





دانشگاه الزهرا (س)

دانشکده فنی و مهندسی

پایان نامه

جهت اخذ درجه کارشناسی ارشد

رشته مهندسی کامپیوتر - گرایش هوش مصنوعی

عنوان

ارائه ی یک سیستم تشخیص نفوذ هوشمند، مبتنی بر رویکرد ماشین مجازی

استاد راهنما

دکتر رضا عزمی

دانشجو

بشری پیشگو

اسفند ماه سال ۱۳۹۰

کلیه دستاوردهای این تحقیق متعلق
به دانشگاه الزهرا (س) می باشد.

تقدیم

تقدیرم جی تخلصی را فی کہ ہر رندی شان آرزوی من است.

بہ پاس تبویہ مرغہم انسا نشان از کدی ایثار

و بہ پاس مجربت های درین شان کہ ہرگز ایمان نہی پذیرد

این مجہد و عہد را بہ پدر غنی ذم تقدیرم من کہ نم

قدردانی

ای پللال که بویا کنده این نخلدیشیدن و توانایی آم و خدقن عطا فرمود و با تبه یی نر آدمی به این دو زیرو، اورا اثرشف منخلوقات خود
به لطف و نیکدیش، اراده کرد تا در مریم کرکب عام و دوازش قدم بگزارم و در سایه اللہ صلی اللہ علیہ وسلم و منین علی علیہ السلام و به
دعای نیر محمد صلی اللہ علیہ وسلم این طریق یار عشق بویش و شویاد که مری را در قدم نمایند. فرمود و وارتادی کنه اندیشند و
حلاق جوانی که تمرضا عن علی به را به نایم کار که اندر ایشان کی گنیز کت و چراغ راه پوشو هشتم کرد.

پروردگارا، با تمام عشق و اشتیاقم یکایک استند استیاب به رحمت به من عطا فرمودی و یا به حکمت از من دریغ داشتی، رپاس مکنو یم
ورتا یشت کفم که توبه حق شایسته تارشی.

چکیده

یکی از چالش‌های اساسی در مبحث امنیت سیستم‌های کامپیوتری، تشخیص فعالیت‌های خرابکارانه و نفوذی به سیستم می‌باشد. به همین جهت تشخیص نفوذ بر پایه‌ی متدهای یادگیری، هم‌اکنون به یکی از رویکردهای فعال در تحقیقات امنیتی تبدیل شده است. این تحقیقات دارای دو چالش عمده می‌باشند. نخست جمع‌آوری امن اطلاعات و دوم ایجاد متدی دقیق برای تشخیص ناهنجاری‌ها. در اینجا ما از فراخوان‌های سیستمی و آرگومان‌های آن‌ها، به عنوان الگویی برای توصیف رفتار فرایندهای سیستمی استفاده می‌نماییم.

در کاربردهای امنیتی، الگوهای رفتاری می‌بایست به صورت کامل امن و به دور از دستبرد و تغییر مهاجمان و روت‌کیت‌ها جمع‌آوری گردند. به همین منظور، در این پژوهش، سیستمی نوین با عنوان SHADuDT¹ ارائه می‌دهیم. این سیستم در حقیقت نوعی معماری امن و مبتنی بر ناظر² است که از یک سو با بکارگیری مؤلفه‌ای تحت عنوان رویدادنگار امن (SHADuDT_SA³) به جمع‌آوری امن اطلاعات مرتبط با فراخوان‌های سیستمی در لایه‌ی مانیتور ماشین مجازی می‌پردازد و از سویی دیگر، رفتارهای هنجار و ناهنجار را از طریق متدی تحت عنوان SHADuDT_DM⁴ تفکیک می‌نماید. SHADuDT_DM الگوریتمی نوین در حوزه‌ی سیستم‌های ایمنی مصنوعی می‌باشد که از تئوری خطر به عنوان یکی از جدیدترین تئوری‌های ایمنی‌شناسی الهام می‌پذیرد و به شبیه‌سازی همزمان عملکردهای هر دو سیستم ایمنی طبیعی و انطباقی می‌پردازد.

این پژوهش پس از ارائه‌ی معماری SHADuDT، به ارزیابی کارایی آن بر اساس معیارهای گوناگون و نیز مقایسه‌ی SHADuDT_DM با متدهای تشخیص نفوذ کلاسیک در حوزه‌ی سیستم‌های ایمنی مصنوعی می‌پردازد. آزمایشات صورت‌گرفته نشان می‌دهند که معماری امن ارائه‌شده با بهره‌گیری همزمان از مزایای سیستم‌های ایمنی مصنوعی و تکنولوژی ناظر، قادر است با تحمل میزان کمی سرباره‌ی زمانی و حافظه‌ای، به رشد قابل توجهی در دقت تشخیص دست یابد.

