APJ. 4(2):11-20, Summer 2023

Arman Process Journal (APJ)

Homepage: armanprocessjournal.ir

Examining Data Warehouse and Big Data Security Aspects

Z. Nasiri*,1

[1] Computer Engineering Department, Sharif University of Technology,Tehran,Iran

Received: 12 June 2023
Accepted: 25 August 2023

KEYWORDS:
Security Management
Reliability
Data Warehouse
Big Data
Data Encryption

[1] Corresponding author
za.nasiri@sharif.edu

ABSTRACT

Currently, the significant growth of the number of users of online services and distributed systems generate a large amount of database information at the level of data warehouses and big data. Data warehouses are large collections of business data that help organizations and businesses make more accurate and intelligent decisions. Big data is also a very large collection of data collected from multiple sources. These data can be the results of evaluating the performance of an organization or the interactions of its audience in social networks. Due to the increasing importance of big data functions and data warehouses, the issue of maintaining data security is the biggest threat these systems face. Basically, dealing with the category of information security and security of internet networks, data warehouses and services based on big data, requires special attention of the organization to the position of information security and all-round security of these systems, and this category should be considered at the macro level and from the perspective of benefits and benefits. it looked The existence of security weaknesses in these systems, the lack of proper training and justification of users regardless of their point of view regarding the position and importance of security, the absence of necessary instructions to prevent security defects, the absence of specific and codified policies in order to properly and timely deal with Security flaws will lead to issues that harm all users and the operational capabilities of these systems and actually expose the information infrastructure of organizations to serious damage and threats. In this article, we are going to research the security aspects of data warehouses and big data and examine effective and logical solutions in this area.

NUMBER OF REFERENCES	NUMBER OF FIGURES	NUMBER OF TABLES
24	0	0

نشریه تخصصی آرمان پردازش، دوره ۴، شماره ۲، تابستان ۱۴۰۲

فصلنامه تخصصی آرمان پردازش
(APJ)
Homepage: www.armanprocessjournal.ir

بررسی جنبه های ایمن سازی انبار داده و کلان داده

زهرا نصیری[*،۱]
[۱] دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شریف، تهران، ایران

چکیده

در حال حاضر، رشد چشمگیر تعداد کاربران خدمات برخط و سیستم های توزیعی، مقدار زیادی از اطلاعات پایگاه داده ای در سطح انبارهای داده و کلان داده ها را تولید می نمایند. انبارهای داده مجموعه بزرگی از داده‌های تجاری هستند که به سازمان‌ها و کسب‌وکارها کمک می‌کند تا در تصمیم‌گیری‌های خود دقیق‌تر و هوشمندانه‌تر عمل کنند. کلان داده نیز مجموعه‌ای بسیار بزرگ از داده‌ها است که از منابع متعددی جمع‌آوری می‌شود. این داده‌ها می‌توانند نتایج ارزیابی عملکرد یک سازمان یا تعاملات مخاطبان آن در شبکه‌های اجتماعی باشند. بنا به اهمیت روزافزون کارکردهای کلان داده و انبارهای داده، مسئله ای حفظ امنیت داده ها، بزرگترین تهدیدی است که این سیستم ها با آن مواجه می شوند. اساسا پرداختن به مقوله امنیت اطلاعات و ایمن سازی شبکه‌های اینترنتی، انبارهای داده و خدمات مبتنی بر کلان داده، مستلزم توجه ویژه سازمان به جایگاه امنیت اطلاعات و ایمن‌سازی همه جانبه این سیستم ها بوده و می‌بایست به این مقوله در سطح کلان و از بعد منافع و فواید آن نگریست. وجود ضعف امنیتی در این سیستم ها، عدم آموزش و توجیه صحیح کاربران صرف نظر از دیدگاه آنان نسبت به جایگاه و اهمیت ایمن سازی، عدم وجود دستورالعمل‌های لازم برای پیشگیری از نقایص امنیتی، عدم وجود سیاست‌های مشخص و مدون به منظور برخورد مناسب و بموقع با اشکالات امنیتی، مسائلی را به دنبال خواهد داشت که ضرر آن متوجه همه کاربران و قابلیت های عملیاتی این سیستم ها می گردد و عملاً زیر ساخت اطلاعاتی سازمان ها را در معرض آسیب و تهدید جدی قرار می‌دهد. در این مقاله قصد داریم درخصوص جنبه های ایمن سازی انبارهای داده و کلان داده ها تحقیق نموده و راهکارهای موثر و منطقی این حوزه را بررسی نمائیم.

واژگان کلیدی:
مدیریت امنیت
قابلیت اطمینان
پایگاه داده تحلیلی
اطلاعات بزرگ
رمزگذاری داده ها

