رایانش ابری

مؤلفین:
دکتر مهدی خلیلی
مهندس مهسا نظری
اعضای هیئت علمی دانشگاه پیام نور

بنام کسی که آفرید از قلم
انتشارات امید مجد
به انسان عطا کرد فکر و قلم

تابستان ۱۳۹۴

سرشناسه: خلیلی، مهدی، ۱۳۵۴ -
عنوان و نام پدیدآور: رایانش ابری/ مولفین مهدی خلیلی،مهسا نظری.
مشخصات نشر: تهران: امید مجد،۱۳۹۴.
مشخصات ظاهری: ۲۶۵ ص.
شابک: ۲-۵۵۰-۷۹۱-۹۶۴-۹۷۸:۶۰۰۰۰ ریال
وضعیت فهرست نویسی: فیپا
موضوع : محاسبات ابری
شناسه افزوده: نظری، مهسا، ۱۳۵۵ -
رده بندی کنگره: ۱۳۹۴/QA۷۶ ۸خ۳م/ ۵۸۵
رده بندی دیویی: ۰۰۶/۷۸
شماره کتابشناسی ملی: ۳۹۶۲۲۳۵

عنوان: رایانش ابری
مؤلفین: دکتر مهدی خلیلی، مهندس مهسا نظری
(اعضای هیئت علمی دانشگاه پیام نور)
ناشر: نشر امید مجد
طراح جلد: علیرضا لکزیان
چاپخانه: موسسه نشاط
صحافی: موسسه نشاط
تیراژ: پانصد نسخه
نوبت چاپ: اول، تابستان ۱۳۹۴
قیمت: شش هزار تومان
شابک: ۲-۵۰-۷۹۱۵-۹۶۴-۹۷۸

آدرس : تهران صندوق پستی ۱۴ ۶۶۵۱۵۵ ۸
تلفن : ۲۲ ۰۶۵۵ ۶۹
Www.OmidMajd.Com

بنام کسی کافریداز عدم

انتشارات امید مجد

به انسان عطا کرد فکر و قلم

فهرست مطالب

چکیده

با گسترش روز افزون تبادل اطلاعات و استفاده از سیستم‌های بر خط، میزان حملات و نفوذ در سیستم‌های اطلاعاتی افزایش یافته است. رایانش ابری گام بعدی تکامل سرویس های فناوری اطلاعات برحسب تقاضا می‌باشد. به منظور جلوگیری از سوءاستفاده از اطلاعات و رخنه در سیستم‌های حساس و مهم باید سیستم‌های مذکور امن شوند.

با گسترش سیستمهای پایگاهی و حجم بالای داده‌های ذخیره شده در این سیستم ها، به ابزاری نیاز است تا بتوان این داده ها را پردازش کرده و اطلاعات حاصل از آن را در اختیار کاربران قرار داد. یکی از روش های بسیار مهمی که با آن می توان الگوهای مفیدی را در میان داده ها تشخیص داد، داده‌کاوی است، این روش که با حداقل دخالت کاربران همراه است اطلاعاتی را در اختیار آنها و تحلیل گران قرار می دهد تا براساس آنها تصمیمات مهم و حیاتی را در سازمانشان اتخاذ نمایند.

در امنیت اطلاعات، کار عملیات تشخیص، تشخیص نفوذ است که برای امکان شناسایی یکپارچگی، قابلیت اطمینان یا قابلیت دسترسی به منابع تلاش می‌کند. بنابراین، این فرآیند بازبینی و تحلیل وقایع به وجودآمده در یک سیستم رایانه به منظور تشخیص مشکلات امنیتی می‌باشد.

در این کتاب ، پس از تشریح تکنیک‌های مختلف خوشه‌بندی داده کاو محور و تشخیص نفوذ، بطور اجمالی الگوریتم‌های داده‌کاوی برای سیستم تشخیص نفوذ را بیان نموده و در پایان، الگوریتم‌های داده‌کاوی خوشه‌بندی K-Means، خوشه‌بندی Y-Means و خوشه‌بندی Fuzzy C-Means را به منظور ارتقاء سیستم‌های تشخیص نفوذ و افزایش سطح امنیت اطلاعات در ابرها پیشنهاد می‌نماییم.

