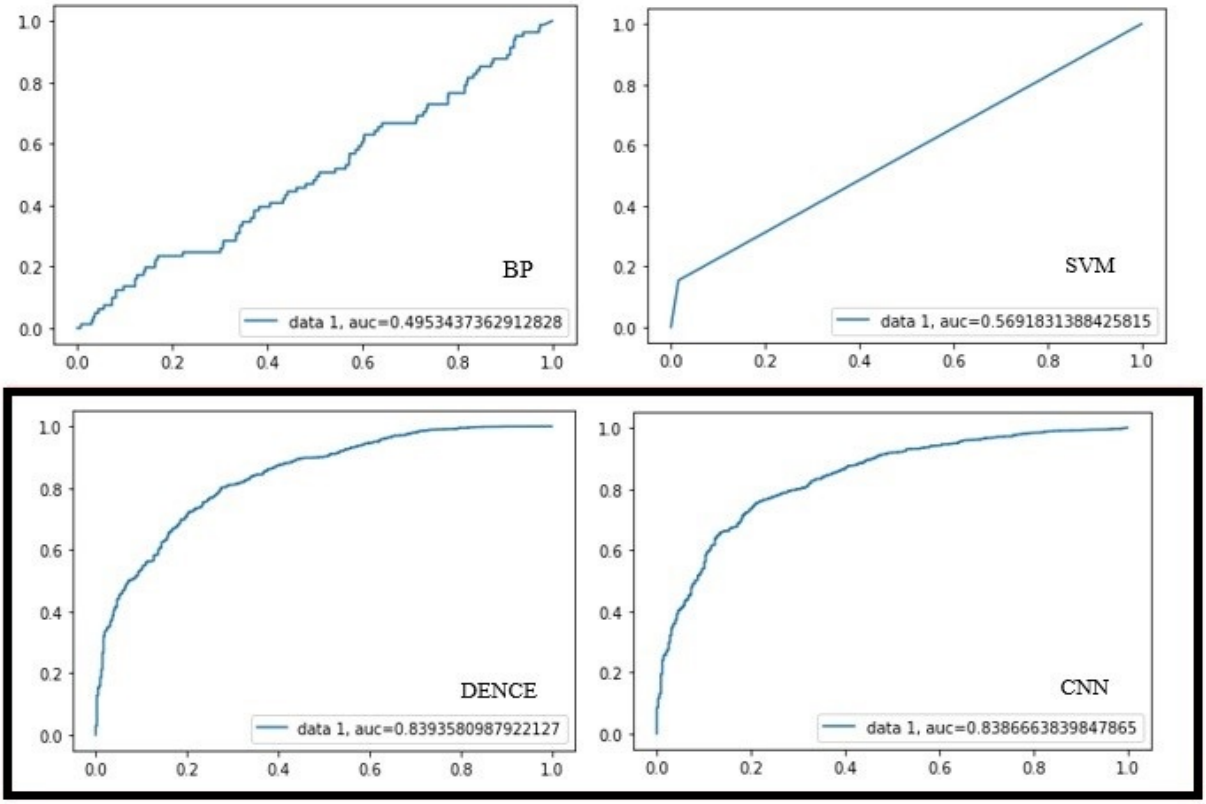
به‌طور کلی تمام معیارها حاکی از کارا بودن عملکرد روش پیشنهادی با داشتن صحت بالا در مجموعه دادگان اول را دارد. روش پیشنهادی در مجموعه دادگان اول با صحت ۹۹ درصد بالاترین صحت را در مقایسه با سایر روش‌های طبقه‌بندی به‌دست آورده است. نمودار مقایسه صحت روش پیشنهادی مربوط به سه مجموعه دادگان مورد استفاده در شکل (۱۲) آورده شده است. هریک از نمودارها در بازه زمانی مشخص بهترین صحت را برای هر الگوریتم نشان می‌دهد. کادر سیاه‌رنگ روش پیشنهادی با دو نمونه یادگیری عمیق DENSE و CNN را نشان می‌دهد.



