Machine Learning Model to Predict the Severity and Provide an Early Warning of Nerve Agents Threats Using Internet of Things Technologies

Syed Nasir Alsagoff
Faculty of Defence Science and Technology,
Universiti Pertahanan Nasional Malaysia
syednasir@upnm.edu.my

Norshahriah Abdul Wahab
Faculty of Defence and Technology,
Universiti Pertahanan Nasional Malaysia
shahriah@upnm.edu.my

Syahaneim Marzuki
Faculty of Defence Science and Technology,
Universiti Pertahanan Nasional Malaysia
syahaneim@upnm.edu.my

Nur Hazimah Nordin
Faculty of Defence Science and Technology,
Universiti Pertahanan Nasional Malaysia
nurhazimah@upnm.edu.my

'Atifah Hanim Rosli
Faculty of Defence Science and Technology,
Universiti Pertahanan Nasional Malaysia
atifahhanim@gmail.com

***Abstract*— Nerve Agents (NAs) are often clear, colourless, and tasteless. The only way to identify if a person has been exposed to NAs is through physical symptoms. Longer exposure can lead to death. Rapid identification and alert of NAs involved in any hazardous material incident or terrorist attacks are vital to the protection of first responders as well as for the effective treatment of victims. In this research, high sensitivity NAs sensors that are capable of detecting various types of dangerous chemical agents along with the response measure to the threat will be developed. To ensure a swift and proper response from the authority, this project will focus on Big Data Analytics and IoT streaming data analytics combined with a machine learning model to predict the severity and provide an early warning of the NAs threats using the Internet of Things (IoT) technologies. Built-in IoT sensors in Android devices such as a barometer and other telemetry sensor data will be merged with the NAs sensor data to be sent to a cloud-based database. To ensure rapid, no false positive, and false-negative alerts, Big Data Analytics and IoT Streaming Data Analytics combined with machine learning will be applied to get the most accurate results. This is especially important in any hazardous material incident or terrorist attack so that the most optimal response for the situation can be executed. The complete outcome of this project is in line with Malaysia's National Defence Policy (MNDP) and will ensure the safety and security of the nation as a whole.**

Keywords—Nerve Agents, National Security, IoT Sensor, Detection System

I. Introduction

In the era of the Internet of Things (IoT) combined with high speed 4G (and above) network, many sensing devices are able to collect various sensory data over time for a wide range of fields and applications. Based on the nature of the application, these devices will result in real-time data streams. By applying Big Data Analytics and IoT Streaming Data Analytics combined with Machine Learning Model, such data streams may be used to discover new information and predict future insight in order to provide control decisions which are the crucial process that makes IoT a worthy paradigm for quality-of-life improving technology [1]–[3].

In this research, we will provide the methodology of combining IoT sensors, NAs sensors along with Big Data Analytics and IoT Streaming Data Analytics combined with machine learning structures for processing data captured by IoT sensors that include parameters such as time and air pressure to be processed in a cloud database. By using an advanced machine learning model, data captured can facilitate the analytics and learning data in the IoT domain. The articulation of IoT data characteristics from a machine learning model perspective will give the advantage on ways to process data in IoT Big Data Analytics and IoT Streaming Data Analytics [4], [5]. The crux of this research is an approach to achieve the desired analytics, which is based on datasets and parameters to identify the best Course of Actions (COAs) in incidents involving NAs for relevant authorities such as the Malaysian Fire and Rescue Department, Ministry of Science, Technology, and Innovation

(MOSTI), Ministry of Health (MOH) and National Disaster Management Agency (NADMA).

II. Literature Review

Monitoring NAs threats is difficult. It is made more complicated when the incident occurs at a remote or inaccessible location. For better use of NAs detection data, there must be a method to alert the authorities before the problem becomes severe. Currently, toxic gas incidents such as the Kim Kim River pollution in Malaysia are handled manually and time-consuming in nature [6]. The response team must be equipped and trained with safety measures when extracting samples and accessing the contaminated location. Monitoring using 4G and IoT sensors is one approach to prevent further causalities, save time, and allow faster action response [7]. It is proposed that the Big Data Analytics and IoT Streaming Data Analytics combined with a machine learning model be utilized to produce early warnings of any impending problem. The system can also solve any false positive or negative warnings that often happen using current detection methods.

Therefore, an integrated system is proposed in this research that consists of i) a system that can process in real time high-speed transfer of NAs data using data collection methods developed by other researchers from other departments of the National Defence University Malaysia (NUDM), ii) a predictive analysis system based on Big Data Analytics and IoT Streaming Data Analytics combined with a machine learning model, and iii) an early detection and warning system for the use of authorities such as the Malaysian Fire and Rescue Department, Ministry of Science, Technology, and Innovation (MOSTI), Ministry of Health (MOH) and National Disaster Management Agency (NADMA).