واژه‌های کلیدی: تشخیص ناهنجاری، سیستم ایمنی مصنوعی، تئوری خطر، تکنولوژی ناظر

¹ Secure Hypervisor based Anomaly Detection using Danger Theory (SHADuDT)

² hypervisor

³ SHADuDT_Secure Auditor (SHADuDT_SA)

⁴ SHADuDT_Detection Method (SHADuDT_DM)

فهرست مطالب

ب	تقدیم
ت	قدردانی
ث	چکیده
ج	فهرست مطالب
د	فهرست جداول
ذ	فهرست اشکال
۱	فصل اول : مقدمه و هدف پژوهش
۲	۱-۱ مقدمه
۳	۲-۱ رویکردهای کلیدی
۶	۳-۱ سازمان‌دهی مستند
۸	۴-۱ خلاصه فصل
۹	فصل دوم : مفاهیم مرتبط
۱۰	۱-۲ مقدمه
۱۰	۲-۲ نفوذ
۱۱	۱-۲-۲ روت‌کیت
۱۵	۳-۲ سیستم‌های تشخیص نفوذ
۱۵	۱-۳-۲ انواع سیستم‌های تشخیص نفوذ بر اساس موقعیت قرارگیری
۱۶	۲-۳-۲ انواع سیستم‌های تشخیص نفوذ بر اساس متد تشخیصی
۱۷	۴-۲ سیستم‌های ایمنی انسانی
۱۷	۱-۴-۲ تئوری‌های ایمنی‌شناسی سنتی
۲۰	۲-۴-۲ فرایند تحریک در سیستم‌های ایمنی انسانی
۲۲	۳-۴-۲ تئوری خطر
۲۵	۵-۲ سیستم‌های ایمنی مصنوعی
۲۶	۱-۵-۲ ویژگی‌های ممتاز سیستم‌های ایمنی مصنوعی برای تشخیص نفوذ
۲۷	۶-۲ مجازی سازی
۲۸	۱-۶-۲ ماشین مجازی و انواع آن
۳۱	۲-۶-۲ خصوصیات مجازی سازی
۳۲	۳-۶-۲ سیستم‌های امنیتی مبتنی بر ماشین مجازی
۳۳	۷-۲ خلاصه فصل

۳۵	فصل سوم : پژوهش‌های مرتبط
۳۶	۱-۳ مقدمه
۳۶	۲-۳ رویدادنگاری فراخوان‌های سیستمی
۳۷	۱-۲-۳ رویدادنگاری در سطح کاربر
۳۸	۲-۲-۳ رویدادنگاری در سطح هسته
۳۹	۳-۲-۳ رویدادنگاری در سطح ناظر
۴۱	۳-۳ متدهای تشخیص ناهنجاری
۴۲	۱-۳-۳ تشخیص ناهنجاری از طریق سیستم ایمنی مصنوعی
۴۷	۲-۳-۳ تشخیص ناهنجاری مبتنی بر اطلاعات سیستمی
۵۱	۴-۳ خلاصه فصل
۵۲	فصل چهارم : ارائه چارچوب پیشنهادی (SHADuDT)
۵۳	۱-۴ مقدمه
۵۵	۲-۴ الهام زیستی
۵۹	۳-۴ معماری مفهومی SHADuDT
۶۲	۱-۳-۴ رویدادنگار امن
۶۴	۲-۳-۴ استخراج‌کننده ویژگی
۶۵	۳-۳-۴ سیستم تشخیص نفوذ زیستی
۷۶	۵-۴ خلاصه فصل
۷۹	فصل پنجم : ارزیابی کارایی چارچوب پیشنهادی
۸۰	۱-۵ مقدمه
۸۲	۲-۵ ارزیابی امنیتی SHADuDT با رویکرد غیر رسمی
۸۴	۳-۵ ارزیابی قدرت تشخیص SHADuDT
۸۴	۱-۳-۵ مجموعه داده Safe Syscalls
۸۸	۲-۳-۵ معیارهای ارزیابی
۹۶	۳-۳-۵ نتایج آزمایشات
۱۰۹	۴-۵ مقایسه الگوریتم‌های انتخاب غیر خودی با الگوریتم SHADuDT_DM
۱۰۹	۱-۴-۵ الگوریتم‌های انتخاب غیر خودی
۱۱۱	۲-۴-۵ مقایسه نتایج
۱۱۵	۵-۵ مقایسه الگوریتم تئوری خطر DCA با الگوریتم SHADuDT_DM
۱۱۵	۱-۵-۵ الگوریتم DCA
۱۱۷	۲-۵-۵ مقایسه نتایج
۱۲۰	۶-۵ مقایسه الگوریتم تئوری خطر TLR با الگوریتم SHADuDT_DM