تعداد جداول	تعداد شکل ها	تعداد مراجع
۰	۰	۲۴

مقدمه

در جهان کنونی پایگاههای داده[1] و متعاقبا سیستم‌های مدیریت پایگاه داده[2] به عنوان یک بخش جدایی‌ناپذیر در سازمان‌ها و شرکت‌های مختلف مورد استفاده قرار می‌گیرند. کلان‌داده (مِه‌داده) دارایی اطلاعاتی در حجم، سرعت و یا تنوع بالا به شمار می‌آید که نیازمند روش نوآورانه و مقرون به صرفه پردازش اطلاعات است که بینش ارتقا یافته، تصمیم‌سازی و خودکارسازی فرآیندها را امکان‌پذیر می‌سازد. اکنون باید شفاف باشد که «کلان» در کلان‌داده تنها به حجم مربوط نیست .در حالیکه کلان‌داده (مِه‌داده) قطعا دربرگیرنده داده‌های زیادی است، اما عبارت کلان‌داده تنها به حجم اشاره ندارد. این یعنی در صورتی که مساله‌ای کلان‌داده باشد، تنها بحث تحلیل حجم انبوهی از داده‌ها مطرح نیست، بلکه داده‌ها با سرعت تولید می‌شوند و در قالب‌های پیچیده از منابع داده گوناگونی هستند . انبار داده یا همان پایگاه داده تحلیلی نیز فرآیندی برای جمع آوری و مدیریت داده‌ها از منابع مختلف برای ارائه بینش معنادار تجاری است. انبار داده به طور معمول برای اتصال و تجزیه و تحلیل داده‌های تجاری از منابع ناهمگن استفاده می‌شود. انبار داده، هسته اصلی سیستم BI (هوش تجاری) است که برای تحلیل و گزارش داده ساخته شده است. مدیریت امنیت در کلان داده و انبارهای داده مرتبط به دلیل نوین بودن زیرساخت ها از پیچیدگی های بخصوصی برخوردار است. به همین دلیل، حفظ امنیت پایگاه داده یکی از مهم‌ترین موضوعاتی به حساب می‌آید که لازم است کسب‌وکارها به آن توجه ویژه داشته باشند. امنیت پایگاه داده به اقدامات مختلفی اطلاق می‌شود که سازمان‌ها از آن‌ها برای اطمینان از حفظ شدن پایگاه‌های اطلاعاتی خود در برابر تهدیدات داخلی و خارجی استفاده می‌کنند. منظور از امنیت پایگاه داده، محافظت از خود پایگاه داده، داده‌های موجود در آن، سیستم مدیریت پایگاه داده مربوطه و برنامه‌های کاربردی مختلفی است که دسترسی به آن‌ها در ارتباط با بانک اطلاعاتی وجود دارد [۱]. سازمان‌ها باید پایگاه‌های اطلاعاتی را در برابر حملات عمدی گوناگون مانند تهدیدات امنیت شبکه و همچنین سو استفاده از داده‌ها و پایگاه‌های اطلاعاتی ایمن کنند. در طول چند سال گذشته، میزان نقض اطلاعات و قانون‌شکنی در این زمینه به طور قابل توجهی افزایش پیدا کرده است. علاوه بر آسیب قابل توجهی که این تهدیدها به شهرت و اعتبار یک شرکت وارد می‌کنند، مقررات و مجازات‌های مختلفی برای نقض داده‌ها وجود دارند و لازم است سازمان‌ها با چالش نقض اطلاعات مقابله کنند. یکی از این موارد، مقررات عمومی حفاظت از داده‌ها به حساب می‌آیند که غالباً بسیار پرهزینه هستند. با توجه به نکات مذکور، می‌توان با قاطعیت، امنیت پایگاه داده موثر را برای سازگاری، حفاظت از اعتبار سازمان‌ها و حفظ مشتریان آن‌ها به عنوان یک امر کلیدی در نظر گرفت [۲]. چالش‌های امنیتی پایگاه داده از طیف وسیعی از تهدیدات امنیتی ناشی می‌شود. خطرات احتمالی را در نظر داشته باشید و قبل از وقوع به هر موضوع بپردازید. مشکلات رایج نگهداری پایگاه داده عبارت‌اند از:

- خطای انسانی و کاربری
- تهدیدهای داخلی که به تنهایی مقوله گسترده‌ای است.
- آسیب‌پذیری‌های نرم‌افزاری
- بدافزارها
- حملات سایبری
- امنیت مکان فیزیکی

همچنین در این حوزه خطرات احتمالی متعدد و مختلفی برای امنیت مخازن داده و پایگاه داده های مرتبط وجود دارند که برخی از پراهمیت ترین آن‌ها در ادامه فهرست شده‌اند [۳-۴]. اولین و به طور بالقوه، خطرناک‌ترین تهدیدی که امنیت پایگاه داده را به خطر می‌اندازد، دسترسی غیرمجاز هکرها و دستکاری‌کنندگان به سیستم‌های امنیتی و ایجاد مخاطره در اطلاعات مهم کاربر خارج از پایگاه داده است .آن‌ها به نوبه خود می‌توانند یا در نهایت به پایگاه داده آسیب برسانند یا سوابق را به گونه‌ای دستکاری کنند تا بتوانند به اهداف شوم خود برسند. حملات مختلف از طریق نرم‌افزار، اسکریپت یا سایر سیستم‌های غیرقانونی بالقوه مضر که شامل استفاده از بدافزارها و ویروس‌ها می‌شوند. این مسئله به هکرها اجازه دسترسی غیرمجاز به سیستم‌های پایگاه داده را می‌دهد. ممکن است تمام تهدیدات فوق منجر به بروز سربار سیستم، عملکرد نادرست برنامه‌های مختلف و قطع دسترسی مدیر مجاز به سیستم شود. اگر فایل‌های آلوده حذف یا از سیستم سرور پاک نشوند، ممکن است منجر به بروز آسیب‌های فیزیکی مختلفی مانند داغ شدن بیش از حد یا حتی خرابی کامل در موارد شدید شوند. علاوه بر موارد فوق، خرابی داده‌ها می‌تواند در موارد نقض یا تهدید در کنترل‌های امنیتی مختلفی رخ دهد که در وهله اول برای جلوگیری از وقوع چنین حوادثی به وجود آمده‌اند. به طور کلی، روش‌های متعددی وجود دارند که از طریق آن‌ها می‌توان امنیت پایگاه داده را به خطر انداخت یا هک و دستکاری کرد . این موارد همگی عواقب شدیدی را به دنبال دارند.

جنبه های ایمن سازی

امنیت پایگاه داده به اقدامات مختلفی اطلاق می‌شود که سازمان‌ها از آن‌ها برای اطمینان از حفظ شدن پایگاه‌های اطلاعاتی خود در برابر تهدیدات داخلی و خارجی استفاده می‌کنند .منظور از امنیت پایگاه داده، محافظت از خود پایگاه داده، داده‌های موجود در آن، سیستم مدیریت پایگاه داده مربوطه و برنامه‌های کاربردی مختلفی است که دسترسی به آن‌ها در ارتباط با بانک اطلاعاتی وجود دارد [۵] . سازمان‌ها باید پایگاه های اطلاعاتی را در برابر حملات عمدی گوناگون مانند تهدیدات امنیت شبکه و همچنین سو استفاده از داده‌ها و پایگاه‌های اطلاعاتی ایمن کنند. به طور کلی، ایمن سازی پایگاه داده سه جنبه کلیدی را در بر می گیرد که در ادامه به آن‌ها پرداخته می‌شود:

[1] Database

[2] Database Management System (DBMS)

حفظ محرمانگی اطلاعات[1]

در مفاهیم امنیت پایگاه داده، حفظ محرمانگی اطلاعات به عنوان اولین معیار و جنبه ایمن سازی در نظر گرفته می‌شود. امکان ایجاد محرمانگی از طریق رمزنگاری داده‌های ذخیره شده در پایگاه داده امکان‌پذیر است. رمزنگاری یک روش یا فرآیندی است که در آن داده‌ها کدگذاری می شوند. این کدگذاری به گونه‌ای انجام می‌شود که تنها کاربران مجاز امکان خواندن داده‌ها را داشته باشند. به بیان دیگر، رمزنگاری یعنی داده های حساس برای کاربران غیرمجاز به صورت غیرقابل خواندن هستند. الگوریتم‌های رمزنگاری مختلفی مانند DES، AES و Triple DES برای برقراری و حفظ محرمانگی در پایگاه داده استفاده می‌شوند [۶-۷] .