شکل (۹): نمودار ROC مجموعه دادگان اول



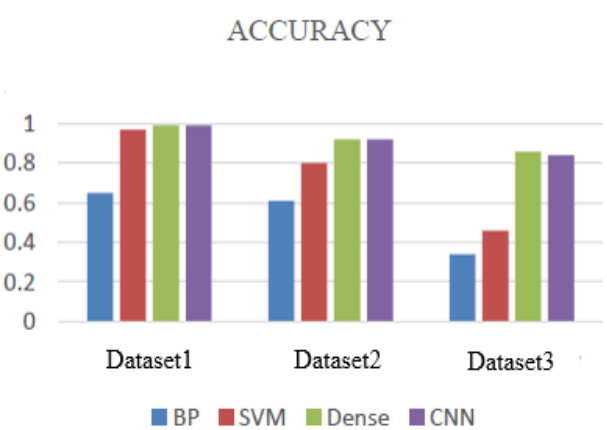
شکل (۱۰): نمودار ROC مجموعه دادگان دوم



شکل (۱۱): نمودار ROC مجموعه دادگان سوم

جدول (۵): مقایسه روش پیشنهادی با کارهای مرتبط

روش	مجموعه دادگان	دقت	ویژگی‌ها
EXPOSURE [۱۹]	داده‌های DNS و لیست دامنه‌های تولید شده توسط DGAConcker	بیش از ۵۰ درصد	استفاده از ۱۵ ویژگی و طبقه‌بندی در ۴ گروه مبتنی بر زمان، مبتنی بر TTL، مبتنی بر نام دامنه و مبتنی بر DNS Answer
ELM [۲۳]	اطلاعات دانشگاه Jiaotong شانگ‌های و DNS Query	۹۶/۲۹	استخراج ویژگی و طبقه‌بندی به ۴ دسته، مبتنی بر ساخت، مبتنی بر IP، مبتنی بر TTL و مبتنی بر WHOIS.
استفاده از یادگیری ماشین و ژنتیک [24]	۱۰۰۰۰ آدرس وب سایت	۹۷/۲۲	استفاده از نرم افزار WEKA و الگوریتم‌های طبقه‌بندی J48، SMO، JRIP ،Random Forest ،PART
LSTM, RNN [۲۸]	دامنه‌های تولید شده توسط DGA، DNS log و OSINTfeed	۰/۹۶،۰/۹۷	استفاده از یادیگیری عمیق و ابزار تحلیل‌گر شبکه Tcpdump و استفاده از Hadoop
LSTM دوجهته [۲۹]	دامنه‌های تولید شده توسط DGA و Alexa و OSINTfeed	۹۸/۹	استفاده از نمونه‌های شبکه عصبی عمیق و استفاده از Character embeddings
خوشه‌بندی [۳۰]	DNS query-response	۰/۹۶	استفاده از ویژگی‌های TTL، TLD و SLD
سیستم شناسایی دامنه [۳۱]	Passive Dns	بیش از ۰/۹۷	استفاده از مقیاس شباهت و خوشه‌بندی سلسله مراتبی
تحلیل نمودار DNS [۳۳]	داده‌های DNS	تشخیص ۲۸/۱درصد از تروجان‌ها و ۲۵/۲باتنت‌ها در طول ۳ ماه ردیابی	DNS failure subgraph
سیستم شناسایی دامنه‌های مشکوک با گزارش‌های پروکسی HTTP [۳۴]	DNS request logs, HTTP proxy logs	۹۰/۲	استخراج ویژگی از Logs و تجزیه و تحلیل پروکسیHTTP
مبتنی بر استباط [۳۸]	Active DNS	نرخ مثبت حقیقی ۰/۹۸	تعریف دو نوع ارتباط بین دامنه‌ها و ساخت سه گراف G-Baseline، G-New و G-Relaxed
مبتنی بر تحلیل گراف با یادگیری عمیق (روش پیشنهادی مقاله حاضر)	داده‌های DNS	۰/۹۹	استخراج IP، محاسبه وزن بین یال‌ها در گراف و تبدیل داده‌ها به بردار با Node2vec

