DBD Collaborative Discussion 1

The internet of things (IOT) provides access to a variety of data sources resulting in heterogeneous data forms that can be semi-structured or unstructured (Azad et al., 2019). In 1995 the digital universe was roughly 130 billion gigabytes which was then predicted to reach 40 trillion gigabytes in 2020 (Haider, 2016). This vast amount of data can be seen as a great opportunity or a difficult challenge to overcome. On one hand connectivity provides a vast amount of data to obtain from multiple sources of data. On the other a large proportion of this data will be of no interest which raises the challenge of filtering through the abundance of data to find “the needle in the haystack” (Agrawal et.al 2012).

Many big data technologies available over the IOT are open source providing a cost effective way to gather data, however the quality of this data can vary significantly. Hajjaji et al. (2021) states that the veracity of this data is the most crucial challenge when compared with the other V’s of big data (i.e., variety, volume and velocity).

Analysing data via IOT poses a multitude of threats in relation to privacy/security especially around personal data, leading to IOT users to be prone to privacy infringements (Tariq et al., 2019). To counter security issues multiple security techniques are being employed however the complexity and heterogeneity of big data remains an issue (Tariq et al., 2019).

Bibliography

Agrawal, D.*et al.* (2012). Challenges and Opportunities with Big Data: A white paper prepared for the Computing Community Consortium committee of the Computing Research Association. Available at: http://cra.org/ccc/resources/ccc-led-whitepapers/

Azad, P. *et al.* (2019). “The role of structured and unstructured data managing mechanisms in the internet of things,” *Cluster Computing*, 23(2), pp. 1185–1198. Available at: https://doi.org/10.1007/s10586-019-02986-2.

Hajjaji, Y. *et al.* (2021) “Big Data and IOT-based applications in Smart Environments: A systematic review,” *Computer Science Review*, 39, p. 100318. Available at: https://doi.org/10.1016/j.cosrev.2020.100318.

Haider, M., (2016). *Getting started with data science*. 1st ed. Boston: IBM Press, Pearson, p.4.

Tariq, N. *et al.* (2019) “The security of big data in fog-enabled IOT applications including Blockchain: A survey,” *Sensors*, 19(8), p. 1788. Available at: https://doi.org/10.3390/s19081788.

Peer Response: Tasweem Beelunkhan

Thank you Shaun for this insightful post. Indeed, analysing data via IOT poses a multitude of threats in relation to privacy/security especially around personal data. Privacy and security issues are the main challenges when performing data cleaning on a large volume of data. Since data cleaning processes must be performed using automated data cleaning tools, inconsistent software may be used that will clean the data based on faulty models and thus decrease the data quality and create the possibility of breaches.

Also, another reason is these big data tools are open source and are not fully secure. Thus, it requires more security tools to protect data from DDoS attacks, ransomware, theft and other malicious activities. Thus, companies need to invest in big data security and privacy tools to protect their data and understand the newest threats. Failure to do so could result in severe financial losses and other organizational problems as severe fines and restrictions could be imposed under the GDPR.

References:

The internet of things: Opportunities and challenges: Think tank: European parliament Think Tank | European Parliament. Available at: https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2015)557012 (Accessed: 20 May 2023).

Sanie, M. (2023) Big Data Security: Biggest challenges and best practices, Cprime. Available at: https://www.cprime.com/resources/blog/big-data-security-biggest-challenges-and-best-practices/ (Accessed: 20 May 2023).

Maayan, G.D. (2023) Big Data Security: Challenges and solutions, DATAVERSITY. Available at: https://www.dataversity.net/big-data-security-challenges-and-solutions/# (Accessed: 20 May 2023).

Summary Post

After receiving responses from my peers during this collaborative discussion I have improved my understanding of the risk and opportunities associated with the IOT.

Tasweem Beelunkhan's response correctly mentioned that a major risk in the IOT is data breaches which can result in major financial losses/restrictions that can occur as a result of the General Data Protection Regulation (GDPR). In 2019, tech giants Google were fined €50 million for a violation of GDPR obligations. The fine was issued after Google failed to provide customers with enough information about the data consent policies they have in place to be able to gain consent from the users (European Commission, 2019).

In addition to compliance with the GDPR in relation to consent, companies also have to ensure that data is secure. Analysing data via IOT poses a multitude of threats in relation to privacy/security especially around personal data, leading to IOT users to be prone to privacy infringements (Tariq et al., 2019). To counter security issues multiple security techniques are being employed however the complexity and heterogeneity of big data remains an issue (Tariq et al., 2019).

The veracity of data remains the most crucial challenge in comparison to the other "V's" of big data (i.e., variety, volume and velocity). Many big data technologies available over the IOT are open source providing a cost effective way to gather data, however the quality of this data can vary significantly. In the context of IOT data veracity is connected to the usability and quality of the data (Assiri, 2020). Assessing the veracity of data on IOT is challenging due. IOT provides access to a variety of data sources resulting in heterogeneous data forms that can be semi-structured or unstructured (Azad et al., 2019).

Reference:

Assiri, F. (2020) 'Methods for assessing, predicting, and improving data veracity: A survey', *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, 9(4), pp. 5–30. doi:10.14201/adcaij202094530.

Azad, P. *et al.* (2019). "The role of structured and unstructured data managing mechanisms in the internet of things," *Cluster Computing*, 23(2), pp. 1185–1198. Available at: https://doi.org/10.1007/s10586-019-02986-2.

European Commission (2019) *Communication from the Commission to the European Parliament and the Council. Data protection rules as a trust-enabler in the EU and beyond – taking stock* , *eur-lex.europa.eu*. Available at: https://eur-lex.europa.eu/EN/legal-content/summary/general-data-protection-regulation-gdpr.html (Accessed: October 27, 2022).

Hajjaji, Y. *et al.* (2021) “Big Data and IOT-based applications in Smart Environments: A systematic review,” *Computer Science Review*, 39, p. 100318. Available at: https://doi.org/10.1016/j.cosrev.2020.100318.

Haider, M., (2016). *Getting started with data science*. 1st ed. Boston: IBM Press, Pearson, p.4.

Tariq, N. *et al.* (2019) “The security of big data in fog-enabled IOT applications including Blockchain: A survey,” *Sensors*, 19(8), p. 1788. Available at: https://doi.org/10.3390/s19081788.