تمامیت[2]

تمامیت و جامعیت پایگاه داده‌ها به معنی صحت، دقت و سازگاری داده‌های ذخیره شده در پایگاه داده‌ها در تمام لحظات است. هر سیستم پایگاه داده باید بتواند جامعیت پایگاه داده را کنترل و تضمین کند. برای اطمینان از تمامیت پایگاه داده روش‌هایی وجود دارند که در ادامه به آن پرداخته می‌شود [۸-۱۰] :

- پس از نصب پایگاه داده، باید رمز عبور تغییر داده شود . علاوه بر این، بررسی‌های دوره‌ای گوناگونی لازم است تا این اطمینان به وجود بیاید که رمز عبور در خطر قرار نگرفته است.
- باید آن دسته از حساب‌های کاربری که استفاده نمی‌شوند، قفل شوند. در شرایطی که یک حساب کاربری به طور قطعی هیچ‌گاه دوباره استفاده نخواهد شد، بهترین اقدام حذف آن است.
- لازم است سیاست‌های پیشرفته مختلفی برای توسعه رمزهای عبور قوی ایجاد شوند. یکی از ایده‌های کارامد در این خصوص، الزام در تغییر رمز عبور به صورت ماهانه است.
- بررسی نقش‌ها و تنظیم دسترسی‌ها بر اساس آن‌ها بسیار اهمیت دارد. در واقع، باید این اطمینان حاصل شود که کاربران تنها به مواردی دسترسی دارند که مجاز به استفاده از آن‌ها هستند. با وجود اینکه بررسی این موضوع برای پایگاه داده‌های بزرگ بسیار زمان‌بر است، اما اگر دسترسی‌ها به درستی تنظیم شوند، ورود یا دسترسی غیرمجاز به راحتی قابل بررسی خواهد بود. بررسی اینکه آیا کسب و کار مربوطه چندین ادمین پایگاه داده دارد یا خیر؛ در صورتی که پاسخ این سوال مثبت باشد، بهتر است وظایف میان این مدیران پایگاه داده تقسیم شوند.

مفهوم تمامیت در امنیت پایگاه داده از طریق تنظیمات مربوط به کنترل های دسترسی کاربری اعمال می‌شود. با استفاده از این مفهوم، به هر کاربر دسترسی به پایگاه داده تا سطح مورد نیاز داده خواهد شد. به عنوان مثال، ممکن است به یک کارمند اجازه دیدن رکوردها و تغییر بخش‌هایی از اطلاعات، مثل جزییات شماره تماس داده شود، اما کارمند بخش منابع انسانی دسترسی‌های بیش‌تری داشته باشد.

دسترس پذیری[3]

دسترس پذیری به این معنی می باشد که داده ها، پایگاه های داده و سیستمهای حفاظت امنیت، در زمان مناسب نیاز به اطلاعات در دسترس باشند. در یک سیستم مدیریت پایگاه داده کارآمد، نباید پایگاه داده از کارافتادگی بازه‌ای داشته و نرخ دسترس پذیری آن باید قابل قبول باشد . مهمترین تمهیداتی که می توان درجهت دستیابی به دسترس پذیری در پایگاه داده اتخاذ نمود شامل موارد زیر می باشد [۱۱-۱۳] :

- محدود کردن میزان فضای ذخیره‌سازی برای کاربران در پایگاه داده
- ایجاد محدودیت در تعداد نشست‌های های موازی قابل دسترسی برای هر کاربر پایگاه داده
- پشتیبانی‌گیری از داده‌ها به صورت دوره‌ای به منظور کسب قابلیت بازیابی داده در صورت بروز مشکلاتی در اپلیکیشن
- ایجاد ایمنی در پایگاه داده در برابر آسیب‌های امنیتی
- استفاده از پایگاه داده‌های خوشه‌ای با هدف افزایش دسترسی پذیری

جنبه های ایمن سازی انبار داده و کلان داده

انبار داده[4] مجموعه بزرگی از داده‌های تجاری است که به سازمان‌ها و کسب‌وکارها کمک می‌کند تا در تصمیم‌گیری‌های خود دقیق‌تر و هوشمندانه‌تر عمل کنند. انبار داده مفهوم جدیدی نیست و از دهه ۱۹۸۰ وجود داشته است. حجم زیادی از داده‌های موجود در انبارهای داده از منابع مختلفی جمع‌آوری می‌شوند؛ برنامه‌های کاربردی داخلی مانند بازاریابی، فروش و امور مالی نمونه‌هایی از این منابع هستند. لازم به ذکر است، مراکز داده، مکان‌های فیزیکی هستند که سرورها در آنها نگهداری می‌شوند؛ در حالی که انبار داده، یک مفهوم نرم‌افزاری و در واقع یک داده ساختار بر روی یک یا چند سرور است. به بیان ساده‌تر، مرکز داده یک اتاق فیزیکی یا ساختمانی است که سرورهای داده و کامپیوترها در آن قرار می‌گیرند. در حالی که یک انبار داده فقط نوعی پایگاه داده نرم‌افزاری است که برای گزارش‌گیری و تجزیه‌وتحلیل داده‌ها استفاده می‌شود و یکی از اجزای اصلی هوش تجاری به شمار می‌آید [۱۴] . از دیگر مواردی که ممکن است با انبار داده اشتباه گرفته شود پایگاه داده است. اساسا پایگاه داده سامانه‌ای اطلاعاتی است که وضعیت حال حاضر یک سامانه نرم‌افزاری را ثبت می‌کند و مقادیر داده‌ها

