УДК 004.056.5:6

Луцков М. П., студент 2 курсу
спеціальності 122 «Комп'ютерні науки»
Новицький М. О., студент 2 курсу
спеціальності 122 «Комп'ютерні науки»
Римар П. В., старший викладач
кафедри інформаційних технологій

ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ ДАНИХ У BIG DATA

Донецький національний університет імені Василя Стуса, м. Вінниця

З кожним роком сфера Big Data дедалі більше проникає в наше життя. На сьогоднішній день обсяги великих даних почали вимірюватись у ексабайтах. Але у багатьох людей, які займаються сферою захисту даних, постало запитання як забезпечити повний захист цієї інформації, щоб вони не потрапили до не бажаних рук, адже щороку ринок Big Data збільшується на 16% [1]. Оскільки разом з ринком значно збільшуються і обсяги даних, проблема їх безпеки виходить на перший план. Щоб з'ясувати як на сьогоднішній день захищаються Big Data, ми провели дослідження, як організовували безпеку цих даних різні компанії світу.

Подібні дослідження в області безпеки великих даних проводились багатьма іноземними вченими [2][3].

Термін «великі дані» (Big Data широко застосовується з кінця 2000-х. Популярність цієї теми все ще досить висока. Термін Big Data асоціюється з структурованими та неструктурованими даними великого обсягу та великою різноманітністю форматів. При цьому їх формування, накопичення та обробка, як правило, здійснюються в режимі он-лайн, завдяки роботі розподілених систем, використанню хмарних сервісів.

Прикладами того, що може бути джерелом даних про великі обсяги, є: GPS-сигнали для транспортної компанії; – дані датчиків промислового підприємства; – цифрові книги в електронній бібліотеці; – Банківські операції; – інформація про товари чи покупки з великої роздрібної мережі тощо. Основне завдання методів Big Data – обробляти величезну кількість даних та будувати прогнозні моделі, виявляти приховані зв'язки та взаємодіяти з ними. Це актуальне питання, яке вирішують, наприклад, такі компанії, як Informatica (розробка сховищ ETL та технології управління даними); Hewlett Packard Enterprise; Imperva (розробка та виробництво продукції для захисту систем управління базами даних та веб-додатків); Apache Software Foundation (Apache – технології розробки програмного забезпечення з відкритим кодом).

Великі проблеми безпеки даних виникають на всіх етапах обробки даних – у процесі генерації, передачі, зберігання, зберігання, аналізу та візуалізації. Відомі компанії на ринку IT приділяють значну увагу розробці спеціалізованих рішень щодо захисту даних Big Data, найвідоміші з яких перераховані в таблиці

1. Найбільший внесок у розвиток технологій безпеки даних Big Data робить International Cloud Security Альянс (CSA), Національний інститут стандартів і технологій США (NIST) та Агентство Європейського Союзу з безпеки мережі та інформації (ENISA).

Таблиця 1 – Рішення з захисту BD

Розробник	Рішення
IBM	Top tips for Big Data Security
Oracle	Enterprise Security for Big Data Environments
Forrester	Big Data Security Strategies For Hadoop Enterprise Data Lakes
Cloudera	Cloudera Security
Securosis	Securing Hadoop: Security Recommendations for Hadoop Environment
CSA	The Big Data Security and Privacy Handbook
NIST	Big Data Interoperability Framework

Сучасні технології дозволяють розгорнути власний хмарний сховище та розгорнути його у ваш корпоративний простір, віртуальний хостинг, VPS або віддалений сервіс. Інформація може надходити з офісних ПК, Інтернету, промислових датчиків або з мобільних пристроїв.

Існує безліч інструментів для розгортання хмарного сховища. До таких інструментів належать Seafile, ownCloud, Pydio, BitTorrent Sync, Syncthing, HRCloud2, SparkleShare, Storage Made Easy, AeroFS, TeamDrive, arXshare, LimboMedia, EncBox, помічник git-annex, Tonido, Nextcloud, ownCloud, Cozy.