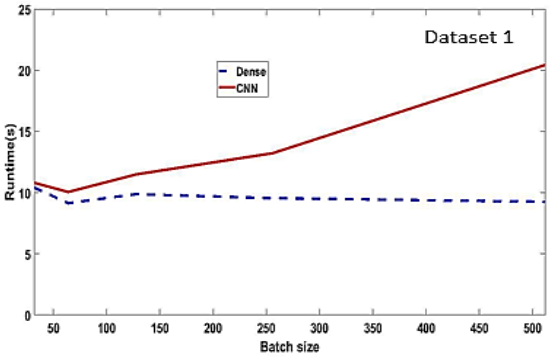


شکل (۱۲): نمودار مقایسه صحت روش پیشنهادی سه مجموعه دادگان

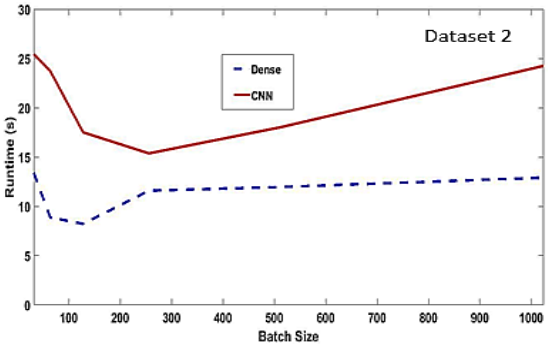
در توضیح عملکرد و کارایی بهتر شبکه DENSE از CNN می‌توان گفت با توجه به اینکه معماری DENSE شبکه‌ایی کاملا متصل است که در هر لایه اطلاعات به صورت کامل و دقیق به لایه‌های بعد منتقل می‌شود پس عملا هیچ‌گونه اطلاعات مفید در این شبکه از دست نمی‌رود و داده‌ها به صورت کامل به لایه‌های بعدی انتقال می‌یابند. در مقابل در معماری شبکه‌های کانولوشنی CNN در هرلایه عملیات کانولوشنی به همراه Maxpooling بر روی داده‌ها اعمال می‌شود که در هرلایه به دلیل اعمال عملیات مذکور تعدادی از داده‌ها به لایه بعد منتقل می‌شوند این عمل باعث می‌شود، در هرلایه تعدادی از

داده‌های مفید از بین رفته و به لایه بعد منتقل نشوند و این عمل در پایین آمدن صحت عملکرد و کارایی الگوریتم تاثیر بسزایی خواهد داشت.

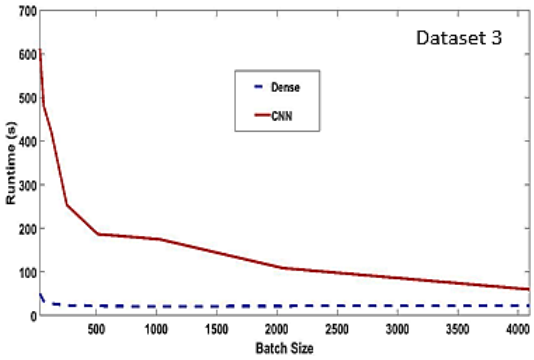
نمودارهای پیچیدگی زمانی دو نمونه DENSE و CNN در شکل‌های (۱۵- ۱۳) با Batchsizeهای مختلف نمایش داده شده است.



