

تشخیص نفوذ در شبکه با استفاده از ترکیب شبکه‌های عصبی مصنوعی به صورت سلسله مراتبی

علی ماروسی^{۱*}، ایمان ذباح^۲، حسین عطائی خباز^۳

۱- استادیار گروه مهندسی کامپیوتر، دانشکده فنی و مهندسی، دانشگاه تربت حیدریه ۲- مربی گروه کامپیوتر، دانشگاه آزاد اسلامی تربت حیدریه ۳- دانش آموخته کارشناسی کامپیوتر، گروه مهندسی کامپیوتر، دانشکده فنی و مهندسی، دانشگاه تربت حیدریه

(دریافت: ۹۸/۰۵/۲۴، پذیرش: ۹۸/۱۰/۲۰)

چکیده

با رشد فناوری اطلاعات، امنیت شبکه به‌عنوان یکی از مباحث مهم و چالش بسیار بزرگ مطرح است. سامانه‌های تشخیص نفوذ، مؤلفه اصلی یک شبکه امن است که حملاتی را که توسط فایروال‌ها شناسایی نمی‌شود، تشخیص می‌دهد. این سامانه‌ها با داده‌های حجیم برای تحلیل مواجه هستند. بررسی مجموعه داده‌های سامانه‌های تشخیص نفوذ نشان می‌دهد که بسیاری از ویژگی‌ها، غیرمفید و یا بی‌تأثیر هستند؛ بنابراین، حذف برخی ویژگی‌ها از مجموعه به‌عنوان یک راه‌کار برای کاهش حجم سربار و در نتیجه بالا بردن سرعت سیستم تشخیص، معرفی می‌شود. برای بهبود عملکرد سیستم تشخیص نفوذ، شناخت مجموعه ویژگی بهینه برای انواع حملات ضروری است. این پژوهش علاوه بر ارائه مدلی بر اساس ترکیب شبکه‌های عصبی مصنوعی برای اولین بار به منظور تشخیص نفوذ، روشی را برای استخراج ویژگی‌های بهینه، بر روی مجموعه داده KDD CUP 99 که مجموعه داده استاندارد جهت آزمایش روش‌های تشخیص نفوذ به شبکه‌های کامپیوتری می‌باشد، ارائه می‌نماید.

کلیدواژه‌ها: شبکه‌های عصبی مصنوعی، انتخاب ویژگی، ترکیب خبره‌ها، سامانه تشخیص نفوذ

۱. مقدمه

کامپیوتری پرداخته است. همچنین در پژوهشی دیگر بهبود تشخیص نفوذ با استفاده از سامانه‌هایی مبنی بر داده کاوی ۴.5 و ماشین بردار پشتیبان ارائه شده است و دقت تشخیص ۹۶/۸٪ گزارش شده است [۴]. الگوریتم‌های دیگر مانند شبکه عصبی LVM نظارت شده [۵]، درخت تصمیم [۶]، الگوریتم‌های بیز [۷]، الگوریتم j48 [۸]، شبکه‌های عصبی مصنوعی [۹]، الگوریتم ژنتیک [۱۰]، ماشین بردار پشتیبان [۱۱]، شبکه‌های عصبی همینگ [۱۲]، شبکه‌های SOM چندگانه [۱۳]، بردار پشتیبان و کلونی مورچه‌ها [۱۴]، ترکیب نزدیک‌ترین همسایه خوشه‌بندی فازی و نظریه Damper-Shafer [۱۵]، الگوریتم PCA [۱۶] مورد توجه محققین بوده است. عموم پژوهش‌های انجام شده مذکور بر روی پایگاه داده KDD^۴ انجام شده است این پایگاه داده از طریق درگاه^۵ در اختیار محققین این حوزه قرار گرفته است. مسئله تشخیص نفوذ را می‌توان به‌عنوان یک مسئله طبقه‌بندی در نظر گرفت که توسط مجموعه‌ای از ویژگی‌ها توصیف می‌شود. طبقه‌بندی، یک نداشت از فضای مقادیر ویژگی‌ها به فضای کلاس است؛ و از آنجایی که تعداد ۴۱ ویژگی در مجموعه KDD مفروض

سامانه‌های تشخیص نفوذ^۱ (IDS) که به منظور فعالیت‌های غیرعادی در شبکه‌ها مورد استفاده قرار می‌گیرند، در سال‌های اخیر مورد توجه بسیاری از محققین قرار گرفته است و مدل‌های متعددی به منظور شناسایی و جلوگیری از حملات و برقراری امنیت ارائه و طراحی شده‌اند. اساسی‌ترین هدف امنیت اطلاعات، جلوگیری از دسترسی غیرمجاز، استفاده، افشا، اختلال، اصلاح و یا تخریب است [۱]. از آنجا که رفتار نفوذکنندگان از یک کاربر قانونی متفاوت است [۲] یک دیواره آتش سنتی به خوبی نمی‌تواند در تمام موارد به تشخیص نفوذ کمک کند در حالی که IDS می‌تواند این وظیفه را به درستی انجام دهد. در این راستا الگوریتم‌های متعددی از سوی محققین پیشنهاد شده است که می‌تواند به پژوهش [۳] اشاره کرد که با استفاده از الگوریتم خوشه‌بندی k-means و طبقه‌بندی کننده‌های ۴.5 و FNN^۲ و ماشین بردار پشتیبان (SVM^۳) به تشخیص نفوذ در سامانه‌های

* رایانه نامه نویسنده مسئول: ali.maroosi@torbath.ac.ir

1- Intrusion Detection System

2- Fuzzy Neural Network

3- Support Vector Machine

4- KDD Cup 1999 Data

5 - <http://kdd.ics.uci.edu/databases/kddcup99/>.html

این‌گونه تهدیدها پرداخت، گامی بزرگ در کاهش خسارات وارده- ی ناشی از این تهدیدها برداشته شده است.

در این پژوهش سعی شده است که ضمن بررسی و تحلیل آماری داده‌های KDD ویژگی‌های زائد و کم تأثیر در تشخیص هر یک از ۲۲ حمله شناسایی شود. لذا این مطالعه، باهدف تشخیص نفوذ در شبکه انجام شده است. که برای رسیدن به این هدف از روش جامع‌تری به نام ترکیب شبکه‌های عصبی مصنوعی یا همان ترکیب خبره‌ها استفاده شد. افزایش خبره‌ها منجر به تخصصی شدن وظیفه آن‌ها شده و ضمن تمرکز یادگیرها به توزیع خطا حول هدف کمک خواهد کرد و درنهایت موجب افزایش صحت عملکرد سیستم تصمیم‌گیر می‌شود؛ لذا با ترکیب نظرات آن‌ها حصول نتیجه‌ای دقیق‌تر نسبت به مطالعات مشابه صورت خواهد گرفت.