۱۲۰ TLR الگوریتم	۱-۶-۵
۱۲۳ مقایسه نتایج	۲-۶-۵
۱۲۷ SHADuDT ارزیابی سرباره کارایی	۷-۵
۱۲۷ ارزیابی زمان اجرا	۱-۷-۵
۱۳۱ ارزیابی میزان حافظه مصرفی	۲-۷-۵
۱۳۲ خلاصه فصل	۸-۵
فصل ششم : ارزیابی SHADuDT_DM به عنوان یک طبقه‌بند عام		
۱۳۴	۱۳۴
۱۳۵ مقدمه	۱-۶
۱۳۶ طبقه‌بندهای سنتی	۲-۶
۱۳۷ معرفی مجموعه داده‌ها	۳-۶
۱۳۷ Safe Syscalls مجموعه داده	۱-۳-۶
۱۳۸ NSL-KDD مجموعه داده	۲-۳-۶
۱۳۹ Breast Cancer Wisconsin (Original) مجموعه داده	۳-۳-۶
۱۴۰ Ecoli مجموعه داده	۴-۳-۶
۱۴۱ نتایج آزمایشات	۴-۶
۱۴۲ Safe Syscalls مجموعه داده	۱-۴-۶
۱۴۳ NSL-KDD مجموعه داده	۲-۴-۶
۱۴۴ Breast Cancer Wisconsin (original) مجموعه داده	۳-۴-۶
۱۴۵ Ecoli مجموعه داده	۴-۴-۶
۱۴۶ تحلیل نتایج	۵-۴-۶
۱۴۸ خلاصه فصل	۵-۶
فصل هفتم : نتیجه‌گیری		
۱۵۰	۱۵۰
۱۵۱ مقدمه	۱-۷
۱۵۳ بحث و ارزیابی نتایج پژوهش	۲-۷
۱۵۴ پژوهش‌های آینده	۳-۷
۱۵۶ خلاصه فصل	۴-۷
I فهرست مراجع	
IX مقالات مستخرج از تز	
X Abstract	

فهرست جداول

- جدول ۳-۱: مقایسه رویکردهای متفاوت رویدادننگاری فراخوان‌های سیستمی ۳۷
- جدول ۳-۲: مقایسه مدل‌های گوناگون سیستم‌های تشخیص نفوذ مبتنی بر AIS ۴۲
- جدول ۳-۳: دسته‌بندی روش‌های موجود در حوزه‌ی تشخیص فراخوان‌های سیستمی ناهنجار ۵۰
- جدول ۵-۱: نتایج ارزیابی الگوریتم SHADuDT_DM بر مجموعه داده Safe Syscalls ۹۸
- جدول ۵-۲: نتایج ارزیابی الگوریتم NSCD با بر مجموعه داده Safe Syscalls ۱۱۱
- جدول ۵-۳: نتایج ارزیابی الگوریتم NSVD بر مجموعه داده Safe Syscalls ۱۱۱
- جدول ۵-۴: نتایج ارزیابی الگوریتم DCA بر مجموعه داده Safe Syscalls ۱۱۹
- جدول ۵-۵: نتایج ارزیابی الگوریتم TLR بر مجموعه داده Safe Syscalls ۱۲۳
- جدول ۵-۶: مقایسه‌ی سرباره‌ی زمانی رویدادننگاری بر زمان اجرای فراخوان‌های سیستمی ۱۲۸
- جدول ۶-۱: نتایج ارزیابی قدرت طبقه‌بندی متدهای یادگیری ماشین و الگوریتم‌های مبتنی بر AIS بر مجموعه داده Safe Syscalls ۱۴۳
- جدول ۶-۲: نتایج ارزیابی قدرت طبقه‌بندی متدهای یادگیری ماشین و الگوریتم‌های مبتنی بر AIS بر جموعه داده NSL-KDD ۱۴۴
- جدول ۶-۳: نتایج ارزیابی قدرت طبقه‌بندی متدهای یادگیری ماشین و الگوریتم‌های مبتنی بر AIS بر جموعه داده Breast Cancer Wisconsin ۱۴۵
- جدول ۶-۴: نتایج ارزیابی قدرت طبقه‌بندی متدهای یادگیری ماشین و الگوریتم‌های مبتنی بر AIS بر جموعه داده Ecoli ۱۴۶
- جدول ۷-۱: مقایسه الگوریتم SHADuDT_DM با الگوریتم‌های مبتنی بر AIS از لحاظ ساختاری ۱۵۴