[1] Confidentiality
[2] Integrity
[3] Availability
[4] Data Warehouse

به صورت مرتب و مکرر در حال تغییر و به روز رسانی است. در حالی که انبار داده یک سامانه اطلاعاتی است که داده‌های تاریخی را از منابع مختلف در خود گردآوری و تجمیع می‌کند. پایگاه داده برای ذخیره و بازیابی مکرر داده‌های معین و انجام تراکنش‌های آنلاین (OLTP) طراحی شده است در حالی که انبار داده برای تجزیه‌وتحلیل تجمیعی داده‌ها (OLAP)کاربرد دارد [۱۶-۱۵] . انبارهای داده این مزیت کلی را ارائه می‌کنند که به سازمان‌ها اجازه می‌دهند تا حجم زیادی از داده‌های مختلف را تجزیه‌وتحلیل کرده و اطلاعات ارزشمند و قابل توجهی از آن‌ها استخراج کنند.

امنیت کلان داده به دلیل تنوع بسیار سبک های آن، در سازمانها از دو نظر داده های شرکتی و داده های مشتری بررسی می شود. سازمان‌ها باید از امنیت داده‌های کسب‌وکار آگاه باشند، زیرا از دست دادن داده‌های اختصاصی و محرمانه برابر است با افشای نقاط ضعف شما در برابر رقبا و دادن مزیت رقابتی به آنها. در کنار داده‌های شرکتی، شرکت‌ها داده‌های حساس مشتریان خود را نیز حفظ می‌کنند. در صورتی که این اطلاعات به دست عوامل مخرب بیفتد، چه بسا اتفاقات بحرانی زیادی رخ خوهد داد. نشت داده‌های کاربران می تواند پیامدهای عمده‌ای برای شرکت ها از نظر هزینه غرامت و از دست دادن مشتریان به رقبایی که خدمات مشابه اما با سطوح امنیتی بالاتر را ارائه می‌دهند، داشته باشد. در ادامه برخی مشکلات رایج امنیت داده در این حوزه و راه حل های آنها به طور خلاصه آمده است.

ناامنی محیط شبکه

هنگامی که اطلاعات وارد شبکه یک شرکت می شود، این اطلاعات از طرق مختلف بررسی امنیتی می شوند. اما اگر سیستم شما نتواند داده‌های بالقوه مخرب را از بین داده‌های ورودی جدا کند، در آینده آسیب های زیادی به زیرساخت و معماری اطلاعات وارد شده و خسارت بسیار به سازمان وارد خواهد شد.

حملات سایبری به طور مداوم در حال توسعه و پیچیده تر شدن هستند، بنابراین سازمان ها باید در مدیریت امنیت کلان داده تجدید نظر و راه حل های معماری اعتماد صفر را اتخاذ کنند. با درنظر گرفتن یک محیط و رویکرد مبتنی بر هویت دیجیتال پویا که در آن هر کاربر یا دستگاه، برنامه یا سیستم، خارج و داخل شبکه، به طور پیش فرض به عنوان غیرقابل اعتماد طبقه بندی می شود. در این حالت، منابع سازمان‌ها بدون در نظر گرفتن موقعیت مکانی آنها ایمن می‌شوند، زیرا کنترل دسترسی از محیط به هر دستگاه و کاربر منتقل می‌شود و شبکه به بخش‌های خرد تقسیم می‌شود تا حمله هکرها را سخت‌تر کند [۱۷] .

حملات مهندسی اجتماعی

یک نمونه کلاسیک از این نوع حمله به عنوان فیشینگ شناخته می شود. فیشینگ زمانی رخ می دهد که مهاجمان پیام‌هایی را برای شما ارسال می‌کنند که به نظر می‌رسد از یک منبع شناخته شده و قابل اعتماد هستند، اما در واقع، مخرب هستند. اگر کارمند شما ناآگاهانه روی یک لینک غیرقابل اعتماد کلیک کند، هکرها می توانند به شبکه شرکت دسترسی پیدا کنند. برای جلوگیری از ایمیل‌های کلاهبرداری که ممکن است حاوی هرزنامه، بدافزار یا تلاش‌های فیشینگ باشند، فایروال پیاده‌سازی کنید. دروازه های امن فایروال به دلیل عملکردهای آنتی ویروس، ضد اسپم و ضد فیشینگ برای شناسایی ایمیل های بد مفید هستند. اما اگر پیوند مخرب در ایمیل نباشد بلکه در یک فایل PDFپیوست شده باشد، فایروالها ایمیل مشکوک را شناسایی و آن را قرنطینه نمی‌کنند. برای رویارویی با مشکلات و تهدیدات امنیتی فعلی داده ها، از ویژگی های اضافی که فایروال ها می توانند در اختیار شما قرار دهند، مانند سند باکسینگ بهره ببرید. در این صورت یک کپی ایمن ایزوله از محیط واقعی شما که در آن می توانید نامه های بالقوه مخرب را بدون تأثیرگذاری بر سیستم یا پلتفرمی که روی آن اجرا می شود باز کنید، خواهید داشت [۱۸] .

مشکلات پاک سازی داده ها

کلان داده به مزایای امیدوارکننده ای مانند بهبود تصمیم گیری معروف است، اما برای به دست آوردن داده های ارزشمند، ابتدا باید گندم را از کاه جدا کنید و با داده های کثیف و نامرتب مقابله کنید. در غیر این صورت، در یک چرخه معیوب از داده های ضعیف و یک تله زباله در زباله گیر خواهید کرد. بنابراین یک فرآیند پاکسازی خودکار داده‌ها وارد میدان جنگ می‌شود تا داده‌های صحیح، کامل و با فرمت مناسب را از میان انبوهی از دادهها به شما ارائه دهد. اما اگر آن ابزارها به درستی پیکربندی نشده باشند، پاکسازی داده ها منجر به داده های متناقض می شود و نگرانی های امنیتی تجزیه و تحلیل داده های بزرگ به جایی نمی رسد. یک الگوریتم ممکن است در مرحله طبقه‌بندی داده‌ها شکست بخورد و یا ممکن است داده‌های حساس را عادی تعریف کند و آن را با طیف وسیعی از افراد به اشتراک بگذارد.