۲. روش تحقیق

داده‌های مورد استفاده در این پژوهش از مجموعه داده KDD، شامل سه مجموعه مستقل، استفاده شده است. قسمت اول "کل نمونه‌ها": مشتمل بر ۴۸۹۸۴۳۱ نمونه، می‌باشد. قسمت دوم: تعداد ۱۰٪ نمونه‌ها یعنی تعداد ۴۹۴۰۲۱ به‌عنوان "نمونه‌های استاندارد" نام‌گذاری شده است به‌طوری‌که شامل تمامی انواع حملات می‌شوند. و قسمت سوم: "نمونه‌های اصلاح‌شده" شامل ۳۱۱۰۲۹ رکورد انتخاب شده‌اند که اکثر محققان این نمونه‌ها را به‌عنوان مجموعه آموزشی و آزمون استفاده کرده‌اند [۲۲]. این داده‌ها توسط گروه IST^۴ از آزمایشگاه MIT LINCLON به‌عنوان اولین داده‌های استاندارد برای تشخیص نفوذ جمع‌آوری شده‌اند [۲۳]. مجموعه داده KDD از هفت هفته شبیه‌سازی فعالیت‌های کامپیوتری گردآوری شده است که دو هفته اول هیچ حمله‌ای صورت نگرفته و در پنج هفته باقیمانده بیشتر حملات آزمایش شده است [۲۴]. در اکثر پژوهش‌های انجام‌شده عمدتاً زیرمجموعه ۱۰٪ از مجموعه داده KDD برای آموزش، آزمون و انتخاب ویژگی به‌جای استفاده از کل مجموعه داده KDD مورد استفاده قرار گرفته است [۲۵]. در این پژوهش از همین مجموعه ۱۰٪ که نماینده‌ای از کل دیتا است، استفاده شده است. جدول (۱) ۴۱ ویژگی مربوط به داده‌های KDD را نشان می‌دهد. این مجموعه به ۴ گروه اصلی و ۲۲ نوع حمله تقسیم‌بندی می‌شود. جدول (۲) تقسیم‌بندی این حملات را نشان می‌دهد. در حمله‌های کلاس DOS، منابع سیستم بیش‌ازحد مورد استفاده قرار می‌گیرند و باعث می‌شود که درخواست‌های نرمال برای در اختیار گرفتن منابع، رد

است انتخاب زیرمجموعه‌ای کوچک‌تر از مجموع ویژگی‌ها یکی از مسائل موردتوجه پژوهشگران این حوزه می‌باشد. کاهش ابعاد و توجه به حذف ویژگی‌های زائد منجر به کاهش هزینه و نیز کاهش محاسبات سربار و استنتاج سریع‌تر و درنهایت بهبود درک مدل طبقه‌بند می‌شود.

مسئله انتخاب زیرمجموعه بهینه از نوع NP-Hard است که نیازمند الگوریتم‌های فرا ابتکاری^۱ [۱۷] مانند الگوریتم ژنتیک [۱۸] است. لذا دسته‌ای دیگر از مطالعات این حوزه مبتنی بر یادگیری‌های بدون ناظر، جهت حذف ویژگی‌های زائد است. در این روش‌ها، ابتدا بر اساس الگوریتم جستجوی روبه‌جلو، زیرمجموعه ویژگی‌ها انتخاب می‌شود. سپس مجموعه داده با استفاده از الگوریتم خوشه‌بندی و زیرمجموعه منتخب، بررسی شده و درنهایت، دقت خوشه‌بندی ارزیابی می‌شود. این روند چندین مرتبه تکرار تا خوشه‌بندی با بهترین معیار به دست آید. مزیت این روش این است که قابل‌اعمال بر روی تمام انواع داده‌های عددی و غیر عددی است اما به دلیل وابستگی به زیرمجموعه انتخابی اولیه همیشه بهترین مجموعه ویژگی‌ها حاصل نمی‌شود. در این خصوص می‌توان به پژوهش [۱۹] اشاره کرد که با استفاده از روش انتخاب ویژگی ChiSquaredAttributEval توانست با دقت ۱۰۰٪ گروه حملات DOS و ترافیک نرمال را تشخیص دهد. همچنین در مطالعه [۲۰-۲۱] برای رتبه‌بندی ویژگی‌ها با حذف یک ویژگی میزان کیفیت خوشه‌بندی اندازه‌گیری شده است. هر چه که حذف یک ویژگی میزان کیفیت پایین‌تری به خوشه‌بندی بدهد آن ویژگی از اهمیت بیشتری برخوردار خواهد بود. در این مطالعه برای یافتن میزان کیفیت خوشه‌بندی از معیارهای تراکم^۲ و پراکندگی^۳ استفاده شده است.

امروزه بسیاری از مردم در سراسر جهان با اتصال به اینترنت، ناخودآگاه با تهدیدات امنیتی بسیاری مانند ویروس‌ها، کرم‌ها و حملات هکرها مواجه می‌شوند. در حال حاضر فایروال‌ها، آنتی‌ویروس‌ها، رمزگذاری پیام، پروتکل‌های شبکه ایمن، حفاظت از رمز عبور و غیره برای تأمین امنیت در شبکه‌های کامپیوتری، کافی نیست، زیرا برخی از نفوذها، از نقص در سامانه‌های حفاظتی کامپیوترها برای حمله استفاده می‌کنند [۲۲]. بنابراین، سامانه‌های تشخیص نفوذ (IDS) به ضرورتی برای زیرساخت‌های امنیتی اکثر سازمان‌ها تبدیل شده است که چنانچه بتوان با استفاده از تکنیک‌های داده‌کاوی به‌دقت قابل‌توجهی از تشخیص

1- Heuristic Algorithm

2- Compactness

3- Separateness

جدول (۲): کلاس‌بندی نوع حملات

Num	Input Attribute
1	Duration
2	Protocol_Type
3	Service
4	Flag
5	Src_Bytes
6	Dst_Bytes
7	Land
8	Wrong_Fragment
9	Urgent
10	Hot
11	Num_Faild_Login
12	Logged_In
13	Num_Compromised
14	Root_Shell
15	Su_Attempted
16	Num_Root
17	Num_File_creations
18	Num_Shells
19	Num_Access_files
20	Num_Outbound_Cmds
21	Is_Host_Login
22	Is_Guest_Login
23	Count
24	Srv_Count
25	Serror_Rate
26	Srv_Serror_Rate
27	Rerror_Rate
28	Srv_Rerror_Rate
29	Same_Srv_Rate
30	Diff_Srv_Rate
31	Srv_Diff_Host_Rate
32	Dst_Host_Count
33	Dst_Host_Srv_Count
34	Dst_Host_Same_Srv_Rate
35	Dst_Host_Diff_Srv_Rate
36	Dst_Host_Same_Src_Port_Rate
37	Dst_Host_Srv_Diff_Host_Rate
38	Dst_Host_Serror_Rate
39	Dst_Host_Srv_Serror_Rate
40	Dst_Host_Rerror_Rate
41	Dst_Host_Srv_Rerror_Rate

شود. در حملات کلاس R2L، حمله‌کننده با نفوذ غیرمجاز از راه دور به ماشین قربانی، اقدام به ارسال بسته بر روی شبکه می‌کند. حمله‌های کلاس U2R، به‌طور موفقیت‌آمیزی در ماشین قربانی اجرا می‌شوند و ریشه را در اختیار می‌گیرند و در حمله‌های کلاس PROBE، حمله‌کننده با بررسی و پویش سیستم به دنبال یافتن راه‌های نفوذ می‌باشد. از میان ۴۱ ویژگی مندرج در جدول (۱)، ۳ ویژگی به‌صورت رشته‌ای هستند. برای استفاده از این ۳ ویژگی باید در مرحله پیش‌پردازش این مقادیر کدبندی شوند. ویژگی سوم (Service) طبق جدول (۳)، ویژگی دوم (Protocol) طبق جدول (۴) و ویژگی چهارم (Flag) بر اساس جدول (۵) کدبندی شده‌اند.