شکل (۱۳): نمودار مقایسه پیچیدگی زمانی DENSE و CNN بر روی دادگان اول



شکل (۱۴): نمودار مقایسه پیچیدگی زمانی DENSE و CNN بر روی دادگان دوم



شکل (۱۵): نمودار مقایسه پیچیدگی زمانی DENSE و CNN بر روی دادگان سوم

زمان اجرا در شبکه DENSE بهتر از CNN است. به دلیل عملیات CONV و POOLING برروی داده‌ها در هر لایه زمان اجرا در CNN افزایش و پیچیگی محاسباتی در CNN به مراتب از DENSE بیشتر و زمانبرتر می‌باشد. همچنین انتخاب مناسب Batchsize در زمان اجرای الگوریتم‌های موثر می‌باشد. در انتخاب این مولفه‌ها باید دقت کرد نه زیاد بزرگ باشد و خیلی هم کوچک نباشد، از طرفی اگر Batchsize خیلی بزرگ باشد، شبکه به کندی همگرا می‌شود و از طرف دیگر اگر Batchsize خیلی کوچک انتخاب شود گرادیان ناپایدار خواهد شد. بنابرین انتخاب سایز بزرگ برایBbatchsize خیلی مناسب نیست و بهتر هست که کوچک‌تر انتخاب شود، البته راه‌حل مشخصی برای محاسبه‌ی تعداد مناسب Batchsize وجود ندارد، پس بهترین راه این هست که با حدس و آزمایش Batchsize مناسب را به‌دست آورد.

۵- نتیجه‌گیری

هدف این مقاله شناسایی حملات تله‌گذاری و ارائه راه‌کارهای آن می‌باشد. در این مقاله یک سیستم تشخیص هوشمند دامنه‌های مشکوک از سالم مبتنی بر یادگیری عمیق ارائه شد. درابتدا هدف به‌دست‌آوردن IP دامنه‌ها و سپس ساختن گراف دو بخشی دامنه- IP و تبدیل آن به گراف وزن‌دار و محاسبه ارتباط بین دامنه‌ها و وزن بین آن‌ها می‌باشد. در ادامه با استفاده از الگوریتم Node2vec داده‌ها تبدیل به بردار می‌شوند و سپس عمل طبقه‌بندی با نمونه‌های یادگیری عمیق DENSE و CNN صورت می‌پذیرد.

روش پیشنهادی با ارائه دقت ۹۹٪ در مجموعه دادگان اول در تشخیص دامنه‌های مشکوک در مقایسه با الگوریتم‌های دسته‌بند دیگر BP و SVM بهتر عمل کرده است. نتایج نشان دادند نمونه‌های یادگیری عمیق یک روش شناسایی قابل اعتماد در شناسایی دامنه‌های مشکوک هستند. ضعف عمده شناسایی دامنه مشکوک پیشین در عدم توانایی آن‌ها در استخراج و ساماندهی ویژگی‌های متمایزکننده و پویا از داده‌ها است.

به‌عنوان کارهای بعدی، استفاده از محتوای سایت در کنار دامنه و IP با استفاده از الگوریتم‌های پردازش زبان طبیعی می‌تواند در استخراج ویژگی موثرتر، توانمند عمل نماید. انتخاب داده‌های مناسب کمک بسیار زیادی به بالارفتن دقت شناسایی خواهد کرد. از این رو به عنوان کار آینده می‌توان با استفاده از پردازش زبان طبیعی و متن کاوی در کنار ارتباطات دامنه-IP به منظور یافتن مولفه‌های موثرتر برروی شبکه عمیق و ترکیب یادگیری عمیق با سایر الگوریتم‌های هوشمند دقت را تا حد قابل توجهی بهبود بخشید. همچنین مقایسه روش پیشنهادی در این مقاله با کارهای مرتبط در جدول (۵) آورده شده است.

۶- تشکر و قدردانی

نویسندگان از حمایت مالی پارک علم و فناوری اطلاعات و ارتباطات از این پایان‌نامه در قالب کد اعتباری ۰۰۰۰۴۳- ۰۱- ۹۹- ۱۶ قدردانی می‌نماید.