معیارهای پوشش داده های ناقص

سازمان‌ها سیاست‌های ماسکینگ داده‌ها یا داده پوشانی را اتخاذ می‌کنند تا داده‌هایی را که مشتریان را شناسایی می‌کند مانند ویژگی‌های انحصاری یک فرد که شامل تاریخ تولد، نام، سن و ... می باشد را از اطلاعات محرمانه مربوط به آن‌ها که اغلب شامل داده‌هایی که به مشتری مرتبط ، اما قابل تغییر هستند متمایز می‌کنند. مانند آدرس منزل، شماره گواهینامه رانندگی، شماره حساب بانکی. هدف از داده پوشانی پیشگیری از خطرات امنیتی کلان داده از طریق متوقف کردن مجرمان سایبری در تطبیق مشتریان با اطلاعات حساس آنها است. اگر داده پوشانی به اشتباه انجام شود، می تواند توسط هکرها معکوس شود. یک راه حل کلیدی برای ایمن سازی کلان داده ها، رمزگذاری داده ها است .اما اجرای موثر آن زمان می‌برد. بنابراین چگونه می‌توان سرعت پردازش داده‌ها را افزایش داد (از آنجایی که سرعت پردازش داده یکی از پنج هسته اصلی داده‌های بزرگ به همراه تنوع، حجم، ارزش و صحت است). از این تکنیک های رمزگذاری داده ها در

آخرین مرحله پردازش داده ها برای اطمینان از امنیت داده های بزرگ استفاده کنید. سایر رویکردهای مفید عبارتند از [۱۹]:

- Data scrambling از تکنیک‌های رایج داده‌پوشانی می باشد .در این تکنیک کاراکترهای داده ورودی به طور تصادفی سازماندهی مجدد شده و در ذخیره سازی داده جایگزین می شوند.
- Data substitution یا تکنیک جایگزینی داده ها ، داده‌های جعلی جایگزین اطلاعات واقعی می‌شوند، می‌توانید به جای نام‌های واقعی مشتریان از نام‌های تصادفی از دفترچه تلفن استفاده کنید.

در هر دو مورد، داده های اصلی هنوز در انبار یا دریاچه داده شما موجود است. شما از اطلاعات جایگزین یا درهم برای تصمیم‌گیری استفاده می‌کنید، زیرا نمی‌توانید از داده‌های حساس برای آن مقاصد طبق مقررات حفاظت از داده‌ها از جمله GDPR و D-DPA استفاده کنید. علاوه بر این، برای اینکه تلاش‌های امنیتی شما از بین نرود، مطمئن شوید که کنترل‌های دسترسی را به گونه‌ای تنظیم کنید که تنظیمات الگوریتم داده پوشانی خاص فقط برای دارندگان داده در بخش‌های مربوطه در دسترس باشد و نه هیچ‌کس دیگری.

تولید داده های جعلی

حملات جرایم سایبری می‌تواند منجر به اعداد جعلی و در داشبورد نمایش داده شود. همانطور که در مورد آمازون اتفاق افتاد، زمانی که الگوریتم‌های سایت با بررسی‌های محصولات جعلی دستکاری شدند و به طور مصنوعی محصولات و فروشندگان خاصی را بیش از حد ارزیابی می‌کردند. اگر نتوانید تصویر واقعی را ببینید، در نهایت بینش های اشتباهی دریافت خواهید کرد و تصمیمات ناقصی می گیرید که می تواند منجر به مشکلات امنیتی کلان داده شود. داده‌های قدیمی نیز می توانند جعلی شوند و بر فرآیند تصمیم‌گیری و در نتیجه بر عملیات تجاری شما تأثیر منفی بگذارند. نمونه ای از تاثیر منفی دادههای قدیمی ، خطوط هوایی یونایتد بود .آنها سالانه یک میلیارد دلار از دست دادند زیرا یک مدل قیمت گذاری نادرست مبتنی بر ترجیحات صندلی مسافران مربوط به ۱۰ سال قبل را استفاده می کردند. اگر نتوانید داده های جعلی یا قدیمی را در مخزن مرکزی خود شناسایی کنید، توانایی شما برای ایمن سازی داده ها و محافظت از داده های مشتری و شرکت کاهش می یابد. از مدل های یادگیری ماشین (ML) برای یافتن ناهنجاری ها در داده های خود استفاده کنید و رویکرد تشخیص تقلب را اعمال کنید. سیستم‌های تشخیص کلاهبرداری مبتنی بر ML می‌توانند نرخ دقت تشخیص را تا ۹۰٪ بهبود بخشند و زمان بررسی تقلب را تا ۷۰٪ کاهش دهند [۲۰].

تغییرات غیرمجاز در ابرداده

با در نظر گرفتن حجم عظیمBig Data ، تغییرات غیرمجاز متادیتا ، مدیریت تغییرات در «داده‌های مربوط به سایر داده‌ها» و شناسایی اطلاعات مربوطه پس از آن را چالش‌برانگیز می‌کند، زیرا مطمئناً نمی‌دانید کدام تغییرات قابل اعتماد هستند. برای حذف دسترسی غیرمجاز و پیامدهای ناخوشایند آن، مانند مجموعه داده های اشتباه و منابع داده غیرقابل ردیابی اجرای کنترل دسترسی کاربر و فرآیندهای مجوز اجباری برای کارمندان مفید است. همچنین استفاده از nulling outیا جایگزینی داده های حساسِ با ارزشِ تهی تا در محیط های آزمایشی یا برنامه نویسی. از آنجایی که حتی اطلاعات حساس اولیه مانند نام نویسنده سند، تاریخچه ویرایش، نوع نرم‌افزار، در دستان اشتباه می تواند منجر به نقض احتمالی داده‌ها شود، بهتر است از یک اقدام دیگر برای ایمن‌سازی ابرداده‌های خود استفاده کنید . پاک‌سازی. این فرآیند حذف داده های حساس از سند است. پس از پاکسازی، فایل ممکن است برای مخاطبان گسترده تری توزیع شود. اگر با کپی کردن فایل قبلی، پیشنهادی را برای مشتری آماده کنید، تغییرات غیرمجاز در متادیتا نیز می‌تواند امنیت داده‌ها را به خطر بیندازد و اوضاع را ناخوشایند کند. اگر سند از طریق پاکسازی نرود، مشتری فرضی شما به تاریخچه تغییرات دسترسی خواهد داشت و اصلاحات بودجه اصلی یا محدوده کار مشتری قبلی شما را پیدا می کند. بنابراین، اهمیت این فعالیت را دست کم نگیرید. از ابزارهای پاکسازی خودکار داده ها استفاده کنید تا اطمینان حاصل کنید که فقط می توان به اطلاعات مورد نظر دسترسی داشت [۲۱].