جدول (۱): مجموعه ویژگی‌های KDD

نام حمله	تعداد حمله	نوع حمله	کد حمله
Smurf	28079	DOS	1
Neptune	107201		2
Back	2203		3
Teardrop	797		4
Pod	264		5
Land	21		6
Normal	97278	NORMAL	7
Satan	1589	PROBE	8
Ipsweep	1247		9
Portsweep	1040		10
Nmap	231		11
Warezclient	1020		12
Guess_Pass	53	R2L	13
Warezmaster	20		14
Imap	12		15
FTP_Write	8		16
Multihop	7		17
Phf	4		18
Spy	2		19
Buffer_Overflow	30		U2R
Rootkit	10	21	
Loadmodule	9	22	
Perl	3	23	

جدول (۳): کدبندی ویژگی Service

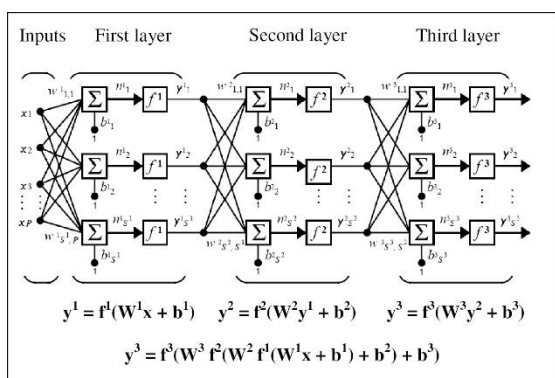
کد ویژگی	نام ویژگی	کد ویژگی	نام ویژگی
34	Netbios_s	1	Auth
35	Netstat	2	Bgp
36	Nnsp	3	Courier
37	Nntp	4	Csnet_N
38	Ntp_U	5	Ctf
39	Other	6	Daytime
40	Pm_Dum	7	Discard
41	Pop_2	8	Domain
42	Pop_3	9	Domain_U
43	Printer	10	Echo
44	private	11	Eco_I
45	Red_I	12	Ecr_I
46	RemoteJob	13	Efs
47	Rje	14	Exec
48	Shell	15	Finger
49	Sntp	16	Ftp
50	Sql_Net	17	Ftp_Data
51	Ssh	18	Gopher
52	Sunrpc	19	Hostname
53	Supdup	20	Http
54	Systat	21	Http_443
55	Tftp_U	22	Irc
56	Telnet	23	Imap4
57	Tim_I	24	Iso_Tsap
58	Time	25	Klogin
59	Urh_I	26	Kshell
60	Urp_I	27	Ldap
61	Uucp	28	Link
62	Whois	29	Login
63	Vmnet	30	Mtp
64	Uucp_Path	31	Name
65	Z39_50	32	Netbios_Dgm
66	Xll	33	Netbios_Ns

جدول (۵): کدبندی ویژگی پرچم (Flag)

کد ویژگی	نام ویژگی	کد ویژگی	نام ویژگی
7	S1	1	OTH
8	S2	2	REJ
9	S3	3	RSTO
10	SF	4	RSTOS0
11	SH	5	RSTR
-	-	6	S0

۱-۲. شبکه‌های عصبی مصنوعی

شبکه‌های عصبی مصنوعی^۱ آن دسته از سامانه‌هایی هستند که با الگوبرداری از کار مغز انسان ساخته می‌شوند. همان‌طور که مغز انسان متشکل از میلیون‌ها نورون است که توسط سیناپس‌ها متصل هستند، یک شبکه عصبی مصنوعی مجموعه‌ای از واحدهای ورودی یا خروجی متصل است که در آن هر اتصال دارای وزن مرتبط با آن واحد است. شبکه با تنظیم وزن‌ها در زمان آموزش می‌تواند برچسب کلاس ورودی را پیش‌بینی کند. به‌عبارتی دیگر وزن اتصالات تعیین می‌کند چگونه یک واحد بر دیگری اثر می‌گذارد و در نتیجه خروجی موردنظر تولید می‌شود. زیرمجموعه‌ای از این واحدها به‌عنوان گره‌های ورودی و خروجی عمل می‌کنند و گره‌های باقیمانده لایه‌های پنهان را تشکیل می‌دهند. در این پژوهش از یکی کارآمدترین ساختارهای پیشنهادی برای استفاده در مدل‌سازی به نام مدل پرسپترون چندلایه (MLP) استفاده شده است. شکل (۱) ساختار شبکه عصبی مورد استفاده در این پژوهش که از نوع ۳ لایه است را نشان می‌دهد.



شکل (۱): معماری شبکه عصبی پرسپترون چندلایه مورد استفاده در این مطالعه

تابع خروجی شبکه در لایه آخر با رابطه (۱) محاسبه می‌شود که در آن، h و o به ترتیب نشان‌دهنده لایه پنهان و لایه خروجی

جدول (۴): کدبندی ویژگی پیکربندی (Protocol)

پیکربندی	کد ویژگی
TCP	1
UDP	2
ICMP	3

موفقیت و کارایی این مدل‌ها از ماتریس آشفتگی^۲ استفاده می‌شود. برای محاسبه نرخ تشخیص در طبقه‌بندی از روابط ۳ و ۴ استفاده می‌شود:

$$DR(x_i) = \frac{T_p}{F_N + T_p} \quad (۳)$$

$$Accuracy = \frac{\sum_{i=1}^{|N|} DR(x_i) * N_i}{|N_i|} \quad (۴)$$

در رابطه (۴)، N کل نمونه‌هایی است که برای داده آزمون در نظر گرفته می‌شود. اگر در این رابطه نمونه x_i به صورت صحیح کلاس‌بندی شود مقدار $DR(x_i)$ یک و در غیر این صورت مقدار صفر را برمی‌گرداند. علاوه بر محاسبه دقت شبکه دو شاخص حساسیت Sensitivity به معنی نسبت نفوذهای کشف شده اشتباه به کل نفوذها و شاخص صحت Specificity به معنی نسبت نفوذهای کشف شده صحیح به کل افراد نیز محاسبه شده است که از طریق روابط (۵-۶) به دست می‌آید:

$$Sensitivity = \frac{TP}{TP + FN} \quad (۵)$$

$$Specificity = \frac{TN}{FP + TN} \quad (۶)$$

به طوری که:

TP: کل حملاتی که به درستی تشخیص داده شده‌اند.

FP: کل حمله‌هایی که اشتباهاً نرمال تشخیص داده شده‌اند.

TN: کل نرمال‌هایی که درست تشخیص داده شده‌اند.