فهرست اشکال

- شکل ۱-۲: محیط ماشین مجازی نوع اول و نوع دوم..... ۳۰
- شکل ۱-۳: راهکار سیستم ایمنی مصنوعی با مدل تمایز خودی و غیر خودی. الف) عملیات مربوط به ایجاد مجموعه تشخیص دهنده‌ها و ب) عملیات کشف..... ۴۳
- شکل ۲-۳: مدل انتخاب غیر خودی با تشخیص دهنده‌هایی با طول الف) متغیر و ب) ثابت..... ۴۴
- شکل ۳-۳: چرخه عمر تشخیص دهنده‌ها در سیستم ایمنی مصنوعی..... ۴۴
- شکل ۴-۳: مدل سیستم ایمنی مصنوعی مبتنی بر پردازش تکاملی..... ۴۶
- شکل ۱-۴: شمایی از حضور سلولهای دندریتی و سلولهای T در حالت‌های گوناگون درون نودهای لنفاوی و بافت غیرلنفاوی..... ۵۷
- شکل ۲-۴: شمایی از معماری مفهومی چارچوب ارائه شده (SHADuDT)..... ۶۰
- شکل ۳-۴: ساختار الگوی خروجی از مؤلفه‌ی استخراج کننده‌ی ویژگی..... ۶۵
- شکل ۴-۴: نمایش یک nTC به همراه گیرنده‌های آن..... ۶۸
- شکل ۵-۴: نمایش یک کره‌ی خطر..... ۶۹
- شکل ۶-۴: انواع گوناگون سلول‌های دندریتی به همراه گیرنده‌های مختص هر نوع..... ۷۰
- شکل ۷-۴: انواع گوناگون سلولهای T..... ۷۰
- شکل ۸-۴: نمایش مراحل گوناگون فاز تشخیص SHADuDT_DM..... ۷۶
- شکل ۱-۵: بخشی از مجموعه داده‌ی Safe Syscalls قبل از عملیات پیش پردازش..... ۸۵
- شکل ۲-۵: بخشی از مجموعه داده‌ی Safe Syscalls بعد از عملیات پیش پردازش..... ۸۸
- شکل ۳-۵: نمایش چهار مفهوم پایه در ارزیابی قدرت تفکیک یک الگوریتم..... ۸۹
- شکل ۴-۵: گراف ROC برای نمایش قدرت یک طبقه‌بند فرضی..... ۹۵
- شکل ۵-۵: نمودار ROC برای الگوریتم SHADuDT_DM در اثر افزایش تعداد nTC..... ۱۰۰
- شکل ۶-۵: نمودار میزان پیشرفت قدرت الگوریتم SHADuDT_DM در اثر افزایش تعداد nTC بر اساس مقادیر FPR و TPR..... ۱۰۱
- شکل ۷-۵: نمودار میزان پیشرفت قدرت الگوریتم SHADuDT_DM در اثر افزایش تعداد nTC بر اساس مقادیر FNR و TNR..... ۱۰۱
- شکل ۸-۵: نمودار میزان پیشرفت مقادیر PPV، NPV و ACC در الگوریتم SHADuDT_DM بر اثر افزایش تعداد nTC..... ۱۰۲

شکل ۵-۹: نمودار میزان پیشرفت مقادیر PPV، TPR و F1 در الگوریتم SHADuDT_DM بر اثر افزایش تعداد nTC..... ۱۰۴

شکل ۵-۱۰: نمودار میزان پیشرفت مقادیر Acc، MCC و F1 در الگوریتم SHADuDT_DM بر اثر افزایش تعداد nTC..... ۱۰۴

شکل ۵-۱۱: شمای کلی یک اتوماتای یادگیر در تعامل با محیط..... ۱۰۶

شکل ۵-۱۲: نمودار بهینه‌سازی شعاع کره‌های خودی از طریق اتوماتای یادگیر..... ۱۰۸

شکل ۵-۱۳: نمودار بهینه‌سازی فاصله‌ی حاشیه‌ای کره‌های خطر از طریق اتوماتای یادگیر..... ۱۰۸

شکل ۵-۱۴: نمودار مقایسه‌ی میزان پیشرفت قدرت سه الگوریتم NSVD، NSCD و SHADuDT_DM بر اساس مقادیر FPR و TPR..... ۱۱۲

شکل ۵-۱۵: نمودار شکل ۵-۱۲ برای $FPR \leq 0.1$ ۱۱۳

شکل ۵-۱۶: نمودار مقایسه‌ی میزان پیشرفت معیار F1 برای سه الگوریتم NSVD، NSCD و SHADuDT_DM بر اثر افزایش تعداد تشخیص‌دهنده‌ها..... ۱۱۴

شکل ۵-۱۷: نمودار مقایسه‌ی میزان پیشرفت معیار Acc برای سه الگوریتم NSVD، NSCD و SHADuDT_DM بر اثر افزایش تعداد تشخیص‌دهنده‌ها..... ۱۱۴

شکل ۵-۱۸: نمودار مقایسه‌ی میزان پیشرفت معیار MCC برای سه الگوریتم NSVD، NSCD و SHADuDT_DM بر اثر افزایش تعداد تشخیص‌دهنده‌ها..... ۱۱۵

شکل ۵-۱۹: ضرایب مربوط به سیگنال‌های سه‌گانه در الگوریتم DCA..... ۱۱۷

شکل ۵-۲۰: نمودار مقایسه‌ی الگوریتم SHADuDT_DM و الگوریتم DCA در سه حالت تک‌گام، دو گام و n گام بر اساس معیارهای F1، Acc و MCC..... ۱۲۰

شکل ۵-۲۱: نمودار مقایسه‌ی میزان پیشرفت قدرت دو الگوریتم TLR و SHADuDT_DM بر اساس مقادیر FPR و TPR..... ۱۲۵

شکل ۵-۲۲: نمودار مقایسه‌ی میزان پیشرفت معیار F1 برای دو الگوریتم TLR و SHADuDT_DM بر اثر افزایش تعداد تشخیص‌دهنده‌ها..... ۱۲۶

شکل ۵-۲۳: نمودار مقایسه‌ی میزان پیشرفت معیار Acc برای دو الگوریتم TLR و SHADuDT_DM بر اثر افزایش تعداد تشخیص‌دهنده‌ها..... ۱۲۶