بی احتیاطی کاربران

گاهی اوقات هکرها و سایر تهدیدات امنیتی خارجی در رتبه دوم پس از خطاهای انسانی از نظر آسیب وارد شده به شرکت قرار می گیرند. یک سطح سهل انگاری تصادفی کارکنان نیز همراه با افزایش تعداد دستگاه هایی که به داده های حساس شرکت و مشتریان دسترسی دارند، افزایش می یابد. اگر کارکنان شما بتوانند به داده‌ها از دستگاه‌های شخصی و از طریق شبکه‌های ناامن دسترسی داشته باشند، اشتراک‌گذاری فایل با طرف‌های غیرمجاز، به‌طور تصادفی یا بدخواه، به آسانی انگشتان دستتان است. جنبه‌های فنی راه‌حل که به شما کمک می‌کند نقض‌های امنیتی را به موقع شناسایی کنید و امنیت داده‌های بزرگ را تضمین کنید، شامل اجرای احراز هویت چندلایه و رویکرد تشخیص تهدید داخلی برای دریافت اعلان‌ها درباره تهدیدات امنیتی توسط کارکنان است. در مورد بخش «افراد»، فضایی ایجاد کنید که در آن کارکنان احساس امنیت کنند تا درباره یک دستگاه گم شده یا دزدیده شده گزارش دهند و بلافاصله این کار را انجام دهند.

ایمن سازی در پایگاه اوراکل

در دنیای امروز بسیاری از کسب و کارها بر مبنای اطلاعات بنا شده است، اهمیت دسترسی به اطلاعات بر هیچکس پوشیده نیست و محافظت همه جانبه از اطلاعات امری اجتناب ناپذیر می باشد. امنیت در بانک اطلاعاتی نیازمند یک استراتژی و برنامه مدون و دقیق است که بر

اساس آن نیازهای امنیتی شناخته شده و برای هریک راهکار مناسب پیاده سازی شود. رعایت نکات امنیتی در پیکربندی بانک اطلاعاتی، اعطای دسترسی مناسب به کاربران، اطمینان از پیکربندی‌های مناسب برای اتصالات، رمز نگاری حساس، نصب وصله‌های امنیتی از جمله وظایف مدیران بانک اطلاعاتی برای پاسخگویی به نیازهای امنیتی می باشد. بانک اطلاعاتی اوراکل یکی از قدرتمند ترین بانک‌های اطلاعاتی به لحاظ توانایی‌های امنیتی می‌باشد .این بانک با ارائه مجموعه وسیعی از امکانات بسیاری از نیازهای امنیتی را تا حد بسیار زیادی پوشش می دهد. بکارگیری دقیق این راهکارها می‌تواند درصد بالایی از امنیت را بدنبال داشته باشد [۲۲]:

رمزگذاری شفاف داده[1]

رمزگزاری شفاف داده که به اختصار بنام TDE نیز شناخته می‌شود یک راه کار ساده و شفاف برای رمز نگاری اطلاعات حساس می‌باشد. از مهمترین مزیت‌های این راهکار عدم نیاز به تغییر در برنامه کاربردی است. در این روش اطلاعات مورد نظر قبل از نوشته شدن در دیسک رمز نگاری شده و قبل از خواندن رمز گشایی می‌شوند. استفاده از این راهکار تاثیر بسیار ناچیزی در کارایی بانک اطلاعاتی خواهد داشت. TDE الگوریتم‌های استاندارد شامل AES, Triple DES را پشتیبانی می کند . نگارش اولیه TDE در اوراکل 10gR2 معرفی شده که در آن رمزنگاری برای جداول و ستونهای آنها ممکن بود . در اوراکل 11gR1 رمزنگاری درسطح Tablespace نیز پشتیبانی شد که به علت سادگی و عدم وابستگی به نوع ستون‌های جداول استفاده از آن توصیه شده و پیش فرض اوراکل می‌باشد [۲۳].

رمزگزاری شبکه[2]

با استفاده از این امکان می‌توان اطلاعات ارسالی بین سرور بانک اطلاعاتی و کلاینت‌ها را با الگوریتم‌های استاندارد مانند AES, Triple DES و RC4 رمز نگاری نمود .این مکانیزم همچنین از SSL[3] پشتیبانی می‌نماید.

احراز هویت قوی[4]

رمزهای عبور معمولی عموما مکانیزم‌های سختگیرانه ای برای اعتبارسنجی نیستند . رویکرد احراز هویت قوی یا SA با استفاده از مفاهیم امنیتی مانند Ticket،card، PIN entry و Token عملیات اعتبارسنجی را انجام می‌دهد . برخی تکنیک‌های SA اجازه می‌دهند که اعتبارسنجی چند فاکتوری را بکار ببریم . Oracle Advance Security انواع مختلفی از تکنیک‌های Strong Authentication را پشتیبانی می کند از جمله موارد زیر:

- Kerberos
- PKI (certificate-based authentication and encryption)
- RADIUS (Remote Authentication Dial-In User Service)

Oracle Databese Vault

تحقیقات نشان می‌دهد بیش از هشتاد درصد دسترسی‌های غیر مجاز به اطلاعات از طریق کارمندان داخلی شرکت‌ها انجام می‌شود. همچنین شصت و پنج درصد تهدیدات داخلی شناسایی نمی‌شوند Oracle Database vault . عمومی‌ترین راه کار برای مقابله و جلوگیری از تهدیدات داخلی و کاهش خطرات آن می‌باشد . از جمله کاربردهای آن جلوگیری از دسترسی کاربران با دسترسی بالا در سطح دیتابیس (DBA)بدون دسترسی از طریق Application می‌باشد . با استفاده از این راهکار امکان تعریف دسترسی‌های مجزا و تفکیک وظایف فراهم می گردد به بیان دیگر با بکارگیری آن به طور دقیق تعیین می‌گردد چه کسی، چه زمانی، کجا و چگونه می‌تواند به اطلاعات دسترسی داشته باشد.