FN: کل نرمال‌هایی که اشتباهاً حمله تشخیص داده شده‌اند.

پارامتر حساسیت و دقت برای هر یک از شبکه‌ها محاسبه گردید.

۲-۳. فاز اول: طراحی خبره‌ها به منظور کلاس‌بندی

نوع حمله

مدل پیشنهادی از ترکیبی از دسته‌بندی قوی برای دسته‌بندی استفاده می‌کند. این مکانیزم که اصطلاحاً ترکیب خبره‌ها نامیده می‌شود، فضای ورودی را به زیر فضاهایی تقسیم کرده و سپس هر زیر فضا را به یک دسته‌بند محول می‌کند. این زیر فضاها بر اساس بردار خروجی (Target) تقسیم‌بندی می‌شوند. به این معنی که داده‌هایی که در کلاس Normal هستند در یک زیر فضا قرار می‌گیرند و به عنوان نمونه‌هایی با طبقه ۱ و بقیه نمونه‌ها با طبقه ۲ مشخص می‌شوند. این امر باعث سهولت آموزش می‌شود. زیرا به جای آموزش الگوریتم یادگیر بر روی یک فضا با رفتار پیچیده، الگوریتم بر روی یک فضا با رفتاری ساده‌تر آموزش می‌یابد. مسلم

بوده و منظور از W همان وزن‌های لایه‌ها می‌باشد.

$$O_i = \text{sgm} \left(\sum_m \text{sgm} \left(\sum_l x_i w_{lm}^h \right) w_{mi}^o \right) \quad (۱)$$

Sgm نیز تابع سیگموئید است که به صورت رابطه (۲) تعریف می‌گردد:

$$\text{sgm}(x) = \frac{1}{1 + e^{-x}} \quad (۲)$$

این شبکه با قابلیت تقریب زنی عمومی و با الگوریتم یادگیری پس انتشار خطا^۱ در بین سایر شبکه‌های عصبی منحصربه‌فرد است.

شبکه‌های عصبی به عنوان یکی از بهترین نوع یادگیرها در بسیاری از پژوهش‌ها مورد استفاده محققین بوده است [۲۸]. اگر تعداد یادگیرها افزایش یابد می‌توان با ترکیب آن‌ها به دقت بالاتری دست پیدا کرد. در این پژوهش تلاش شده است که تعداد شبکه‌های عصبی افزایش یابد و به هر یک از شبکه‌ها یک وظیفه تخصصی واگذار گردد. این امر موجب می‌شود که حوزه تخصصی هر شبکه با شبکه دیگر متفاوت شود ضمن اینکه دقت آن افزایش یابد. به هر یک از این شبکه‌های عصبی به عنوان یادگیر، یک خبره پایه گفته می‌شود. این تنوع باعث می‌شود که نوعی از واریانس در عملکرد سیستم تشخیص نفوذ به وجود آید در نتیجه اگر خبره‌های مختلف وجود داشته و وظایف هر یک محدود شود احتمال توزیع خطا حول هدف متمرکز شده و نتایج بهتری کسب می‌شود.

به عنوان مثال، اگر وظیفه یک شبکه فقط تشخیص نفوذ و شناسایی دو نوع حمله نرمال از غیر نرمال باشد به دلیل تمرکز یادگیری دقت تشخیص آن نسبت به حالتی که باید تمامی حملات را تشخیص دهد بیشتر است. برای این که بتوان نتیجه مناسبی از ترکیب خبره‌ها گرفت باید هر یک از شبکه‌های عصبی طراحی شده شرایط ذیل را داشته باشد:

الف- هر یک از خبره‌ها به تنهایی در حد قابل قبولی (و نه کامل) دقیق باشند.

ب- هر کدام مکمل دیگری باشند یعنی نباید مشابه هم بوده و نتیجه یکسانی تولید کنند.

با در نظر گرفتن شرایط فوق طراحی شبکه‌های عصبی مختلف انجام گردید.

۲-۲. بررسی عملکرد یادگیرها در تشخیص نوع نفوذ

به طور کلی در سیستم‌های دسته‌بندی، برای بررسی میزان

باید زیرمجموعه‌ای مؤثر از ویژگی‌ها، انتخاب شود، که کار آیی قابل قبولی برای سیستم ایجاد کند. به‌منظور کاهش ویژگی، راه‌حل‌ها و الگوریتم‌های فراوانی ارائه شده است. به‌عنوان مثال کاهش ابعاد می‌تواند توسط فیلتر داده، خوشه‌بندی داده‌ها یا انتخاب ویژگی انجام شود [۲۹]. مشکل بعضی از الگوریتم‌ها این است که در زمانی که ارائه شده‌اند، بار محاسباتی زیاد به سیستم تحمیل کرده‌اند. اگرچه امروزه با ظهور کامپیوترهای سریع و حافظه‌های بزرگ، این مشکل کمرنگ شده است ولی مجموعه داده‌های بسیار بزرگ همچون KDD باعث شده‌اند که همچنان پیدا کردن یک الگوریتم سریع، برای این کار مهم باشد. در این پژوهش پس از پیش‌پردازش داده و حذف نمونه‌های تکراری، به‌منظور یافتن بهترین ویژگی، دقت هر شبکه عصبی مورد بررسی قرار گرفت. جدول (۶) تأثیر هر یک از ۴۱ ویژگی مندرج در جدول (۳) بر روی دقت شبکه عصبی را نشان می‌دهد.

جدول (۶): نتایج ساخت شبکه با یک ویژگی

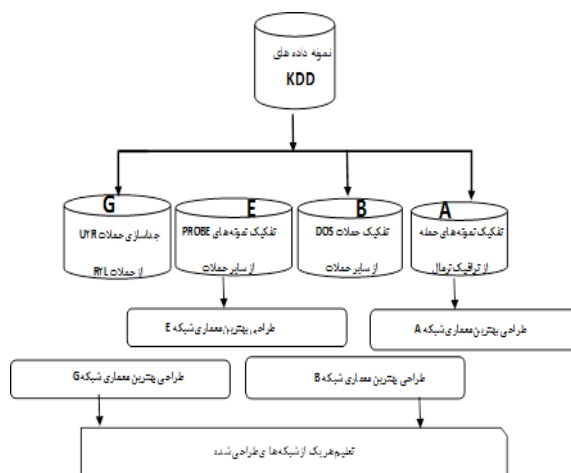
شماره ویژگی	صحت تشخیص	شماره ویژگی	صحت تشخیص
1	60.4%	22	60.3%
2	60.8%	23	95.2%
3	95.8%	24	66.8%
4	93.6%	25	89.5%
5	60.4%	26	89.3%
6	91.2%	27	64.8%
7	60.3%	28	64.7%
8	61.1%	29	96.1%
9	60.3%	30	95.9%
10	61.2%	31	61.4%
11	60.4%	32	76.3%
12	86.2%	33	92%
13	60.9%	34	92.7%
14	60.3%	35	93.1%
15	60.3%	36	77.4%
16	60.3%	37	72.9%
17	60.3%	38	89.6%
18	60.3%	39	89.3%
19	60.3%	40	64.8%
20	60.3%	41	66.2%
21	60.3%	-	-

همان‌طور که انتظار می‌رود وقتی تعداد پارامترهای ورودی شبکه‌ها افزایش می‌یابد در مجموع دقت آن‌ها نیز زیاده‌تر می‌شود. جدول (۷) این مهم را نشان می‌دهد.