❖ **کلمات کلیدی:** رایانش ابری- حملات امنیتی به ابر- روش‌های تشخیص نفوذ به ابر- سیستم‌های تشخیص نفوذ- داده‌کاوی- سیستم‌های تشخیص نفوذ داده‌کاو محور- الگوریتم K-Means- الگوریتم Y-Means- الگوریتم Fuzzy C-Means.

رایانش ابری، به گونه‌ای از سیستمهای توزیع شده و موازی اطلاق می‌گردد که مجموعه‌ای از کامپیوترهای مجازی را که به یکدیگر متصل هستند شامل می شود. این کامپیوترها به طور پویا عرضه شده و به عنوان یک یا چند منبع محاسباتی یکپارچه بر اساس توافقات سطح سرویس دیده می شوند و این توافقات در طول مذاکرات سرویس دهندگان و مصرف کنندگان برقرار می گردند. رایانش ابری سعی دارد ایجاد پویای نسل جدیدی از مراکز داده ای را، با ارائه کردن سرویس‌ها و خدمات در ماشین های مجازی شبکه شده به صورت پویا، به گونه ای ممکن سازد که کاربران بتوانند از هر جایی از دنیا به برنامه‌های کاربردی دسترسی داشته باشند.
واضح است که رایانش ابری گام بعدی تکامل سرویس های فناوری اطلاعات برحسب تقاضا می‌باشد. با این حال مسئله امنیت در رایانش ابری یکی از مسائل پیچیده به شمار رفته که تمامی سه لایه ابر، مسئولیت هایی که بین کابران و ارائه دهندگان تقسیم می‌شود و حتی شخص ثالث را درگیر می‌کند. مشکلات امنیتی به عنوان یک مانع بزرگ در مقابل استفاده کاربران از سیستم های رایانش ابری قلمداد می‌شود.
سیر تکاملی پردازش ابری به گونه‌ای است که می‌توان آن را پس از آب، برق، گاز تلفن به عنوان عنصر اساسی پنجم فرض نمود. در چنین حالتی، کاربران سعی می کنند بر اساس نیازهایشان و بدون توجه به اینکه یک سرویس در کجا قرار دارد و یا چگونه تحویل داده میشود، به آن دسترسی یابند. نمونه های متنوعی از سیستم های محاسباتی ارائه شده است که سعی دارند چنین خدماتی را به کاربران ارئه دهند. برخی از آنها عبارتند از: محاسبات کلاستری، محاسبات توری و اخیراً رایانش ابری.
در پایین ترین سطح که لایه زیرساخت به عنوان سرویس (Iaas) نامیده می‌شود، پردازنده، حافظه و اجزای سخت افزاری توسط فراهم کننده سرویس ابری ارائه می گردد. لایه میانی یا پلتفرم به عنوان سرویس (Paas) میزبان محیط های مختلف برای ارائه خدمات می‌باشد. در آخر بالاترین لایه نرم افزار کاربردی به عنوان سرویس (SaaS) است که دسترسی به این خدمات از طریق وب سرویس ها و مرورگرهای وب صورت می گیرد. Google App, Amazon Ec2 و Salesforce به ترتیب نمونه های لایه های ارائه شده رایانش ابری می‌باشند.

رایانش (پردازش/محاسبات) ابری ساختاری شبیه یک توده ابر دارد که به واسطه آن کاربران می توانند به برنامه های کاربردی از هر جایی از دنیا دسترسی داشته باشند. بنابراین، رایانش ابری می تواند با کمک ماشین های مجازی شبکه شده، به عنوان یک روش جدید برای ایجاد پویای نسل جدید مراکز داده مورد توجه قرار گیرد. بدین ترتیب، دنیای محاسبات به سرعت به سمت توسعه نرم‌افزارهایی پیش می رود که به جای اجرا بر روی کامپیوترهای منفرد، به عنوان یک سرویس در دسترس میلیون ها مصرف کننده قرار می گیرند. بررسی‌ها نشان می دهد که، محاسبات کلاستری در حال حاضر نسبت به دو مورد دیگر مقبولیت کمتری دارد، محاسبات توری در رتبه دوم قرار گرفته است، و پردازش انبوه با فاصله زیادی از آنها در حال افزایش جلب توجه افراد بیشتری است.