پایگاه خصوصی مجازی

این راهکار در اوراکل 8i معرفی گردید. با استفاده از آن امکان سیاستگذاری برای دسترسی به داده‌ها در سطح سطرها و ستون‌های جداول فراهم می‌گردد .به بیان دیگر می‌توان برای هر کاربر تعیین کرد به چه بخشی از داده‌های موجود در Tableها،View ها و Synonymها دسترسی داشته باشد .بعنوان مثال می‌توان برای یک کاربر تعریف نمود در جدول مربوط به حقوق کارمندان در صورت استفاده از دستور Select فقط به اطلاعات برخی از کاربران خاص دسترسی داشته باشد و حقوق سایر کارمندان برای او نمایش داده نشود .همچنین با استفاده از این راهکار می‌توان دسترسی به داده‌ها را به ساعت خاصی در روز محدود نمود، بعنوان مثال کاربران فقط در ساعت اداری به داده‌ها دسترسی داشته باشند.

امنیت برچسب زنی اوراکل[5]

از این امکان بعنوان یک ابزار قدرتمند و در عین حال ساده برای کنترل دسترسی به داده‌های جداول پس از طبقه بندی و برچسب زدن بر روی انها استفاده می‌گردد . با این روش سطرهای حاوی داده‌های حساس را می‌توان از دسترس افراد حتی با داشتن دسترسی بر روی جداول بانک اطلاعاتی خارج ساخت . به بیان دیگر هر کاربر در صورت پرس و جو از جدول فقط داده‌هایی را که به طبقه بندی آن دسترسی دارد را مشاهده می‌نماید.

پشتیبان گیری ایمن اوراکل[6]

[1] Transparent Data Encryption
[2] Network Encryption
[3] Secure Sockets Layer (SSL)
[4] Strong Authentication
[5] Oracle Label Security
[6] Oracle Secure Backup

این ابزار یک سیستم مدیریت یکپارچه برای تهیه پشتیبان گیری حفاظت شده بر روی رسانه های خارجی مانند نوارها می‌باشد که از مزایای آن می‌توان به موارد زیر اشاره نمود [۲۴-۲۳]:

- ایجاد Backupامن با استفاده از مکانیزم کدگذاری 256 AESدر سطح سیستم عامل
- یکپارچگی با محیط Oracle Enterprise Manger و RMAN
- بهبود کارایی پشتیبان گیری بین 10تا 25%

ORACLE AUDIT VAULT AND DATABASE FIREWALL

این امکان اولین خط تدافعی برای بانک اطلاعاتی بوده و امکان ممیزی یکپارچه را بین سیستم عامل و بانک اطلاعاتی فراهم می‌نماید. از مزایای آن می‌توان به موارد زیر اشاره نمود:

- شناسایی و جلوگیری از حملات SQL injection
- محیط کاربری آسان برای مشاهده و تحلیل نتایج ممیزی‌ها
- شناسایی خودکار فعالیت‌های بدون مجوز که امنیت را تهدید می‌نماید

تحلیلگران کسب‌وکار، مهندسان داده و متخصصان ابزارهای هوش تجاری، کاربران SQL و سایر برنامه‌های تحلیلی به داده‌ها دسترسی دارند .تجزیه‌وتحلیل داده‌ها به کسب‌وکارها کمک می‌کند تا بتوانند در بازار پررقابت امروز، موفق‌تر عمل کنند و در صنعت خود بدرخشند. متخصصان، برای استخراج اطلاعات از داده‌ها، نظارت بر عملکرد کسب‌وکار و پشتیبانی، به گزارش‌ها، داشبوردها و ابزارهای تحلیلی متکی هستند. انبارهای داده با ذخیره‌سازی کارآمد داده‌ها این گزارش‌ها، داشبوردها و ابزارهای تحلیلی را در اختیار متخصصان داده قرار می‌دهند. در پایان لازم به ذکر است، کلان داده یک واژه برای مجموعه داده های حجیم، دارای ساختار بزرگ، بسیار متنوع و پیچیده با سختی هایی برای ذخیره سازی، تجزیه وتحلیل و بصری سازی می باشد .کلان داده کارآمد نباید تنها روی حجم، سرعت یا تنوع داده هاتمرکز کند، بلکه باید روی بهترین روش حفاظت داده ها تمرکز کند. با این حال، یک تناقض آشکار بین امنیت و حریم خصوصی کلان داده و استفاده گسترده از آن وجود دارد. تاکنون بسیاری ازتکنیک ها مانند تکنیک های مبتنی بر رمزنگاری و مبتنی بر گمنام سازی و سایر تکنیک ها برایحفظ حریم خصوصی و امنیت کلان داده پیشنهاد و پیاده سازی شده اند. اما متاسفانه به علت ویژگیهای اساسی کلان داده یعنی حجم، تنوع و سرعت بالا تمام این تکنیک ها به طور کامل مناسب نیستند. لذا امکان سنجی شبک ایمن سازی براساس ویژگی های سازمان و انبارداده های مرتبط عامل تعیین کننده در این حوزه می باشد.

نتیجه‌گیری

داده‌ها برای تصمیم‌گیری آگاهانه سازمان‌ها بسیار ضروری‌اند، بنابراین منطقی است که انبارهای داده برای هر سازمانی مهم باشند، زیرا همه داده‌ها را در خود ذخیره می‌کنند. بدون انبار داده نمی‌توان به جریان اطلاعات دسترسی داشت و از هوش تجاری در کسب‌وکار های تجاری و محیط های دانش محور بهره‌مند شد.

