الأمن السيبراني في البيانات الضخمة

إعداد: وداع عثمان – نغم رمضان

إشراف: م. منال زكار – أ.د. عمار زقزوق

قسم هندسة الحاسوب – كلية الهندسة

الملخص:

مع التطور المستمر للتكنولوجيا والبيانات، أصبح من المهم حماية المعلومات الحساسة. ومع زيادة حجم البيانات، أصبحت البيانات الضخمة محاوراً أساسية للمؤسسات في مختلف الصناعات. ومع ذلك، مع الإمكانات الهائلة للبيانات الضخمة، تأتي تحديات أمنية مختلفة تتطلب اهتماماً دؤوباً وحلولاً استراتيجية. في هذا البحث، سنتعمق في أمن البيانات الضخمة ونستكشف أفضل الممارسات وطرائق التحليل للمؤسسات لتحصين بياناتها واكتشاف التهديدات المحتملة والأنشطة الضارة، بالإضافة إلى كيفية عمل أمان البيانات الضخمة، وما هي بعض التحديات التي تواجهنا في هذا المجال.

الكلمات المفتاحية: البيانات الضخمة – الأمن السيبراني – حماية البيانات.

المقدمة:

مع تطبيق تقنيات الإنترنت والهواتف المحمولة السائدة في الحياة اليومية، بما في ذلك الشبكات الاجتماعية وإنترنت الأشياء (Internet of Things: IoT) والخدمات الشخصية، يتم بشكل مستمر جمع كميات هائلة من البيانات وتخزينها وتحليلها واستخدامها في منصات مختلفة بما في ذلك السحابة من قبل الأفراد والمنظمات. البيانات الضخمة، هو مصطلح يستخدم لمثل هذه المجموعة من مجموعات البيانات الكبيرة، والتي تمتلك خصائصاً نموذجية مثل: سريعة الحركة ومتعددة المصادر وكبيرة للغاية وغير منظمة. تحدد هذه الميزات الأبعاد الثلاثة المعروفة للبيانات الضخمة، وهي السرعة والتنوع والحجم، والتي يشار إليها باسم Vs3. يتم إنتاج البيانات الضخمة من مواقع الويب المختلفة وأرشيفات الوسائط المتعددة والشبكات الاجتماعية وشبكات إنترنت الأشياء. هذه البيانات الضخمة تواجه تحديات خطيرة تتعلق بالأمن والخصوصية بسبب الخصائص النموذجية التي تتمتع بها Vs3 المرتبطة بالبنى التحتية السحابية واسعة النطاق وإنترنت الأشياء وآليات الأمن والخصوصية التقليدية غير الكافية وغير القادرة على التعامل مع الانفجار السريع للبيانات في مثل هذه البيئة الحاسوبية الموزعة.

مفهوم البيانات الضخمة:

هي عبارة عن مجموعة من البيانات الكبيرة جداً والمعقدة، تجمع بين البيانات المنظمة وشبه المنظمة وغير المنظمة

التي تجمعها المنظمات. يمكن استخراجها للحصول على معلومات واستخدامها في مشاريع التعلم الآلي (Machine Learning: ML) والنمذجة التنبؤية وتطبيقات التحليلات المتقدمة الأخرى.

يتم استخدام ثلاثة أساسيات لوصف البيانات الضخمة، وهي:

1- الحجم الكبير من البيانات في العديد من البيئات.

2- مجموعة واسعة من أنواع البيانات المخزنة بشكل متكرر في أنظمة البيانات الضخمة.

3- السرعة التي يتم بها إنشاء الكثير من البيانات وجمعها ومعالجتها [1,2].

يتم الحصول على البيانات الضخمة من مصادر ووسائط مختلفة عبر الخدمات والأعمال التجارية عبر الإنترنت.

مفهوم الأمن السيبراني:

الأمن السيبراني (أو الأمن الإلكتروني) يشير إلى مجموعة من الإجراءات والتدابير التي تهدف إلى حماية الأنظمة الإلكترونية والشبكات الحاسوبية والبيانات من التهديدات السيبرانية. يعني الأمن السيبراني الحفاظ على سرية المعلومات الرقمية وسلامتها وتوفرها وتقنيات المعالجة الآلية والبنية التحتية للمعلومات.

يشمل الأمن السيبراني حماية الأنظمة والشبكات من الاختراق والوقاية من الهجمات الإلكترونية والتحكم في حماية البيانات والتعرف إلى التهديدات والتحقق من الهوية وإدارة الحوادث الأمنية وتعزيز الوعي الأمني للمستخدمين.