با گسترش روز افزون تبادل اطلاعات و استفاده از سیستم‌های بر خط، میزان حملات و نفوذ در سیستم‌های اطلاعاتی افزایش یافته است. به منظور جلوگیری از سوءاستفاده از اطلاعات و رخنه در سیستم‌های حساس و مهم باید سیستم‌های مذکور امن شوند. مفهوم امنیت به معنای محافظت از محرمانگی، یکپارچگی و دسترس پذیری است. اما محرمانگی، یکپارچگی و دسترس پذیری چیست؟

محرمانه بودن داده ها به معنای این است که داده‌های در حال انتقال در شبکه تنها باید توسط افرادی که به شیوه مناسب احراز هویت شده اند مورد دستیابی قرار بگیرد. حفظ جامعیت داده به این معنی است که داده ها از لحظه ای که منتقل می‌شوند تا لحظه ای که دریافت می‌شوند نباید هیچ اختلال و فقدانی چه به دلیل رویدادهای تصادفی و چه به دلیل فعالیت‌های خرابکارانه داشته باشند. دسترس پذیری در حالت مشهود به معنای آن است که سیستم در مقابل حملات انکار سرویس باید قدرتمند و ضدضربه باشد.

در دستگاه ها و سایر سیستم‌های اطلاعاتی جهت حفظ امنیت از ۵ مرحله حسابرسی استفاده می‌کنند که به ترتیب زیر است:

۱. **هویت:** به معنای داشتن یک شناساگر برای افراد به جهت دسترسی کاربران معتبر به اطلاعات است که نمونه بارز آن در حال حاضر نام کاربری است.

۲. **احراز هویت**: به معنای بررسی صحت نام کاربری فردی است که مدعی مالکیت آن است، به کمک پسورد یا رمز عبور می‌توان صحت هویت فرد مدعی را بررسی کرد.

۳. **اجازه**: فرایندی است که طی آن کاربران و یا تجهیزات متقاضی دسترسی به منابع، امکان استفاده از منبع یا منابع مستقر بر روی شبکه داده می‌شود. به بیان دیگر این عمل برای مدیران شبکه امکان تعیین نوع دسترسی به هر یک از منابع شبکه، برای تک تک متقاضیان دسترسی و یا گروهی از آنها را فراهم می‌کند. لیست‌های کنترل دسترسی مانند دیوارهای آتش و سیستم‌های تشخیص نفوذ در این مرحله قرار دارند. در واقع لیست‌های کنترل دسترسی می‌توانند ترافیک داخلی و خارجی شبکه را کنترل کنند از این رو سیستم‌های تشخیص نفوذ در این لیست قرار می‌گیرند.

۴. **حسابرسی یا ممیزی کردن**: به معنای جمع آوری اطلاعات شبکه و رکوردهای ترافیک شبکه است تا به کمک ابزارهای مناسب بتوان این داده ها را بررسی کرد و در جهت تشخیص عملیات نفوذ و حتی پی بردن به عملکرد سیستم از آنها استفاده کرد. این داده ها می‌توانند درباره سیستم عامل، کاربردها و یا فعالیت کاربران باشد.

۵. **پاسخگویی**: در حالت کلی به معنای بیان نتایج حاصل از درخواست کاربر احراز هویت شده است. با توجه به اهمیت سیستم‌های تشخیص نفوذ در ارتقا امنیت سیستم‌های اطلاعاتی و موقعیت آن در مراحل ایمن سازی می‌توان به کمک این سیستم ها، عملیات نفوذ در یک سیستم را نافرجام باقی گذاشت. اهمیت سیستم‌های تشخیص نفوذ در آنجا به اوج خود می‌رسد که قبل از اینکه حمله به اتمام برسد وقوع آن را گزارش کند و مانع از آسیب‌های احتمالی شود. بنابراین تشخیص نفوذ عبارت است از فرایند شناسایی و پاسخ به فعالیت های مخرب که به صورت هدفمند، منابع شبکه را مورد تهاجم قرار می‌دهد.

براساس تعریف مؤسسه استاندارد و تکنولوژی «فرآیند نظارت بر رویدادهایی که در سیستم کامپیوتر و یا شبکه رخ می‌دهد و تجزیه و تحلیل آنها در جهت یافتن