انبارهای داده کمک می‌کنند تا حجم زیادی از داده‌ها را در یک پایگاه داده مرکزی ذخیره نمود، آن‌ها را در یک مکان امن نگه‌داشت و در زمان نیاز، داده‌ها را برای نیازهای تجاری تجزیه‌وتحلیل نمود. امنیت پایگاه داده به اقدامات مختلفی اطلاق می‌شود که سازمان‌ها از آن‌ها برای اطمینان از حفظ شدن پایگاه‌های اطلاعاتی خود در برابر تهدیدات داخلی و خارجی استفاده می‌کنند. در این مقاله درخصوص ایمن سازی کلان داده و انبارهای داده و همچنین ایمن سازی پایگاه داده توزیعی اوراکل بررسی نموده و راهکارهای مفید را بیان نمودیم. در حال حاضر در ایران بسیاری از افراد و نهادها دربارهٔ کلان‌داده صحبت می‌کنند، اما دستاورد چندانی به چشم نمی‌خورد. نتیجه آنکه وضعیت کلان‌داده در ایران هنوز در حد مطلوب نیست و همچنان ضرورت سرمایه‌گذاری در این زمینه به‌شدت احساس می‌شود. شاید به‌رسمیت‌نشناختن اطلاعات و داده‌ها به‌عنوان سرمایه و موضوع کسب‌وکار از سوی دولت مهم‌ترین چالش این حوزه باشد . هر یک از سازمان‌ها، وزارتخانه‌ها و ادارات دولتی دارای بانک اطلاعاتی متنوعی هستند. با تجمیع و یکپارچه‌سازی بانک‌های اطلاعاتی و سرمایه‌گذاری در این رابطه و تحلیل هوشمند آن‌ها می‌توان در مسیری قرار گرفت که سیاست‌گذاری‌ها و برنامه‌ریزی‌های کشور واقع‌بینانه و با دقت بالایی صورت پذیرد.

تعارض منافع

«هیچ‌گونه تعارض منافع توسط نویسندگان بیان نشده است»

منابع و مآخذ

[1] Paul P, Aithal PS. Database security: An overview and analysis of current trend. International Journal of Management, Technology, and Social Sciences (IJMTS). 2019 Oct 30;4(2):53-8.

[2] Omotunde H, Ahmed M. A Comprehensive Review of Security Measures in Database Systems: Assessing Authentication, Access Control, and Beyond. Mesopotamian

[12] Krishnan K. Data warehousing in the age of big data. Newnes; 2013 May 2.

[13] Cardenas AA, Manadhata PK, Rajan SP. Big data analytics for security. IEEE Security & Privacy. 2013 Dec 12;11(6):74-6.

[14] Venkatraman S, Venkatraman R. Big data security challenges and strategies. AIMS Mathematics. 2019 Jan 1;4(3):860-79.

[15] Tiwari AK, Chaudhary H, Yadav S. A review on Big Data and its security. In2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS) 2015 Mar 19 (pp. 1-5). IEEE.

[16] Alguliyev R, Imamverdiyev Y. Big data: Big promises for information security. In2014 IEEE 8th International Conference on Application of Information and Communication Technologies (AICT) 2014 Oct 15 (pp. 1-4). IEEE.

[17] Cuzzocrea A, Bellatreche L, Song IY. Data warehousing and OLAP over big data: current challenges and future research directions. InProceedings of the sixteenth international workshop on Data warehousing and OLAP 2013 Oct 28 (pp. 67-70).

[18] Jain P, Gyanchandani M, Khare N. Big data privacy: a technological perspective and review. Journal of Big Data. 2016 Dec;3:1-25.

[19] Altıntaş B. A security comparison of Oracle, SQL Server and MySQL database management systems against SQL injection attack vulnerabilities.

[20] Rjaibi W, Hammoudeh M. Enhancing and simplifying data security and privacy for multitiered applications. Journal of Parallel and Distributed Computing. 2020 May 1;139:53-64.

[21] Gedam MN, Meshram BB. DATABASE PRIVATE SECURITY JURISPRUDENCE: A CASE STUDY USING ORACLE.

Journal of CyberSecurity. 2023 Aug 7;2023:115-33.

[3] Teimoor RA. A review of database security concepts, risks, and problems. UHD Journal of Science and Technology. 2021 Oct 10;5(2):38-46.

[4] Kothari H, Suwalka AK, Kumar D. Various database attacks, approaches and countermeasures to database security. International Journal of Advance Research in Computer Science and Management. 2019;5(4):357-62.

[5] Samaraweera GD, Chang JM. Security and privacy implications on database systems in Big Data era: A survey. IEEE Transactions on Knowledge and Data Engineering. 2019 Jul 18;33(1):239-58.

[6] Santos RJ, Bernardino J, Vieira M. A survey on data security in data warehousing: Issues, challenges and opportunities. In2011 IEEE EUROCON-International Conference on Computer as a Tool 2011 Apr 27 (pp. 1-4). IEEE.

[7] Pan B, Stakhanova N, Ray S. Data provenance in security and privacy. ACM Computing Surveys. 2023.

[8] Ray S, Mishra KN, Dutta S. Big data security issues from the perspective of IoT and cloud computing: A review. Recent Advances in Computer Science and Communications. 2020;12(1):1-22.

[9] Diaz C. Database Security: Problems and Solutions. Stylus Publishing, LLC; 2022 Jul 28.

[10] Sagiroglu S, Sinanc D. Big data: A review. In2013 international conference on collaboration technologies and systems (CTS) 2013 May 20 (pp. 42-47). IEEE.

[11] Kaushik M, Jain A. Challenges to big data security and privacy. International Journal of Computer Science and Information Technologies (IJCSIT). 2014;5(3):3042-3.

[22] Pamulaparty L, Kumar TP, Varma PV. A Survey: Security Perspectives of ORACLE and IBM-DB2 Databases. International Journal of Scientific and Research Publications. 2013 Jan;3(1).

[23] Nguyễn ĐL. Database security in the cloud: Survey and deployment of oracle database security in the Amazon cloud.

[24] Kaleem M, Shi W. Demystifying pythia: A survey of chainlink oracles usage on ethereum. InFinancial Cryptography and Data Security. FC 2021 International Workshops: CoDecFin, DeFi, VOTING, and WTSC, Virtual Event, March 5, 2021, Revised Selected Papers 25 2021 (pp. 115-123). Springer Berlin Heidelberg.