Окремим питанням щодо безпеки та захисту великих даних є криптографія Big Data. На додаток до відомих і поширених алгоритмів шифрування, таких як AES або RSA, на різних етапах захисту великих даних рекомендується шифрування паролем, шифрування прохідних даних, асоційоване шифрування, шифрування на основі атрибутів (ABE), шифрування на основі ідентичності (IBE), конвергентне шифрування. Увага приділяється безпеці на транспортному шарі (TLS), шифруванню на рівні захищених сокетів (SSL) та наявності надійних механізмів захисту для зберігання, зберігання та обробки хмар даних. Особливості криптографічних підходів до обробки даних Big Data – це роздільна обробка інформації та метаданих, організація пошуку за допомогою булевих запитів зашифрованих даних, порівняння даних без їх розшифровки, виявлення дублікатів даних у великих масивах на основі клавіш [4] і інші. Можна зробити висновок, що дані повинні бути повністю зашифровані на кожному етапі обробки, зберігання або передачі. Але при цьому криптографічні процедури повинні бути швидкодіючими. Такими, що не завадять основній задачі – аналізу даних, а крім того, до них повинен бути організований ефективний та безперервний доступ [5].

Наприкінці наведемо ще деякі програмні й алгоритмічні підходи, що застосовуються для захисту файлів, що зберігаються в хмарних сховищах. Це:

криптографічна файлова система EncFS; пропрієтарні програми, що позиціонуються як засіб для шифрування даних в хмарі, наприклад, Boxcryptor, Cloudfogger Truecrypt і ін.; локальне шифрування (наприклад, утиліта CryptSync); програми для резервного копіювання (наприклад, Duplicati); клієнт CarotDav.

Ми дослідили питання безпеки Big Data. Визначили як різні компанії світу боролись із проблемою безпеки великих даних. Встановили, що для забезпечення захисту необхідний багаторівневий захист та застосування ефективного програмного забезпечення.

Список літератури

1. Аналітичний розбір ринку Big Data [Електронний ресурс]. Режим доступу – https://habr.com/ru/company/moex/blog/256747/

*2. Tian, Y. (2017) Towards the Development of Best Data Security for Big Data. Communications and Network, **9**, 291-301. doi: 10.4236/cn.2017.94020.*

3. Bashari Rad, Babak & Akbarzadeh, Nafisseh & Ataei, Pouya & Khakbiz, Yasaman. (2016). Security and Privacy Challenges in Big Data Era. International Journal of Control Theory and Applications. 9. 437-448.

4. Big Data Security and Privacy Handbook: 100 Best Practices in Big Data Security and Privacy. Cloud Security Alliance. URL: https://downloads.cloudsecurityalliance.org/assets/research/big-data/BigData_Security_and_Privacy_Handbook.pdf

5. Проблеми безпеки великих даних [Електронний ресурс]. Режим доступу – https://www.osp.ru/os/2017/04/13053380/

УДК 004.056

Нескородєва А. Р., студентка 1 курсу спеціальності 113 «Прикладна математика»
Римар П. В., старший викладач кафедри інформаційних технологій

ВИКОРИСТАННЯ BIG DATA У СУЧАСНІЙ МЕДИЦИНІ

Донецький національний університет імені Василя Стуса, м. Вінниця

Сьогодні «великі дані» використовуються у багатьох сферах людського життя. Вони набувають популярності та стають невід’ємною частиною багатьох досліджень. Вікіпедія[1] дає «великим даним» таке визначення – це структуровані та неструктуровані данні великих розмірів та великої різнорідності, а також методи їх обробки, які дозволяють людині аналізувати інформацію.

Одне з найважливіших завдань перед вченими це можливість максимально продовжити життя людини, зробити його безтурботним та самостійним. У цьому питанні «великі дані» будуть досить корисними, адже вони є основним джерелом інформації.