نشانه‌هایی از نفوذ، که به معنای دورزدن محرمانگی و جامعیت و دسترسی پذیری و یا عبور از مکانیزم‌های امنیتی یک کامپیوتر یا شبکه است، می‌باشد.
با توجه به تکامل در عرصه رایانش، روش های بسیاری جهت توزیع منابع و پیشرفت استفاده از
داده ها از قبیل خوشه‌بندی داده ها، رایانش توری و سیستم مدیریت پایگاه داده های توزیع شده معرفی شده اند. امروزه رایانش ابری مکانیزم در حال ظهور برای محاسبات سطح بالا به عنوان یک سیستم
ذخیره سازی تلقی می‌شود که در آن ابرها به کاربران خود بر مبنای میزان استفاده از منابع هزینه دریافت کرده و سرویس های خود را در اختیار آن‌ها قرار می دهند. از این رو می توان سرویس های ابری را در ایجاد انگیزه برای شروع یک کسب و کار با هزینه های مالی پایین تر سهیم دانست. رایانش ابری بسته به نوع توزیع منابع از سه لایه زیرساخت به عنوان سرویس، پلتفرم به عنوان سرویس و نرم افزار کاربردی به عنوان سرویس تشکیل شده است.
واضح است که رایانش ابری گام بعدی تکامل سرویسهای فناوری اطلاعات برحسب تقاضا می‌باشد. با این حال مسئله امنیت در رایانش ابری یکی از مسائل پیچیده به شمار رفته که تمامی سه لایه ابر، مسئولیت هایی که بین کابران و ارائه دهندگان تقسیم می‌شود و حتی شخص ثالث را درگیر می کند. مشکلات امنیتی به عنوان یک مانع بزرگ در مقابل استفاده کاربران از سیستم های رایانش ابری قلمداد می‌شود.
با گسترش سیستم های پایگاهی و حجم بالای داده‌های ذخیره شده در این سیستم ها، به ابزاری نیاز است تا بتوان این داده ها راپردازش کرد و اطلاعات حاصل از آن را در اختیار کاربران قرار داد. معمولا" کاربران پس از طرح فرضیه ای بر اساس گزارشات مشاهده شده به اثبات یا رد آن می‌پردازند، در حالی که امروزه به روشهایی نیاز داریم که به اصطلاح به کشف دانش بپردازند یعنی روش هایی که با کمترین دخالت کاربر و به صورت خودکار الگوها و رابطه های منطقی را بیان نمایند.
یکی از روش های بسیار مهمی که با آن می توان الگوهای مفیدی را در میان داده ها تشخیص داد، داده‌کاوی است، این روش که با حداقل دخالت کاربران همراه است اطلاعاتی را در اختیار آنها و تحلیل گران قرار می دهد تا براساس آنها تصمیمات

مهم و حیاتی را در سازمانشان اتخاذ نمایند. باید توجه داشت که اصطلاح داده کاوی زمانی به کار برده می شود که با حجم بزرگی از داده ها، در حد مگابایت یا ترابایت، مواجه باشیم. در تمامی منابع داده کاوی بر این مطلب تاکید شده است. هرچه حجم داده‌ها بیشتر و روابط میان آنها پیچیده تر باشد دسترسی به اطلاعات نهفته در میان داده ها مشکلتر می شود و نقش داده کاوی به عنوان یکی از روش های کشف دانش، آشکارتر می‌گردد.

داده‌کاوی از چندین رشته علمی (نظیر: تکنولوژی پایگاه داده، هوش مصنوعی، شبکه‌های عصبی، آمار، سیستم های مبتنی بر دانش، بازیابی اطلاعات و غیره) به طور همزمان بهره می‌برد. که برای پرهیز از اطاله کلام می‌توان به لحاظ تاریخی به اختصار به مراحل زیر تقسیم کرد:

- **مرحله اول**: گردآوری وایجاد پایگاه اطلاعاتی (تا دهه ۱۹۶۰)
- **مرحله دوم**: نظام های مدیریتی مبنی بر پایگاه اطلاعاتی (دهه ۱۹۷۰ و اوایل دهه ۱۹۸۰)
- **مرحله سوم**: نظام های پایگاه اطلاعاتی پیشرفته (اواسط دهه ۱۹۸۰ تا زمان حاضر)
- **مرحله چهارم** : انبارش اطلاعات و داده‌کاوی (اواخر دهه ۱۹۸۰ تا به امروز)
- **مرحله پنجم**: نظام پایگاه اطلاعاتی مبنی برشبکه (دهه ۱۹۹۰ تا کنون)
- **مرحله ششم**: نسل نوین نظام های اطلاعاتی یکپارچه شده (از ۲۰۰۰ به بعد)

بدین ترتیب فعالیتی که از دهه ۱۹۶۰ شروع شده بود در دهه ۱۹۹۰ گام های بلندی برداشت و انتظار می‌رود در این قرن به رشد و بالندگی خود ادامه دهد.