شکل ۵-۲۴: نمودار مقایسه‌ی میزان پیشرفت معیار MCC برای دو الگوریتم TLR و SHADuDT_DM بر اثر افزایش تعداد تشخیص‌دهنده‌ها..... ۱۲۷

شکل ۵-۲۵: نمودار مقایسه‌ی زمان فراخوانی فراخوان‌های سیستمی در حالت عدم وجود رویدادننگاری و رویدادننگاری از طریق مؤلفه‌ی رویدادننگار امن..... ۱۲۹

شکل ۶-۱: نمودار مقایسه‌ی قدرت طبقه‌بندی متدهای یادگیری ماشین و الگوریتم‌های مبتنی بر AIS روی مجموعه داده‌ی Safe Syscalls بر اساس معیارهای F1، Acc و MCC..... ۱۴۳

شکل ۶-۲: نمودار مقایسه‌ی قدرت طبقه‌بندی متدهای یادگیری ماشین و الگوریتم‌های مبتنی بر AIS روی مجموعه داده‌ی NSL-KDD بر اساس معیارهای F1، Acc و MCC. ۱۴۴.....

شکل ۶-۳: نمودار مقایسه‌ی قدرت طبقه‌بندی متدهای یادگیری ماشین و الگوریتم‌های مبتنی بر AIS روی مجموعه داده‌ی Breast Cancer Wisconsin بر اساس معیارهای F1، Acc و MCC. ۱۴۵.....

شکل ۶-۴: نمودار مقایسه‌ی قدرت طبقه‌بندی متدهای یادگیری ماشین و الگوریتم‌های مبتنی بر AIS روی مجموعه داده‌ی Ecoli بر اساس معیارهای F1، Acc و MCC. ۱۴۶.....

فصل اول : مقدمه و هدف پژوهش

۱-۱ مقدمه

یکی از چالش‌های اساسی در مبحث امنیت سیستم‌های کامپیوتری، تشخیص فعالیت‌های خرابکارانه و نفوذی به سیستم می‌باشد. این عملیات نفوذگرانه، صحت^۱، محرمانگی^۲ و در دسترس بودن^۳ منابع سیستم و شبکه را تهدید می‌نمایند و علیرغم تلاش‌های فراوان در جهت مقابله با آنها، هر روز با شکلی نوین و سیاستی تازه‌تر، اهداف خویش را پی می‌گیرند.

فعالیت‌های نفوذگرانه به سیستم‌های کامپیوتری، از طریق نرم‌افزارهای بدخواه^۴ یا بدافزارها^۵ صورت می‌پذیرد. این بدافزارها که با نیت ورود، اعمال تغییر و یا تخریب دیگر نرم‌افزارهای موجود در کامپیوتر، بدون مجوز وارد آن می‌شوند، انواع متنوعی نظیر ویروس‌ها، کرم‌ها، اسب-های تراوا^۶، شماره‌گیرها^۷، درهای پشتی^۸، نرم‌افزارهای جاسوسی^۹ و روت‌کیت‌ها^{۱۰} را شامل می‌شوند. رشد پرسرعت این بدافزارها موجب گردیده است که تشخیص و مقابله با آنها یکی از ضروری‌ترین گام‌ها در جهت ایمن‌سازی سیستم‌های کامپیوتری محسوب گردد.

روت‌کیت‌ها یکی از خطرناک‌ترین انواع بدافزارها به حساب می‌آیند که قادر به حفظ حضوری ثابت، استوار و غیرقابل کشف بر روی سیستم بوده و به منظور به خطر انداختن و کنترل سیستم از راه دور، به همراه بیشتر انواع بدافزار بکارگرفته می‌شوند. زمانی که یک روت-کیت، به عنوان کاربر ریشه اجرا می‌شود، می‌تواند به تمامی منابع ماشین تحت نفوذ دسترسی یابد و آنها را به صورت دلخواه تغییر دهد. همین قدرت مطلق روت‌کیت‌ها سبب شده است که دفاع و مقابله با آنها در سال‌های اخیر، به عنوان یکی از چالش برانگیزترین موضوعات در جوامع امنیتی مطرح شود.

در همین راستا، هدف اصلی این پژوهش، ارائه‌ی یک چارچوب مناسب برای تشخیص نفوذ و مقابله با روت‌کیت‌هاست که تا حد امکان قادر به کاهش نواقص و افزایش کارایی سیستم‌های

¹ integrity

² confidentiality

³ availability

⁴ Malicious software

⁵ Malware

⁶ Trojan Horse

⁷ dialer

⁸ backdoor

⁹ Spyware

¹⁰ Rootkit

پیشین باشد. در ادامه، رویکردهای کلیدی این پژوهش را در بخش ۱-۲ و شیوه‌ی سازماندهی مستند را در بخش ۱-۳ بررسی می‌نماییم. در نهایت بخش ۱-۴ خلاصه‌ای از مطالب ذکر شده در این فصل را ارائه می‌دهد.