۷- مراجع

[1] P. Gopinath, S. Sangeetha, B. Rajendran, S. Goyal, and B. S. Bindhumadhava, "Malicious Domain Detection Using Machine Learning On Domain Name Features, Host-Based Features and Web-Based Features. " Procedia Computer Science, vol. 171, pp. 654-661, 2020.

[2] S. Soholian, "E-Commerce in the Oil and Gas Industry, " Master Thesis, Arak, Mahallat Branch, Islamic Azad University, 2015. (In Persion)

[3] N, Langari and M Abdolrezzagh-Nezhad, "Phishing Website Detection for e-Banking by Inclined Planes Optimization Algorithm" Journal of Electronical & Cyber Defence, vol. 3, no. 1, pp. 29-40, 2015. (In Persion)

[4] L. Dennis and M. Shain, "Dictionary of information technology," Macmillan International Higher Education, 1988.

[5] M. Antonakakis, R. Perdisci, D. Dagon, W. Lee, and N, Feamster, "Building a dynamic reputation system for dns," In USENIX security symposium, pp. 273-290, 2010.

[6] L. Bilge, E. Kirda, C. Kruegel, and M. Balduzzi, "EXPOSURE: Finding Malicious Domains Using Passive DNS Analysis," In Ndss, pp. 1-17, 2011.

[7] B. Rahbarinia, R. Perdisci, and M. Antonakakis, "Efficient and accurate behavior-based tracking of malware-control domains in large ISP networks," ACM Transactions on Privacy and Security (TOPS), vol. 19, no. 2, pp. 1-13, 2016.

[8] S. Smadi, N. Aslam, and Li. Zhang, "Detection of online phishing email using dynamic evolving neural network based on reinforcement learning," Decision Support Systems, vol. 107, pp. 88-102, 2018.

[9] S. Gupta, A. Singhal, and A. Kapoor, "A literature survey on social engineering attacks: Phishing attack," international conference on computing, communication and automation (ICCCA), IEEE, pp. 537-540, 2016.

[10] B. Rajendran and P. Shetty, "Domain Name System (DNS) Security: Attacks Identification and Protection Methods," Proceedings of the International Conference on Security and Management (SAM), The Steering Committee of the World Congress in Computer Science, Computer Engineering and Applied Computing (World Comp), pp. 27-33, 2018.

[11] C. Y. Tejaswini Yadav, B. Rajendran, and P. Rajani, "An Approach for Determining the Health of the DNS," Sree Vidyanikethan Engineering College INDIA (IJCSMC), 2014.

[12] A. Kountouras, P. Kintis, C. Lever, Y. Chen, Y. Nadji, D. Dagon, M. Antonakakis, and R. Joffe, "Enabling network security through active DNS datasets," International Symposium on Research in Attacks, Intrusions, and Defenses, pp. 188-208, 2016.

[13] M. Antonakakis, R. Perdisci, Y. Nadji, N. Vasiloglou, S. Abu-Nimeh, W. Lee, and D. Dagon, "From throw-away traffic to bots: detecting the rise of DGA-based malware," In Presented as part of the 21st {USENIX} Security Symposium (USENIX, Security 12), pp. 491-506, 2012.

[14] J. Lee, and H. Lee, "GMAD: Graph-based Malware Activity Detection by DNS traffic analysis, " Computer Communications, vol. 49, pp. 33-47, 2014.

[15] I. Khalil, T. Yu, and B. Guan, "Discovering malicious domains through passive DNS data graph analysis, "In Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security, pp. 663-674, 2016.

[16] K. A. Messabi, M. Aldwairi, A. A. Yousif, A. Thoban, and F. Belqasmi, "Malware detection using dns records and domain name features, " In Proceedings of the 2nd International Conference on Future Networks and Distributed Systems, pp. 1-7, 2018.

[17] T. F. Yen and M. K. Reiter, "Traffic aggregation for malware detection," International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment. Springer Berlin, Heidelberg, pp. 207-227, 2008.