تشمل التهديدات السيبرانية التي يتعرض لها الأمن السيبراني هجمات القرصنة والفيروسات والديدان الإلكترونية وبرامج التجسس والاحتيال الإلكتروني والاختراقات الهجينة والهجمات الموجهة والهجمات المنسقة والتهديدات الداخلية والاعتداءات على الخصوصية.

يتطلب الأمن السيبراني التوازن بين الأمان والوظائف والسهولة في استخدام التكنولوجيا الرقمية. وبالتالي، يتضمن الأمن السيبراني تقنيات وأدوات متنوعة مثل تشفير البيانات وجدران الحماية (Firewalls) وأنظمة الكشف عن التسلل (Intrusion Detection Systems) والمصادقة متعددة العوامل (Multi-Factor Authentication) وتحديثات البرامج الأمنية المنتظمة وتدابير الوقاية الأخرى.

تهدف جهود الأمن السيبراني إلى حماية المعلومات الحساسة والحفاظ على الأنظمة الحاسوبية الحيوية والشبكات التي تدعم العمليات الحيوية للأفراد والشركات والمؤسسات والحكومات. الأمن السيبراني أصبح أمراً حيوياً في عصر التكنولوجيا المتقدمة، إذ تتزايد التهديدات السيبرانية وتعقيد الهجمات، وبالتالي يلعب الأمن السيبراني دوراً حاسماً في حماية الأنظمة والبيانات الحساسة وضمان الاستدامة والاستقرار في العالم الرقمي.

أمن البيانات الضخمة:

هو مجموعة التدابير والممارسات المطبقة لحماية كميات كبيرة من البيانات من الوصول غير المصرح به والانتهاكات والأنشطة الضارة.

يتضمن تأمين البيانات الضخمة ثلاث مراحل رئيسة هي:

1- ضمان النقل الآمن للبيانات من مواقع المصدر. عادةً، في السحاب للتخزين أو الاستيعاب في الوقت الفعلي.

2- حماية البيانات داخل طبقات التخزين لخط أنابيب البيانات الضخمة.

3- الحفاظ على خصوصية بيانات المخرجات، بما في ذلك التقارير ولوحات المعلومات التي تحتوي على رؤى تم الحصول عليها من تحليل البيانات باستخدام أدوات مثل Apache Spark [1,3].

تشير بنية الأمن السيبراني في البيانات الضخمة إلى الهيكل والمكونات الموضوعة لضمان الأمن والحماية. يتضمن مراحل وتدابير مختلفة لتقليل المخاطر وحماية البيانات الحساسة [4].

على الرغم من أن البنية المحددة قد تختلف وفقاً للمؤسسة ومتطلباتها، لكن هناك بعض المكونات والاعتبارات القياسية، وهي:

- يعد تشفير البيانات أمراً حيوياً في أمن البيانات الكبيرة. فهو يحول البيانات إلى تعليمات برمجية تتطلب فك تشفير الوصول ويعزز حماية البيانات أثناء التخزين والنقل والمعالجة، ما يمنع الوصول غير المصرح به أو التلاعب.
- يقوم التحكم في الوصول بإدارة الوصول إلى البيانات والإجراءات عبر المصادقة وأدوار المستخدم والأذونات.
- يعمل إخفاء البيانات وإخفاء الهوية على حماية البيانات الحساسة عن طريق استبدالها بمعلومات وهمية أو مشوشة.
- تعمل إجراءات منع فقدان البيانات (Data Loss Prevention: DLP) على منع فقدان البيانات أو تسربها.
- يعمل التخزين الآمن للبيانات على حماية البيانات غير النشطة من خلال الأنظمة الآمنة والتشفير والنسخ الاحتياطية وخطط التعافي من الكوارث.
- يعد أمان الشبكة أمراً مهماً لحماية البيانات أثناء النقل. وهي تتضمن بروتوكولات اتصال آمنة، وجدران الحماية، ومنع التطفل.
- تدقيق الأنشطة المتعلقة بالبيانات ومراقبتها، والكشف عن الانتهاكات المحتملة.
- تستخدم التحليلات الأمنية أساليباً متقدمة لاكتشاف التهديدات والمخالفات الأمنية ومعالجتها [5,6].