۲-۱ رویکردهای کلیدی

رویکرد اصلی این پژوهش، ارائه‌ی یک چارچوب مناسب برای تشخیص و مقابله با انواع نفوذ و به ویژه روت‌کیت‌های سطح هسته می‌باشد. ارائه‌ی یک چارچوب جدید در حوزه‌ی سیستم‌های تشخیص نفوذ، نیازمند بررسی دقیق مسائل و چالش‌های گوناگونی نظیر موقعیت قرارگیری این سیستم‌ها و نیز متد تشخیصی مناسب برای کشف نفوذ می‌باشد.

این سیستم‌ها بر اساس موقعیتی که در آن قرار می‌گیرند، به دو دسته‌ی مبتنی بر شبکه و مبتنی بر میزبان تفکیک می‌شوند. سیستم‌های تشخیص نفوذ مبتنی بر شبکه^۱ (NIDS) [۵۸]، بر روی کامپیوتر و یا ابزار ویژه‌ای در شبکه قرار می‌گیرند و ترافیک شبکه را به منظور کشف نفوذ، مورد بررسی قرار می‌دهند؛ در حالیکه سیستم‌های تشخیص نفوذ مبتنی بر میزبان^۲ (HIDS) [۱۲۹]، رفتارهای میزبان تحت نظارت را ارزیابی می‌نمایند. این سیستم‌ها دارای دید سیستمی بهتر و در نتیجه قدرت بالاتری در کشف نفوذ نسبت به سیستم‌های مبتنی بر شبکه می‌باشند اما به دلیل قرارگیری بر روی میزبان تحت نظارت، نسبت به حملات روت‌کیت‌هایی که ممکن است بر روی میزبان قرار گیرند، آسیب‌پذیر هستند.

ایزوله‌سازی سیستم‌های تشخیص نفوذ مبتنی بر میزبان از میزبان تحت نظارت، یک رویکرد عملی در جهت حفظ این سیستم‌ها در برابر روت‌کیت‌ها به شمار می‌رود و بدین ترتیب می‌توان HIDSها را به عنوان گزینه‌ای مناسب در ایمن‌سازی سیستم‌های کامپیوتری مطرح نمود.

سیستم‌های تشخیص نفوذ مبتنی بر میزبان را می‌توان در لایه‌های متفاوتی از نرم‌افزار سیستم پیاده‌سازی نمود. سیستم‌های سطح کاربر [۱۴۸] به دلیل دارا بودن دیدی محدود نسبت به سیستم، قادر به کشف مناسب پردازش‌های بدرفتار نمی‌باشند و به راحتی می‌توانند از طریق پردازش‌هایی با مجوز بالاتر دور زده شوند. در مقابل سیستم‌های تشخیص نفوذ سطح

¹ Network based Intrusion Detection System (NIDS)

² Host based Intrusion Detection System (HIDS)

هسته [۹۹]، علاوه بر داشتن دیدی جامع نسبت به تمامی زیرسیستم‌های هسته، از گزند فرایندهای سطح کاربر نیز در امان هستند. اما همچنان نسبت به حملات روت‌کیت‌های سطح هسته، آسیب‌پذیر می‌باشند.

به طور کلی، تعبیه‌ی HIDSها در لایه‌های هسته یا کاربر، به معنای افزودن یک تابع جدید، به کدی با حجم بالا و بسیار پیچیده می‌باشد که عمل قابل اعتمادی محسوب نمی‌شود. بنابراین ما در این پژوهش به دلایل گوناگون از قبیل وجود روت‌کیت‌های سطح هسته و نیز نیاز به اجازه‌ی دسترسی سطح بالا به حافظه‌ی هسته و ساختار داخلی آن، سیستم تشخیص نفوذ پیشنهادی خود را درون لایه‌ی ناظر تعبیه می‌نماییم. این ناظر از طریق واسط معین و شناخته‌شده‌ی x86 ISA با لایه‌های دیگر ارتباط برقرار می‌نماید. استفاده از تکنولوژی مجازی‌سازی پردازنده‌های x86، سیستم تشخیص نفوذ را به وسیله‌ی ایزوله‌سازی کامل آن از سایر لایه‌های نرم‌افزار سیستم، در برابر حملات مقاوم‌تر می‌سازد [۱۱۳].

از سویی دیگر برای اعمال متدهای تشخیص ناهنجاری، نخست نیازمند جمع‌آوری داده‌های مناسب برای تفکیک رفتارهای هنجار و ناهنجار می‌باشیم. به منظور جمع‌آوری داده‌ها، سیستم‌های تشخیص نفوذ می‌توانند از متدهای گوناگونی نظیر پروفایل کردن داده‌ها و یا تحلیل مبتنی بر ویژگی به منظور یافتن دیدی مناسب از فرایندهای در حال اجرا، استفاده نمایند [۱۳۸]. رویدادنگاری فراخوان‌های سیستمی^۱ رویکردی است که به منظور ایجاد الگویی مناسب از رفتار فرایندها، از متد تحلیل مبتنی بر ویژگی استفاده می‌کند. تاکنون رویدادنگاری فراخوان‌های سیستمی در لایه‌های گوناگونی از نرم‌افزار سیستم نظیر لایه‌های کاربر، هسته و ناظر، صورت پذیرفته است [۸۶،۴۵،۸۳،۴۷،۱۳۱،۱۱۶،۶۷]. ما در این پژوهش از رویدادنگاری فراخوان‌های سیستمی در لایه‌ی ناظر بهره می‌گیریم تا عملیات جمع‌آوری داده را به صورت امن و به دور از گزند روت‌کیت‌های هسته به انجام رسانیم.