است که این امر موجب افزایش کارایی (که همان افزایش دقت تشخیص است) خواهد شد.

هدف از این مرحله طراحی دسته‌بندی‌ها به‌منظور کلاس‌بندی یکی از ۵ نوع حمله U2R, R2L, Normal, DOS, Probe است. از آنجایی که جهت تعلیم دسته‌بندی باید نمونه‌ها به زیر فضای مناسبی تقسیم شوند لذا از پایگاه داده مفروض نمونه‌های Normal از سایر حملات جدا گردید. به‌طور مشابه به‌منظور تشخیص حملات نوع DOS کلیه این نمونه‌ها از سایر داده‌ها تفکیک گردید.

شکل (۲) فاز اول الگوریتم ارائه شده در این پژوهش را نشان می‌دهد. به این ترتیب پس از تفکیک نمونه‌ها به زیر فضایی مناسب در فاز اول، اقدام به طراحی ۴ شبکه عصبی A, B, E, G گردید.



شکل (۲): فاز اول الگوریتم به‌منظور طراحی و تعلیم خبره‌ها

از آنجایی که معماری شبکه عصبی از نظر تعداد لایه‌ها و تعداد نرون‌ها در هر لایه متفاوت است و می‌تواند در دقت یادگیری تأثیرگذار باشد، لذا اقدام به طراحی شبکه‌ها با معماری‌های متفاوت گردید تا بهترین معماری برای هر شبکه از نظر دقت شناسایی شود. در پایان این فاز بهترین معماری هر یک از شبکه‌ها به‌دست آمده برای استفاده در فازهای بعدی است.

۲-۴. فاز دوم: انتخاب بهترین ویژگی‌ها جهت تشخیص نفوذ

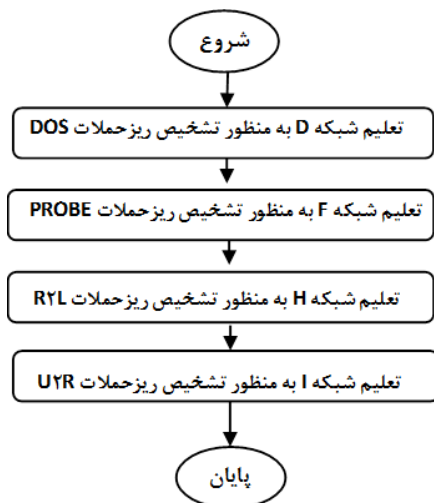
از آنجاکه میزان داده‌های ممیزی مجموعه KDD خیلی بزرگ است، بنابراین تجزیه و تحلیل آن‌ها می‌تواند تشخیص الگوهای رفتاری مشکوک را دشوارتر گرداند [۲۹] در نتیجه طراحی و پیاده‌سازی سیستم‌ها با کمترین تعداد ویژگی، ضروری به نظر می‌رسد. از طرف دیگر، توجه به این موضوع بسیار مهم است که

در این پژوهش دو نوع شبکه عصبی به‌عنوان خبره، در دو فاز مختلف طراحی شده است. ابتدا ۴ شبکه عصبی A, B, E, G با هدف کلاس‌بندی کلی ۵ نوع حمله و سپس، ۵ شبکه عصبی C, D, F, H, I به‌منظور بندی زیر حملات طراحی گردیده است. جدول (۹) بیانگر این واقعیت است که استفاده از ویژگی‌های کلیدی مجموعه داده KDD می‌تواند منجر به ساخت شبکه‌های عصبی شود که ضمن افزایش سرعت به دلیل کاهش ویژگی‌ها، دقت بالایی در تشخیص نوع حمله دارند.

جدول (۹): ویژگی‌های منتخب کلیه شبکه‌ها

شبکه	ویژگی‌های منتخب	درصد تشخیص صحیح
NETWORK A	3,4,10,29,34	%99.5
NETWORK B	3,23,35,36,38	%99.9
NETWORK C	5,7	%100
NETWORK D	2,3,14,33,34	%100
NETWORK E	3,5,6,32,37	%99.9
NETWORK F	2,5,25,36,37	%99.6
NETWORK G	1,5,13,14,23	%99.4
NETWORK H	5,6,7,11,16	%99.6
NETWORK I	5,16,19,23,24	%98.1

در تمامی مراحل فوق برای تعیین نوع زیر حمله (مندرج در جدول ۲) اقدام به تعلیم هر یک از شبکه‌های عصبی شده است. در تمام شبکه‌ها از ۷۰٪ داده‌ها جهت آموزش و ۳۰٪ جهت آزمون شبکه استفاده شده است. شکل (۳) الگوریتم این فاز را نشان می‌دهد.



شکل (۳): فاز دوم الگوریتم به‌منظور طراحی و تعلیم خبره‌ها

جدول (۷): نتایج ساخت شبکه با دو ویژگی

ویژگی‌ها	صحت تشخیص	ویژگی‌ها	صحت تشخیص
1,29	%96.2	21,29	%96.1
2,29	%97	22,29	%96.1
3,29	%97.3	23,29	%96.7
4,29	%96.8	24,29	%96.7
5,29	%96.1	25,29	%96.6
6,29	%96.1	26,29	%96.5
7,29	%96.1	27,29	%96.3
8,29	%96.8	28,29	%96.2
9,29	%96.1	29,30	%96.4
10,29	%96.9	29,31	%96.1
11,29	%96.1	29,32	%96.1
12,29	%96.1	29,33	%96.4
13,29	%96.7	29,34	%96.4
14,29	%96.1	29,35	%96.5
15,29	%96	29,36	%96.4
16,29	%96.1	29,37	%96.5
17,29	%96.1	29,38	%96.6
18,29	%96.1	29,39	%96.5
19,29	%96.1	29,40	%96.2
20,29	%96.1	29,41	%96.1

جدول فوق نشان می‌دهد ویژگی‌های استفاده از ویژگی‌های مختلف در دقت شبکه تأثیرگذار است. همین مراحل تا انتخاب ۵ ویژگی تکرار شده است که نتایج به‌عنوان نمونه برای شبکه اول در جدول (۸) نشان داده شده است.

جدول (۸): مؤثرترین ویژگی‌ها در ساخت شبکه A

تعداد ویژگی	شماره ویژگی‌ها	درصد تشخیص صحیح
1	29	96/1%
2	3,29	97/3%
3	3,29,34	98/5%
4	3,10,29,34	99/2%
5	3,4,10,29,34	99/5%

طراحی شده در فازهای ۱ و ۲ را به صورت سلسله‌مراتبی با یکدیگر ترکیب می‌کنیم. در مجموع ۹ شبکه عصبی طراحی شد. شکل (۴) الگوریتم ارائه شده به روش سلسله‌مراتبی را نشان می‌دهد.