[18] B. Eshete, A. Villafiorita, and K. W. Binspect, "Holistic Analysis and Detection of Malicious Web Pages. Security and Privacy in Communication Networks, " 2012.

[19] L. Bilge, S. Sen, D. Balzarotti, E. Kirda, and C. Kruegel. "Exposure: A passive DNS analysis service to detect and report malicious domains. " ACM Transactions on Information and System Security (TISSEC), vol. 16, no. 4, pp. 1-28, 2014.

[20] P. Zhang, Panpan, T. Liu, Y. Zhang, J. Ya, J. Shi, and Y. Wang. "Domain watcher: detecting malicious domains based on local and global textual features. " Procedia Computer Science, vol. 108, pp. 2408-2412, 2017.

[21] Lei, Kai, Qiuai Fu, Jiake Ni, Feiyang Wang, Min Yang, and Kuai Xu, "Detecting Malicious Domains with Behavioral Modeling and Graph Embedding, " 39th International Conference on Distributed Computing Systems (ICDCS), IEEE, pp. 601-611, 2019.

[22] D. Chiba, T. Yagi, M. Akiyama, T. Shibahara, T. Yada, T. Mori, and S. Goto, "DomainProfiler: Discovering domain names abused in future, " In 2016 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), IEEE, pp. 491-502, 2016.

[23] Y. Shi, G. Chen, and J. Li, "Malicious domain name detection based on extreme machine learning, " Neural Processing Letters, vol. 48, no. 3, pp. 1347-1357, 2018.

[24] H. Akbari and M. Bagheri, "Improving the detection method of fake website based on genetic algorithms and machine learning, " Master. Thesis, Imam Hossein Comprehensive Univ, Nov. 2019.

[25] H. A. Song and S. Y. Lee, "Hierarchical Representation Using NMF, " In International conference on Neural Information Processing, Berlin, Heidelberg, pp. 466-473, 2013.

[26] J. Ahmad, H. Farman, and Z. Jan, "Deep learning methods and applications, " In Deep Learning: Convergence to Big Data Analytics. Springer, Singapore, pp. 31-42, 2019.

[27] D. Li and D. Yu, "Deep learning: methods and applications, " Foundations and trends in signal processing, vol. 7, no. 3–4, pp. 197-387, 2014.

[28] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," Computers and electronics in agriculture, vol. 147, pp. 70-90, 2018.

[29] R. Vinayakumar, K. P. Soman, and P. Poornachandran, "Detecting malicious domain names using deep learning approaches at scale, " Journal of Intelligent & Fuzzy Systems, vol. 34, no. 3, pp. 1355-1367, 2018.

[30] G. S. Josan and J. Kaur, "LSTM Network Based Malicious Domain Name Detection," International Journal of Engineering and Advanced Technology (IJEAT) ISSN, vol. 8, pp. 2249 –8958, 2019.

[31] H. Gao, Hongyu, V. Yegneswaran, J. Jiang, Y. Chen, Ph. Porras, S. Ghosh, and H. Duan, "Reexamining DNS from a

global recursive resolver perspective, ” IEEE/ACM Transactions on Networking, vol. 24, no. 1, pp. 43-57, 2014.

[32] M. Thomas and A. Mohaisen, “Kindred domains: detecting and clustering botnet domains using DNS traffic, ” In Proceedings of the 23rd International Conference on World Wide Web, pp. 707-712, 2014.

[33] T. S. Wang, H. T. Lin, W. T. Cheng, and C. Y. Chen, “DBod: Clustering and detecting DGA-based botnets using DNS traffic analysis,” Computers & Security 64, pp. 1-15, 2017.

[34] N. Jiang, J. Cao, Y. Jin, Li E. Li, and Z. Li. Zhang, “Identifying suspicious activities through dns failure graph analysis, ” In The 18th IEEE International Conference on Network Protocols, IEEE, pp. 144-153, 2010.