پس از جمع‌آوری امن داده‌ها نخست می‌بایست ویژگی‌های مناسب برای تفکیک رفتار هنجار و ناهنجار را از میان داده‌های ذخیره‌شده استخراج نموده و به فرم یک الگو^۲ در بیاوریم و سپس به ساخت یک مدل رفتاری مبتنی بر این الگوها پردازیم. در این حوزه، سیستم‌های

^۱ System call auditing

^۲ Pattern

تشخیص نفوذ با دو رویکرد کلی، عملیات ساخت مدل رفتاری را دنبال می‌نمایند. در رویکرد اول که اصطلاحاً تشخیص سوء استفاده^۱ نامیده می‌شود، رفتارهای ناهنجار تحت عنوان امضای^۲ بدافزار مدل می‌شوند و یک نفوذ، زمانی کشف می‌شود که با یکی از امضاهای بدافزار، تطبیق داشته باشد. اما رویکرد دوم که اصطلاحاً تشخیص ناهنجاری^۳ نامیده می‌شود، به مدل‌سازی رفتارهای هنجار سیستم می‌پردازد و یک نفوذ، زمانی کشف می‌شود که با مدل هنجار تطبیق نداشته باشد.

رویکرد تشخیص سوء استفاده در شناسایی نفوذهای شناخته‌شده بسیار توانمند است و به همین سبب، هم‌اکنون در بسیاری از سیستم‌های تجاری به کار گرفته شده است؛ اما بزرگترین نقص آن در عدم توانایی کشف و شناسایی رفتارهای نفوذی جدیدی می‌باشد که امضای آن برای سیستم شناخته‌شده نیست. رویکرد تشخیص ناهنجاری با هدف برطرف‌سازی این نقص عمده‌ی رویکرد اول ارائه شده و به همین سبب در دهه‌های اخیر با استقبال چشمگیری در مجامع علمی روبروست.

این پژوهش در صدد است تا از طریق سیستم‌های ایمنی مصنوعی^۴ به عنوان یکی از شاخه‌های هوش محاسباتی، به تشخیص ناهنجاری در سیستم‌های کامپیوتری بپردازد. سیستم‌های ایمنی مصنوعی از سیستم ایمنی انسانی^۵ الهام گرفته‌اند و به دلیل خصوصیات نظیر خودسازماندهی، توزیع‌شدگی و سبک بودن، گزینه‌ی مناسبی برای ایجاد سیستم‌های تشخیص ناهنجاری به حساب می‌آیند.

به منظور جمع‌آوری کلیه‌ی خصوصیات ذکرشده در یک چارچوب تشخیص نفوذ، در این پژوهش به ارائه‌ی سیستمی نوین با عنوان SHADuDT می‌پردازیم. این سیستم در حقیقت نوعی معماری امن و مبتنی بر ناظر است که از یک سو با بکارگیری مؤلفه‌ای تحت عنوان رویدادنگار امن (SHADuDT_SA) به جمع‌آوری امن اطلاعات مرتبط با فراخوان‌های سیستمی و آرگومان‌های آن در لایه‌ی ناظر یا مانیتور ماشین مجازی^۶ می‌پردازد و به این ترتیب از گزند

¹ Misuse Detection

² Signature

³ Anomaly Detection

⁴ Artificial Immune System (AIS)

⁵ Human Immune System (HIS)

⁶ Virtual Machine Monitor (VMM)

روت‌کیت‌های هسته در امان می‌ماند و از سویی دیگر از نسل دوم سیستم‌های ایمنی مصنوعی به عنوان متد تشخیص ناهنجاری (SHADuDT_DM) بهره می‌گیرد.

تاکنون بسیاری از سیستم‌های تشخیص نفوذ، عملکرد سیستم ایمنی انطباقی را شبیه‌سازی نموده‌اند اما ایده‌ای جدید در این حوزه که با عنوان تئوری خطر [۹۳،۹۴،۱۴۳] شناخته می‌شود، به منظور ارائه‌ی مدلی دقیق‌تر از فرایندهای زیستی سیستم‌های ایمنی انسانی، علاوه بر سیستم ایمنی انطباقی، عملکرد سیستم ایمنی طبیعی و روابط این دو با یکدیگر را نیز شبیه‌سازی می‌کند.