در امنیت اطلاعات، کار عملیات تشخیص، تشخیص نفوذ است که برای امکان شناسایی یکپارچگی، قابلیت اطمینان، یا قابلیت دسترسی به منابع تلاش می‌کند. درکل، تشخیص نفوذ منجر به جلوگیری از نفوذها نمی‌شود. تقاضاهای تشخیص نفوذ پیشرفته با مشکلات پیچیده‌ای روبرو هستند. این تقاضاها برای مدیریت به توسعه‌پذیری، معتبربودن و آسان بودن نیاز داشته و هزینه نگهداری پایینی دارند.

داده کاوی توجه زیادی به علت افزایش، تولید، انتقال و ذخیره مقدار داده دارد و برای خلاصه‌کردن به جذب اطلاعات و دانش سودمند از آن ها نیاز دارد. در سال‌های گذشته، تحقیق درخصوص امکان استفاده از تکنیک‌های داده‌کاوی در زمینه برآوردن امنیت اطلاعات مخصوصاً در مشکل چالش‌انگیز تشخیص نفوذ شروع شده است. تشخیص نفوذ فرآیند یافتن وقایع مهم به وجودآمده در یک سیستم رایانه و تحلیل آن ها برای وجود نفوذ است. بنابراین، این فرآیند بازبینی و تحلیل وقایع به وجودآمده در یک سیستم رایانه به منظور تشخیص مشکلات امنیتی می‌باشد.

یک IDS براساس کیفیت و کارایی به یک آرایه با مؤلفه‌ها و خصوصیات مختلف نیاز دارد شامل:

- متمرکز کردن دیدگاه داده
- توانایی انتقال داده
- روش‌های داده‌کاوی و تحلیلی
- گسترش کشف‌کننده انعطاف‌پذیر شامل زمان‌بندی که فعال می‌شود.
- تشخیص زمان واقعی و سازنده علامت هشدار
- قابلیت های گزارش‌گیری
- فرآیند توزیع
- قابلیت دسترسی بالا به سیستم
- مقیاس‌پذیری با بارگذاری سیستم

در کل دو نوع حمله وجود دارد:

- حمله داخلی، حمله‌ای است که یک مزاحم، همه امتیازی برای دسترسی به درخواست دارد.
- حمله خارجی، حمله‌ای است که مزاحم، حقوق مناسبی برای دسترسی به سیستم ندارد. تشخیص حمله داخلی معمولاً در مقایسه با حمله خارجی مشکل‌تر است.

در فصل نخست این کتاب، به تشریح ابعاد مختلف پردازش ابری شامل تعریف ابرها و انواع آنها، نقاط قوت و ضعف محاسبات ابری و همچنین معماری لایه‌های ابری و انواع آن، نحوه مدیریت ابرها می‌پردازیم.

در فصل دوم، به تعریف داده‌کاوی، الگوریتم‌های مورداستفاده در آن و ضرورت استفاده از داده کاوی در محاسبات ابری خواهیم پرداخت.
در فصل سوم نیز، نکات امنیتی ابرها شامل حملات امنیتی به ابرها، انواع روش‌های تشخیص نفوذ به ابرها و راهکارهای مقابله، انواع معماری سیستم‌های تشخیص نفوذ، اجزای سیستم‌های تشخیص نفوذ، ارزیابی سیستم‌های تشخیص نفوذ را بیان می‌نماییم.
در فصل چهارم، سیستم‌های تشخیص نفوذ داده کاو محور، تکنیک‌های داده کاوی و درختان تصمیم‌گیری و مدل‌های تشخیص نفوذ را مورد بحث و بررسی قرار می‌دهیم.
و در انتها، به بیان تکنیک‌های خوشه‌بندی داده کاو محور و تشخیص نفوذ پرداخته، به طور اجمالی الگوریتم‌های داده‌کاوی برای سیستم تشخیص نفوذ را بیان نموده و در پایان، الگوریتم‌های داده‌کاوی خوشه‌بندی K-Means، خوشه‌بندی Y-Means و خوشه‌بندی Fuzzy C-Means را به منظور ارتقاء سیستم‌های تشخیص نفوذ و افزایش سطح امنیت اطلاعات در ابرها پیشنهاد کرده و نقاط قوت ضعف هریک را تشریح می‌نماییم.