به‌عنوان مثال "NETWORK A" وظیفه جداسازی ترافیک نرمال را بر عهده دارد. به عبارتی اگر رکورد آزمایشی وارد شده به سیستم از نوع نرمال باشد به‌عنوان ترافیک نرمال به خروجی فرستاده خواهد شد، در غیر این صورت به‌عنوان حمله شناسایی شده و به شبکه NETWORK B تحویل داده می‌شود. "NETWORK B" گروه حملات DOS را از دیگر گروه‌ها جدا می‌کند چراکه شبکه در فاز اول آموزش داده شده و اکنون به‌عنوان یک خیره با دقت بالا نوع حملات DOS را شناسایی می‌کند. اگر رکورد وارد شده به این شبکه از گروه حملات DOS باشد برای جداسازی نوع حمله به "NETWORK C" و در نهایت به "NETWORK D" خواهد رفت. به این ترتیب الگوی تشخیص هر رفتار به شبکه عصبی مخصوص به آن واگذار می‌گردد.

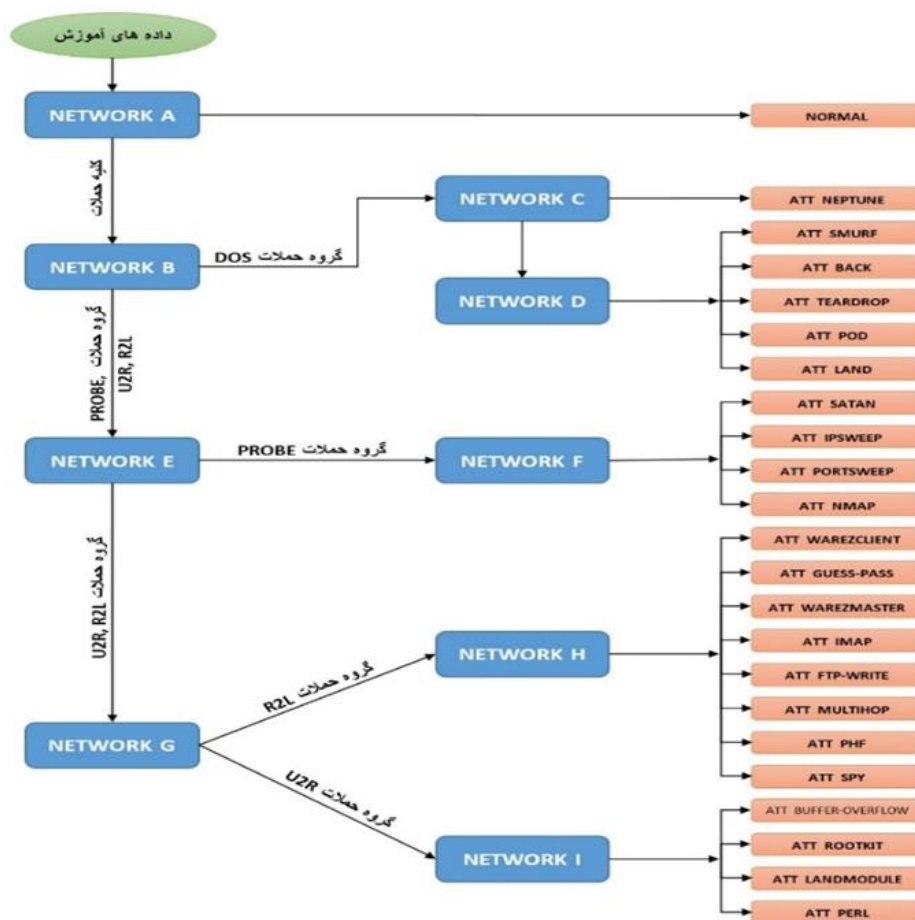
۵-۲. فاز سوم، ترکیب خبره‌ها

روش‌های مختلفی برای ترکیب کردن دسته‌بندها وجود دارد. که دو نمونه از مهم‌ترین آن‌ها عبارت‌اند از:

الگوریتم Bagging: در این الگوریتم تعدادی دسته‌بند یکسان و ناپایدار (مانند درخت تصمیم یک سطحی) با یکدیگر ترکیب می‌شوند. خروجی الگوریتم، رأی‌گیری اکثریت است. در این الگوریتم پایگاه داده به چند قسمت مجزا تبدیل می‌شود و هر قسمت به یکی از طبقه‌بندها واگذار می‌شود.

الگوریتم Boosting: مانند الگوریتم قبلی است با این تفاوت که دسته‌بندها به صورت سری قرار می‌گیرند. یعنی ورودی هر دسته‌بند تحت تأثیر خروجی دیگر دسته‌بندهای قبل از خودش است. به هر دو روش فوق، اصطلاحاً خرد جمعی گفته می‌شود. بدین معنا که از دسته‌بندهای ساده استفاده می‌شود.

در این پژوهش از الگوریتم boosting به منظور تشخیص نفوذ در شبکه استفاده شده است. در این مرحله شبکه‌های عصبی



شکل (۴): مراحل تشخیص و جداسازی حملات به روش سلسله‌مراتبی

۳. نتایج و بحث

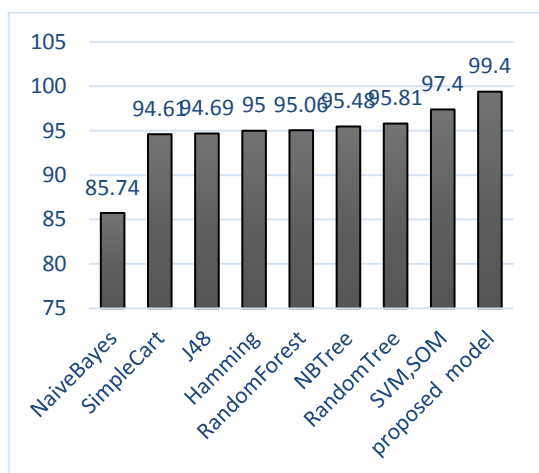
این مطالعه با هدف تشخیص نفوذ در شبکه صورت گرفته است. بدین منظور به جای استفاده از یک سیستم یادگیر شبکه عصبی سعی شده است که از چندین سیستم یادگیر استفاده شود و یادگیرها بتوانند با مشارکت یکدیگر نتایج دقیق‌تری را دست آورند. داده‌های مورد استفاده در این مطالعه از مرجع داده KDD استفاده شده است که مورد توجه بسیاری از محققین این حوزه بوده است. به عنوان نمونه می‌توان به مطالعات [۲۹-۳۲] که در سال‌های ۲۰۱۸ و ۲۰۱۹ بر روی همین دیتاست انجام شده است اشاره نمود.