فوائد أمن البيانات الضخمة [3,4]:

- اتخاذ القرار بشكل أفضل.
- تحسين الكفاءة التشغيلية.
- إدارة المخاطر.
- التسويق والإعلان المستهدف.
- تعزيز فهم العملاء.
- ميزة تنافسية.
- ابتكار المنتجات والخدمات.

عمل الأمن السيبراني في البيانات الضخمة:

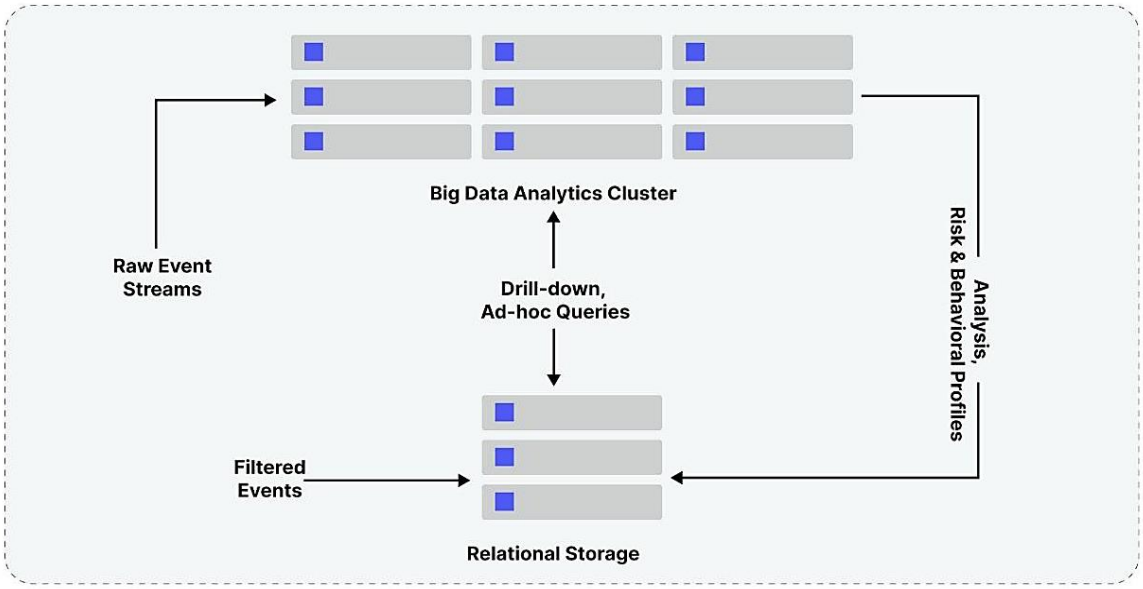
يهدف أمن البيانات الضخمة إلى منع الوصول غير المصرح به والتطفل باستخدام جدران الحماية، والمصادقة القوية للمستخدم، وتدريب المستخدم النهائي. ومع ذلك، تقدم بيئات البيانات الضخمة مستوى أعلى من التعقيد، لأن أدوات الأمان يجب أن تعمل عبر ثلاث مراحل متميزة للبيانات لا تتم مواجهتها عادةً في أمان الشبكات التقليدية، الشكل (1)، وهي:

1- مصادر البيانات الضخمة: مستمدة من مصادر وتنسيقات متنوعة، تشمل البيانات التي ينشئوها المستخدمون.

على سبيل المثال، رسائل البريد الإلكتروني ومنشورات وسائل التواصل الاجتماعي.

2- البيانات المخزنة: تتطلب حماية البيانات المخزنة مجموعة أدوات أمنية، بما في ذلك التشفير أثناء عدم النشاط والمصادقة القوية للمستخدم ومنع التطفل.

3- بيانات الإخراج: تم تصميم منصات البيانات الضخمة لإجراء تحليلات متطورة على مجموعات بيانات واسعة النطاق، وإنشاء رؤى قيمة يتم تقديمها من خلال التطبيقات والتقارير ولوحات المعلومات. ومع ذلك، تصبح هذه المعلومات الاستخباراتية هدفاً جذاباً للتطفل. لهذا السبب، يعد تشفير بيانات المخرجات، إلى جانب إدخال البيانات، وضمان الامتثال في هذه المرحلة أمراً بالغ الأهمية [7].



الشكل (1): دورة الأمن السيبراني في البيانات الضخمة.

يتم توجيه أمن البيانات الضخمة عبر مسار غير مباشر، ومن الناحية النظرية يمكن أن يكون عرضة للخطر في أكثر من نقطة واحدة.