[35] P. K. Manadhata, S. Yadav, P. Rao, and W. Horne, “Detecting malicious domains via graph inference, ” In European Symposium on Research in Computer Security, Springer, Cham, pp. 1-18, 2014.

[36] J. S. Yedidia, W. T. Freeman, and Y. Weiss, “Understanding belief propagation and its generalizations,” Exploring artificial intelligence in the new millennium, vol. 8, pp. 236-239, 2003.

[37] M. A. Jafari Zadeh, F. Ghaffari Joghani, M. Babazadeh, and R. Bayramzadeh, “Quantum Belief Dissemination Algorithm,” Proceedings of the Iranian Physics Conference, University of Tabriz, 2015. (In Persion)

[38] MS. Leifer, and D. Poulin. “Quantum graphical models and belief propagation.” Annals of Physics, vol. 323, no. 8, pp. 1899-1946, 2008.

[39] I. M. Khalil, B. Guan, M. Nabeel, and T. Yu, “A domain is only as good as its buddies: Detecting stealthy malicious domains via graph inference, ” In Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy, pp. 330-341, 2018.

[40] B. Bollobás, “Modern graph theory, ” Springer Science & Business Media, vol. 184, 2013.

[41] H. Chen, S. F. Sultan, Y. Tian, M. Chen, and S. Skiena, “Fast and Accurate Network Embeddings via Very Sparse Random Projection, ” In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, pp. 399-408, 2019.

[42] A. Grover and J. Leskovec, “node2vec: Scalable feature learning for networks, ” In Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 855-846, 2016.

[43] Y. Zhou, Z. M. Fadlullah, B. Mao, and N. Kato, “A deep-learning-based radio resource assignment technique for 5G ultra dense networks,” IEEE Network, vol. 32, no. 6, pp. 28-34, 2018.

[44] O. Brun, Y. Yin, E. Gelenbe, Y. M. Kadioglu, J. Augusto-Gonzalez, and M. Ramos, “Deep learning with dense random neural networks for detecting attacks against iot-connected home environments,” In International ISCIS Security Workshop, Springer, Cham, pp. 79-89, 2018.

[45] I. Zafar, G. Tzanidou, R. Burton, N. Patel, and L. Araujo, “Hands-on Convolutional Neural Networks with TensorFlow: Solve Computer Vision Problems with Modeling in TensorFlow and Python,” Packt Publishing Ltd, 2018.

[46] J. Zhao, X. Mao, and L. Chen, “Speech emotion recognition using deep 1D & 2D CNN LSTM networks, ” Biomedical Signal Processing and Control, vol. 47, pp. 312-323, 2019.

[47] List of Domestic Internet Domains, Information Technology Organization of Iran, 2020. [Online], https://g2b.ito.gov.ir/ index.php/site/page/view/list_ip.

[48] Top 10 million domains, Open PageRank Initiative, 2020. [Online], https://www.domcop.com/top-10-million-domains.

[49] J. T. Townsend, “Theoretical analysis of an alphabetic confusion matrix, ” Perception & Psychophysics, vol. 9, no. 1, pp. 40-50, 1971.

[50] D. G. Altman and J. M. Bland, “Diagnostic tests 3: receiver operating characteristic plots, ” BMJ: British Medical Journal, vol. 309, no. 6948, p. 188, 1994.

[51] A. J. Bowers and X. Zhou, “Receiver operating characteristic (ROC) area under the curve (AUC): A diagnostic measure for evaluating the accuracy of predictors of education outcomes, ” Journal of Education for Students Placed at Risk (JESPAR), vol. 24, no. 1, pp. 20-24, 2019.

[52] Y. Feng, X. Shen, H. Chen, and X. Zhang, “A weighted-ROC graph based metric for image segmentation evaluation," Signal Processing, vol. 119, pp. 43-55, 2016.

[53] L. Zhang and N. Hu, “Roc analysis based condition indicator threshold optimization method, ” Prognostics and System Health Management Conference (PHM-Harbin), IEEE, pp. 1-6, 2017.