استفاده از ترکیب یادگیرها در مطالعات مختلفی مشاهده می‌شود که از این بین می‌توان به مطالعه گاش و همکاران [۳۳] اشاره کرد که با استفاده از الگوریتم خوشه‌بندی grid و الگوریتم دسته‌بندی SVM به مدل‌سازی رفتار نرمال در سیستم تشخیص نفوذ در شبکه پرداختند. به این صورت که اگر رفتار کاربر در هیچ‌یک از خوشه‌های اصلی حمله قرار نگیرد آن را به عنوان یک رفتار نرمال کلاس‌بندی می‌کند. عیب مهم این روش عدم تشخیص رفتارهای جدیدی است که در هیچ‌یک از خوشه‌ها قرار نمی‌گیرد. در مطالعه دیگر هانگ و همکاران [۳۴] با استفاده از خوشه‌بندی C-mean و شبکه عصبی RBF به طراحی یک سیستم IDS پرداختند. هورن و همکاران [۳۵] نیز از درخت تصمیم و شبکه بیزین استفاده نمودند. در تمامی موارد ذکر شده اگرچه از ترکیب روش‌های یادگیر استفاده شده است لیکن در نهایت جمع نظر خبره‌ها به روش رأی اکثریت صورت گرفته است، در حالی که در روش ارائه شده در این مقاله تجمیع نظرات خبره‌ها به طور سلسله مراتبی و نه به صورت موازی، انجام شده است. ضمن این‌که تمامی یادگیرها از یک جنس بوده (شبکه عصبی مصنوعی) و تعلیم آن‌ها به گونه‌ای انجام شده است که با حداقل ویژگی‌ها بالاترین نرخ یادگیری را به همراه داشته باشد.

اگرچه یک شبکه عصبی به دلیل پردازش موازی بسیار مقاوم است و با تنظیم درست تعداد لایه‌های مخفی و تعداد نرون‌های مناسب توانایی تقریب‌زنی هر تابع غیرخطی را دارد؛ اما اگر بخواهیم دقت یک طبقه‌بند را افزایش دهیم می‌توانیم تعداد کلاس‌ها را کاهش و در مقابل تعداد یادگیرها را افزایش دهیم. اگر تعداد یادگیرها (خبره‌ها) افزایش یابد می‌توانیم با ترکیب نتایج بازم به دقت بالاتری دست پیدا کنیم. اساس روش ترکیب شبکه‌ها به روش سلسله مراتبی ترکیب آن‌ها به صورت سری است با این هدف که وظایف هر یک از طبقه‌بندها را محدود کنیم، بدون این‌که در نتیجه کار خللی وارد شود. همان‌طور که یک پزشک متخصص ابتدا به تفکیک بیمار از سالم پرداخته و سپس

به بررسی دقیق‌تر نوع بیماری می‌پردازد.

الگوریتم ارائه شده در این پژوهش از این جهت حائز اهمیت است که با طراحی خبره‌های مختلف اقدام به آموزش هر یک در حوزه تخصص خود می‌نماید. این فرایند تا زمانی که خطای شبکه‌ها به حداقل برسد ادامه می‌یابد. به این ترتیب تمرکز هر یک از شبکه‌ها بر روی وظیفه خود بیشتر می‌شود. سپس با اتصال سری هر یک از شبکه‌ها به یکدیگر فرایند تشخیص نفوذ شکل می‌گیرد. معماری تمامی شبکه‌ها از نوع پرسپترون چندلایه در نظر گرفته شده است جهت اطمینان از صحت عملکرد شبکه‌ها میانگین نتایج بعد از ۳۰ بار آزمون ثبت شده است. سیستم تشخیص نفوذ سلسله مراتبی پیشنهاد شده در این مقاله دقتی برابر ۹۹/۴٪ در تشخیص نوع حملات بر روی این مجموعه داده دارد که از بین مدل‌های ارائه شده در سایر پژوهش‌ها برتری دارد. شکل (۵) دقت مدل ارائه شده در این مطالعه نسبت به مطالعات مشابه را نشان می‌دهد.



شکل (۵): مقایسه مدل ارائه شده با سایر مدل‌ها

۴. نتیجه‌گیری

در مطالعه حاضر علت بهبود نتیجه را می‌توان این‌گونه توصیف کرد که: اگرچه هر یک از شبکه‌های عصبی می‌توانند به تنهایی به پیش‌بینی نوع حمله بپردازند، اما وقتی تعداد ورودی‌های این شبکه‌ها که با انتخاب مناسب ویژگی‌ها تعیین می‌گردد، کاهش یابد باعث تمرکز بیشتر شبکه و در نتیجه افزایش دقت طبقه‌بندی می‌شود. همچنین وقتی وظیفه یک طبقه‌بند تعیین ۲ کلاس به جای ۲۳ کلاس باشد به یقین صحت عملکرد بالاتری خواهد داشت. به این ترتیب به جای اینکه یک شبکه طراحی شود که ۲۳ کلاس را طبقه‌بندی می‌کند، شبکه‌های متفاوتی از حیث معماری و انجام وظیفه طراحی می‌شود که در حوزه طبقه‌بندی خود متخصص باشند. برای هر طبقه‌بندی پایه، توزیع نمونه‌های

- genetic algorithm,” arXiv Prepr arXiv12041336, 2012.
- [11] Xu. Xin and W. Xuening, “An adaptive network intrusion detection method based on PCA and support vector machines,” In International Conference on Advanced Data Mining and Applications, pp. 696-703, 2005.
- [12] R. Naoum, A. L. Abdullah, and Sh. Marwan, “A Hybrid Intrusion Detection System Using Hamming and MAXNET Neural Nets Using NDIS Dataset,” Journal of Emerging Trends in Computing and Information Sciences, vol. 4, pp. 198-203, 2013.
- [13] B. C. Rhodes, A. James. Mahaffey, and D. James, “Multiple self-organizing maps for intrusion detection,” In Proceedings of the 23rd national information systems security conference, pp. 16-19, 2000.
- [14] J. Feng, Y. Sui, and C. Cao, “An incremental decision tree algorithm based on rough sets and its application in intrusion detection,” Artificial Intelligence Review 40, vol. 40, pp. 517-530, 2013.
- [15] C. hou, Te. Shun, Kang K. Yen, and L. Jun, “Network intrusion detection design using feature selection of soft computing paradigms,” International journal of computational intelligence, 2008.
- [16] I. Ahmad, A. B. Abdulah, A. S. Alghamdi, K. Alnafjan, and M. Hussain, “Feature subset selection for network intrusion detection mechanism using genetic eigen vectors,” In: Proceedings of 2011 International Conference on Telecommunication Technology and Applications (ICTTA 2011), 2011.
- [17] F. López, G. T. Miguel, B. Belén, A. Moreno Pérez, and J. Marcos, “Solving feature subset selection problem by a parallel scatter search,” European Journal of Operational Research, vol. 169, pp. 477-489, 2006.
- [18] J. Yang and H. Vasant, “Feature subset selection using a genetic algorithm,” In Feature extraction, construction and selection, Springer, Boston, MA, pp. 117-136, 1998.
- [19] H. Nama and A. Seyyed, “Application of data mining techniques to detect computer network penetration,” The first international conference on the new achievements in electrical engineering and computer science, 2010. (In persian)
- [20] M. J. Asbagh and H. Abolhassani, “Feature-Based Data Stream Clustering,” In: Computer and Information Science, ICIS 2009 Eighth IEEE/ACIS International Conference on. IEEE, 2009.
- [21] M. Dash, K. Choi, P. Scheuermann, and H. Liu, “Feature selection for clustering-a filter solution,” In: Data Mining, ICDM Proceedings 2002 IEEE International Conference on. IEEE, 2002.
- [22] M. D. Hasan, M. AlMehedi, N. Mohammed, A. Shamim, and I. Khademul, “Feature selection for intrusion detection using random forest,” Journal of information security, vol.7, pp. 129-140, 2016.