ممارسات أمن البيانات الضخمة [8,9]:

- التشفير.
- التحكم في وصول المستخدم.
- مراقبة الأمن السحابي.
- إدارة المفاتيح المركزية.
- تحليل حركه مرور الشبكة.
- كشف التهديدات الداخلية.
- صيد التهديد.
- تحقيق الحادثة.
- تحليلات سلوك المستخدم.
- كشف تسرب البيانات.

أنواع ضوابط أمن البيانات:

لتأمين البيانات ومنع اختراقها، يجب اتباع إجراءات التحكم الآتية، الشكل (2) [2,6,3]:

1- صلاحية التحكم وصلاحية الدخول: من المهم الحد من الوصول المادي والرقمي إلى الأنظمة والبيانات المركزية. الهدف هو التأكد من أن جميع أجهزة الحاسوب والأدوات الذكية محمية بكلمة مرور.

2- المصادقة: قبل منح الوصول إلى البيانات، نقوم بتنفيذ إجراءات المصادقة، مثل قيود الوصول والتعرف الصحيح على الأشخاص.

3- النسخ الاحتياطية والتعافي من الكوارث: للوصول إلى البيانات بأمان أثناء فشل النظام أو الكوارث أو تلف البيانات أو الانتهاكات. لتسهيل الاسترداد، يجب تخزين نسخة بيانات احتياطية بتنسيق منفصل، مثل محرك أقراص ثابت.

4- محو البيانات: يعد مسح البيانات الذي يستخدم برنامجاً لمسح البيانات بالكامل من أي جهاز تخزين، طريقة أكثر أماناً من مسح البيانات التقليدي.

5- مرونة البيانات: فرض خصوصية البيانات بشكل فعال من خلال دمج المرونة في الأجهزة والبرامج تحمي البيانات من الاضطرابات الناجمة وانقطاع التيار الكهربائي.

6- التشفير: من خلال مفاتيح التشفير، تقوم خوارزمية الحاسوب بتحويل أحرف النص إلى نموذج غير مفهوم، ما يضمن أن الأفراد المصرح لهم فقط الذين لديهم المفاتيح اللازمة يمكنهم فتح المحتوى والوصول إليه.

الشكل (2): أنواع ضوابط أمن البيانات.

التحديات الرئيسة لأمن البيانات الضخمة:

يوفر حجم البيانات المتزايد باستمرار مزايا وعيوباً. يمكن أن يؤدي تحليل البيانات المحسن إلى اتخاذ قرارات أفضل للشركات، ولكنه يقدم أيضاً مخاوفاً أمنية، وخاصة عند التعامل مع المعلومات الحساسة.

فيمايلي بعض التحديات التي تواجه أمن البيانات الضخمة، والتي تحتاج المؤسسات إلى معالجتها:

1- تخزين البيانات: تعتمد الشركات بشكل متزايد على تخزين البيانات السحابية لإجراء عمليات مبسطة، ولكن هذه

المرحلة تأتي مع مخاطر أمنية. نتيجة لذلك، تختار العديد من شـركات التكنولوجيا الكبرى الجمع بين تخزين البيانات محلياً وسحابياً لتحقيق التوازن بين الأمان والمرونة. يتم تخزين البيانات المهمة في قواعد البيانات المحلية، ويتم وضع المعلومات الأقل حساسية في السحابة لتسهيل الوصول إليها.

2- بيانات وهمية: يشـكل إنشـاء البيانات المزيفة تهديداً كبيراً، لأنه يسـتهلك وقتاً ثميناً يمكن اسـتخدامه لمعالجة المشـكلات الأكثر إلحاحاً، ويجب على الشـركات فحص بياناتها بدقة باسـتخدام مجموعات بيانات الاختبار المختلفة لتقويم نماذج تعلم الآلة واكتشاف الحالات الشاذة.

3- خصـوصـية البيانات: يسـتدعي هذا الاهتمام الكبير في العصـر الرقمي اتخاذ تدابير صـارمة لحماية المعلومات الشـخصـية الحسـاسـة من التهديدات السـيبرانية والانتهاكات وفقدان البيانات، من خلال ممارسـات حاسـمة مثل الوعي الشـامل بالبيانات والإدارة الفعالة لمسـتودع البيانات والنسـخ الاحتياطية وأمن الشـبكة ضـد الدخول غير المصـرح به وتقويمات المخاطر المنتظمة وتدريب المستخدم المستمر على سرية البيانات وأمنها.