SHADuDT_DM در میان الگوریتم‌های موجود در حوزه‌ی نسل دوم سیستم‌های ایمنی مصنوعی جای می‌گیرد اما بر خلاف الگوریتم‌های مبتنی بر تئوری خطر پیشین نظیر DCA [۵۶-۵۰] و TLR [۱۲۳-۱۲۵] که به سیگنال‌های خطر خارجی نیازمند هستند، این روش سیگنال‌های مورد نیاز خود را از طریق ساختار داده‌های هنجار و ناهنجار آموزش تولید می‌نماید و برای تولید سیگنال، به منابع خارجی نیازمند نیست. به دلیل استفاده از ساختار داده‌های آموزش، این روش دارای خطای کمتر و کارایی بهتری نسبت به روش‌های تئوری خطر پیشین است و به همین سبب می‌تواند به عنوان یک طبقه‌بند مناسب برای تفکیک رفتارهای هنجار و ناهنجار مورد استفاده قرار گیرد.

به این ترتیب سیستم SHADuDT می‌تواند با بهره‌گیری همزمان از مزایای تکنولوژی ناظر و تئوری خطر، هم از گزند روت‌کیت‌های هسته مصون بماند و هم با متدی دقیق و الهام گرفته از سیستم ایمنی انسانی، به تفکیک رفتارهای هنجار و ناهنجار سیستم پردازد.

۳-۱ سازمان‌دهی مستند

ادامه‌ی این مستند به صورت زیر سازمان‌دهی می‌شود. فصل دوم به تشریح مفاهیم مرتبط با این پژوهش می‌پردازد. بدافزارها و انواع آن، روت‌کیت‌ها و سیر تکاملی آنها در انجام فعالیت‌های نفوذی، سیستم‌های تشخیص نفوذ و تکنیک‌های رایجی که تاکنون در این حوزه بکار گرفته شده‌اند، سیستم‌های ایمنی انسانی و سیر تکاملی تئوری‌های موجود در این حوزه، سیستم‌های ایمنی مصنوعی و شیوه‌ی الهام‌پذیری آنها از سیستم‌های ایمنی انسانی و در نهایت مجازی-

سازی و خصوصیات ممتاز آن در جهت افزایش امنیت سیستم و پیاده‌سازی سیستم‌های امنیتی، از جمله مطالبی است که در این فصل به آن پرداخته خواهد شد.

فصل سوم به بررسی پژوهش‌های مرتبط با جنبه‌های گوناگون چارچوب پیشنهادی می‌پردازد. از آنجا که رویدادننگاری فراخوان‌های سیستمی از طریق مؤلفه‌ی رویدادننگار امن و تشخیص فعالیت‌های ناهنجار به کمک مکانیزم تشخیص ناهنجاری، دو بخش حساس SHADuDT را تشکیل می‌دهند، بنابراین رئوس مطالب این فصل به بررسی رویکردهای گوناگون موجود در حوزه‌ی رویدادننگاری فراخوان‌های سیستمی و مکانیزم‌های متنوعی که تاکنون به منظور تشخیص ناهنجاری مبتنی بر میزبان بکار گرفته شده‌اند، تخصیص می‌یابد.

فصل چهارم به طور کامل به شرح و توصیف چارچوب پیشنهادی و یکایک مؤلفه‌های اساسی آن می‌پردازد و بیان می‌کند که SHADuDT نوعی معماری امن و مبتنی بر ناظر است که از یک سو با بکارگیری مؤلفه‌های تحت عنوان رویدادننگار امن به جمع‌آوری امن اطلاعات مرتبط با فراخوان‌های سیستمی در لایه‌ی مانیتور ماشین مجازی می‌پردازد و از سویی دیگر از الگوریتمی نوین در حوزه‌ی تئوری خطر، به عنوان متد تشخیص ناهنجاری بهره می‌گیرد.

فصل پنجم به ارزیابی کارایی چارچوب پیشنهادی از جنبه‌های گوناگون می‌پردازد. ارزیابی امنیتی غیررسمی SHADuDT، بررسی قدرت تشخیص این چارچوب در مقابله با انواع نفوذها، مقایسه‌ی قدرت تشخیص روش پیشنهادی با روش‌های کلاسیک موجود در حوزه‌ی سیستم‌های ایمنی مصنوعی نظیر الگوریتم‌های انتخاب غیرخودی و نیز الگوریتم‌های تئوری خطر پیشین نظیر DCA و TLR و در نهایت بررسی سرباره‌های SHADuDT از منظر زمان اجرا و میزان حافظه‌ی مصرفی، رئوس مطالب این فصل را به خود تخصیص می‌دهند.

فصل ششم، به بررسی جزئی‌تر الگوریتم تشخیص نفوذ SHADuDT_DM می‌پردازد. آزمایشات و تحلیل‌های ارائه‌شده در این فصل، الگوریتم SHADuDT_DM را نه به عنوان الگوریتمی ویژه‌ی تشخیص نفوذ، بلکه به عنوان یک طبقه‌بند عام، با قدرت طبقه‌بندهای رایج موجود در حوزه‌ی یادگیری ماشین نظیر بیزین، SVM و k NN مقایسه می‌نمایند.

در نهایت فصل هفتم، کلیه‌ی نتایج حاصل از این پژوهش و راهکارهایی که می‌توانند به عنوان پژوهش‌های آینده مد نظر قرار گیرند را با جزئیات مطرح می‌نماید.