ورودی برای آموزش در جهتی تغییر داده می‌شود که طبقه‌بندی‌ها بر روی نمونه‌ها سخت‌تر متمرکز شوند و در نهایت هر یک از شبکه‌ها بهتر تعلیم داده شوند. مدل‌های داده‌کاوی می‌تواند به‌عنوان ابزاری قدرتمند در تشخیص نفوذ جهت پیش‌گویی نوع حمله در شبکه‌ها مورد استفاده قرار گیرند. مدل ارائه‌شده در این مطالعه می‌تواند در کنار فایروال‌ها به‌عنوان یک دستیار نرم‌افزاری به‌کار گرفته شود.

۵. تقدیر و تشکر:

این مقاله مستخرج از نتایج طرح تحقیقاتی اجرا شده به شماره طرح ۵۰ از محل اعتبارات معاونت آموزشی و پژوهشی دانشگاه تربت حیدریه می‌باشد.

۶. مراجع

- [1] S. Ganapathy, K. Kulothungan, S. Muthurajkumar, M. Vijayalakshmi P. Yogesh, and A. Kannan, “Intelligent feature selection and classification techniques for intrusion detection in networks: a survey,” EURASIP Journal on Wireless Communications and Networking, vol. 1, pp. 271-291, 2013.
- [2] W. Stallings, “Cryptography and network security: principles and practices,” Pearson Education India, 2006.
- [3] M. Solanki and D. Vidya, “Intrusion Detection System by using K-Means clustering C 4.5 FNN SVM classifier,” Int. J. Emerg. Trends Technol. Comput, vol. 3, pp. 6-16, 2014.
- [4] V. Kosamkar and S. Sangita, “Improved Intrusion detection system using C4. 5 decision tree and support vector machine,” PhD diss., Doctoral dissertation, Mumbai University, 2013.
- [5] J. Li, Y. Liu, and L. Gu, “DDoS attack detection based on neural network,” In: Aware Computing (ISAC), 2nd International Symposium on. IEEE, 2010.
- [6] A. Balon-Perin and G. Björn, “Ensembles of decision trees for network intrusion detection systems,” International Journal on Advances in Security, 2013.
- [7] D. M. Farid, H. Nouria, and Z. Mohammad, “Combining naive bayes and decision tree for adaptive intrusion detection,” arXiv preprint arXiv:1005.4496, 2010.
- [8] M. Tavallae, E. Bagheri, W. Lu, and A. Ghorbani, “A detailed analysis of the KDD CUP 99 data set,” In 2009 IEEE symposium on computational intelligence for security and defense applications, IEEE, 2009.
- [9] J. Cannady, “Artificial neural networks for misuse detection,” In: National information systems security conference, 1998.
- [10] M. S. Hoque, M. Mukit, M. Bikas, and A. Naser, “An implementation of intrusion detection system using

- [30] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A Deep Learning Approach to Network Intrusion Detection," In *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 1, pp. 41-50, 2018.
- [31] K. Siddique, Z. Akhtar, F. Aslam Khan, and Y. Kim, "KDD Cup 99 Data Sets: A Perspective on the Role of Data Sets in Network Intrusion Detection Research," In *Computer*, vol. 52, no. 2, pp. 41-51, 2019.
- [32] O. Rashid, Z. Othman, and S. Zainudin, "Features Selection for Intrusion Detection System Based on DNA Encoding," In: *Intelligent and Interactive Computing, Lecture Notes in Networks and Systems*, Springer, vol. 67, 2019.
- [33] A. K. Ghosh, C. Michael, and M. Schatz, "A real-time intrusion detection system based on learning program behavior," In: *Proceedings of the hird International Workshop on Recent Advances in Intrusion Detection Toulouse, France, 2000*.
- [34] L. Hung-Jen, L. Chun-Hung, L. Ying-Chih, and T. Kuang-Yuan, "Intrusion detection system: A comprehensive review," *Journal of Network and Computer Applications*, vol. 36, pp. 16-24, 2013.
- [35] S. Horng, M. Su, Y. Chen, T. Kao, R. Chen, and J. Lai, "A novel intrusion detection system based on hierarchical clustering and support vector machines," *Expert Syst. Appl.*, vol. 38, no. 1, pp. 306-313, 2011.
- [23] A. Das and S. Siva Sathya, "Association Rule Mining For Kdd Intrusion Detection Data Set," *International Journal Of Computer Science And Informatics Issn (PRINT)*, pp. 2231-5292, 2012.
- [24] A. Özgür and H. Erdem, "The impact of using large training data set KDD99 on classification accuracy," *Peer J. Prepr.*, vol. 5, pp. 283-287, 2017.
- [25] A. Ghadiri and N. Ghadiri, "An adaptive hybrid architecture for intrusion detection based on fuzzy clustering and RBF neural networks," In: *Communication Networks and Services Research Conference (CNSR), Ninth Annual. IEEE, 2011*.
- [26] Gharehchopogh, F. Soleimani, M. Molany, and F. Dabaghchi Mokri, "Using artificial neural network in diagnosis of thyroid disease: a case study," *International Journal on Computational Sciences & Applications (IJCSA)*, 2013.
- [27] Y. Chen, A. Ajith, and Ju. Yang, "Feature selection and intrusion detection using hybrid flexible neural tree," In *International Symposium on Neural Networks*, Springer, Berlin, Heidelberg, 2005.
- [28] Rafiqul, et al., "Classification of malware based on integrated static and dynamic features," *Journal of Network and Computer Applications*, vol. 36, no. 2, pp. 646-656, 2013.
- [29] Z. Amirkhani, M. Madani, M. H. Sadipour, and S. Sadat, "Increasing location accuracy in neural network based wireless communications systems," *Cyber Defense and Cyber Defense*, vol. 3, pp. 31-38, 1394. (In persian)

Network Intrusion Detection using a Combination of Artificial Neural Networks in a Hierarchical Manner

A. Maroosi*, E. Zabbah, H. Ataei Khabbaz

*Department of Computer Engineering, University of Torbat Heydarieh, Torbat Heydarieh, Iran
(Received: 15/08/2019, Accepted: 10/01/2020)

ABSTRACT

With the growth of information technology, network security is one of the major issues and a great challenge. Intrusion detection systems, are the main component of a secure network that detect the attacks which are not detected by firewalls. These systems have a huge load of data to analyze. Investigations show that many features are unhelpful or ineffective, so removing some of these redundant features from the feature set is a solution to reduce the amount of data and thus increase the speed of the detection system. To improve the performance of the intrusion detection system it is essential to understand the optimal property set for all kinds of attacks. This research, in addition to presenting a method for intrusion detection based on combining neural networks, also introduces a method for extracting optimal features of the KDD CUP 99 dataset which is a standard dataset for testing computer networks intrusion detection methods.

Keywords: Artificial Neural Networks, Feature Selection, Mixture of Experts, Intrusion Detection System