4- إدارة البيانات: يمكن أن يكون للانتهاك الأمني تداعيات خطيرة، بما في ذلك الكشـف عن معلومات الأعمال المهمة داخل قاعدة بيانات مخترقة. توفر أنظمة إدارة البيانات القوية إجراءات أمنية واسـعة النطاق، بما في ذلك تشفير البيانات وتقسيمها ونقل البيانات بشكل آمن وتنفيذ خادم موثوق به.

5- التحكم في الوصـول إلى البيانات: يعد التحكم الفعال في الوصـول إلى البيانات، خاصـة في المؤسـسـات الكبيرة التي تضم العديد من الموظفين، أمراً صعباً، ولكنه ضرورياً للحفاظ على سلامة البيانات والخصوصية.

6- تسـمم البيانات: تتحسـن حلول تعلم الآلة، مثل برامج الدردشـة الآلية، بشـكل مسـتمر من خلال التفاعل مع مجموعات البيانات الضـخمة، ولكن يمكن استغلال هذا التقدم من خلال هجمات تسميم البيانات. يمكن أن يؤدي هذا التلاعب ببيانات التدريب إلى الإضـرار بقدرة النموذج على عمل تنبؤات دقيقة، ما يؤدي إلى تلف المنطق ومعالجة البيانات وحقن البيانات.

7- سرقة الموظف: إن إضـفاء الطابع الديمقراطي على الوصـول إلى البيانات يعني أن كل موظف لديه مستوى من المعلومات التجارية المهمة، ما يزيد من خطر تسرب البيانات غير المقصود أو المتعمد. تعد سرقة الموظفين مصدر قلق عبر الشـركات، بدءاً من الشـركات الناشـئة وحتى عمالقة التكنولوجيا. لمواجهة هذا التهديد، يجب على الشـركات تنفيذ سياسات قانونية وتأمين الشبكات باستخدام شبكات خاصة افتراضية. بالإضافة إلى ذلك، يمكن لسطح المكتب كخدمة (Desktop as a Service: DaaS) تقييد الوصـول إلى البيانات من محركات الأقراص المحلية وتعزيز الأمان [4,5,6].

الخلاصة:

مما تم ذكره وتحليله ضمن هذه المقالة، نستنتج أن البيانات الضخمة والأمن السيبراني يعتبران من أهم التحديات التي تتطلب تكامل الحلول التقنية وتطوير استراتيجيات قوية من أجل الوصول إلى الغاية السليمة من البيانات. مع تزايد حجم البيانات العملاقة وأهميتها، يصبح تأمينها وإدارتها بشكل فعال أمراً لا بد منه لضمان استدامة الأعمال وحماية الأصول الرقمية.

المراجع:

1- Vailaya, "What's All the Buzz Around "Big Data?"“, IEEE Women in Engineering Magazine, December 2012, pp. 24-31.

2- Brown, M. Chui and J. Manyika, "Are you Ready for the era of 'Big Data'? "McKinsey Quarterly, McKinsey Global Institute, October 2011.

3- Tankard, "Big Data Security”, Network Security Newsletter, Elsevier, ISSN 1353-4858, July 2012.

4- Intel IT Center, "Peer Research: Big Data Analytics “, Intel's IT Manager Survey on How Organizations Are Using Big Data, August 2012.

5- S. Singh and N. Singh, "Big Data Analytics”, 2012 International Conference on Communication, Information & Computing Technology IEEE, October 2011.

6- Bekmyrza Dzhekishev "Big Data Security Best Practices" created at Sep 27, 2023 [online] Available at: https://maddevs.io/blog/big-data-security-best-practices/.

7- Emma Crockett "What is Big Data Security? Challenges & Solutions" created at May 1, 2023 [online] Available at: https://www.datamation.com/big-data/big-data-security/.

8- Sitalakshmi Venkatraman1, and Ramanathan Venkatraman "Big data security challenges and strateg" 25 April 2019.

9- R Rawat and R Yadav "Big Data: Big Data Analysis, Issues and Challenges and Technologies" IOP Conf. Series: Materials Science and Engineering 1022 (2021) 012014.

10- Diana Martinez-Mosquera, Rosa Navarrete and Sergio Lujan-Mora " Modeling and Management Big Data in Databases" Sustainability 15 January 2020.