A. IoT Data Characteristic and Requirements for Analytics

IoT data can be streamed continuously or accumulated as a source of big data. Streaming data refers to the data generated or captured within tiny intervals of time and needs to be promptly analysed to extract immediate insights and to make fast decisions. Big data refers to huge datasets that are commonly used by hardware and software platforms that are not able to store, manage, process, and analyse the data. These two approaches should be treated differently since their requirements for analytics response are not the same. Insight from big data analytics can be delivered after several days of data generation, but insight from streaming data analytics should be ready in a range of a few hundredths of milliseconds to few seconds [8]–[10].

Data fusion and data sharing play a critical role in developing ubiquitous environments based on IoT data. This role is more critical for time-sensitive IoT applications where a timely fusion of data is needed to bring all pieces of data together for analysis and consequently providing reliable and accurate actionable insights [11].

B. Big Data Analytics and IoT Streaming Data Analytics

Beyond the aforementioned properties, researchers [12]–[15], have identified other characteristics. Firstly, big data can be a by-product or footprint of a digital activity or IoT interplay. In addition, big data systems should be horizontally scalable, that is, big data sources should be able to be expanded to multiple datasets. These attributes also lead to the complexity attribute of big data, which in turn imposes other challenges like transferring and cleansing data. Finally, IoT Streaming Data Analytics involve real-time data fetch to analyse the pattern and behaviour of certain parameters.

C. Machine Learning Model and IoT

In the era of Machine Learning and IoT, sensing devices collect and generate various sensory data over time for a wide range of fields and applications. From the collection of data, manipulation from these data is important to transform the data in a meaningful way. Based on the nature of the application, these devices will result in big or real-time data streams. Applying analytics over such data streams will enable faster and more correct COAs for life-threatening NAs incidents.

Using a class of advanced machine learning techniques to facilitate analytics and learning in the IoT domain, articulating IoT data characteristics and identifying two major treatments for IoT data from a machine learning perspective, namely IoT Big Data Analytics and IoT streaming data analytics, may give full potential of using emerging Ma-chine Learning techniques for IoT data analytics [16] – [19]. It is crucial and promising for COAs and the decision-making process in predicting NAs threats in Malaysia using the built-in IoT sensors in Android devices as data collection tools for the early detection of NAs.

III. Methodology

NAs detection data combined with time and air pressure readings of the sensors of the Android IoT sensors will be transferred to a central server running MySQL database. For the prediction to work, base data needs to be collected for comparison for six months. Correlating the various data along with the weather pattern and air pressure of the area, a more accurate prediction of the impending disaster can be made. The description of methodology will be in four main parts:

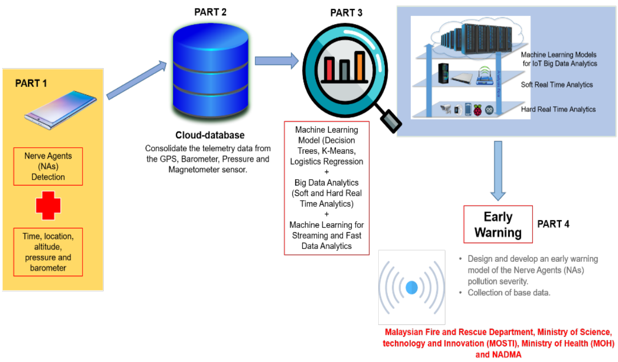

Fig. 1. Main parts of the research project.

Part 1:

Research the various Big Data analytics and IoT Streaming Data Analytics combined with machine learning model to effectively predict the severity, source, and rate of spread of NAs threat. Next, work will be carried out for the integration of sensors in an Android device such as a barometer for high-speed telemetry data transfer. Then, the connectivity of the NAs sensors to the Android device will be developed.

Part 2:

Design a cloud-based database to consolidate the data from the An-droid IoT devices and NAs sensor. The development of an application on the Android device is to: i) Consolidate the telemetry data from the barometer, ii) Receive NAs sensor data, and iii) Combine both sets of data and send them to a cloud-based database.

Part 3:

Design and develop an amalgamation of Big Data Analytics and IoT Streaming Data Analytics combined with machine learning model to effectively predict the severity, source and rate of spread of NAs threats, data filtering and cleansing, develop a machine learning for streaming and fast data analytics, and develop soft and hard real-time analytics.

Part 4:

Design and develop a threat level analysis and early warning model of the NAs threats severity, collection of base data, testing of the prototype and completed system, and reduce false positive and false negative outputs data.

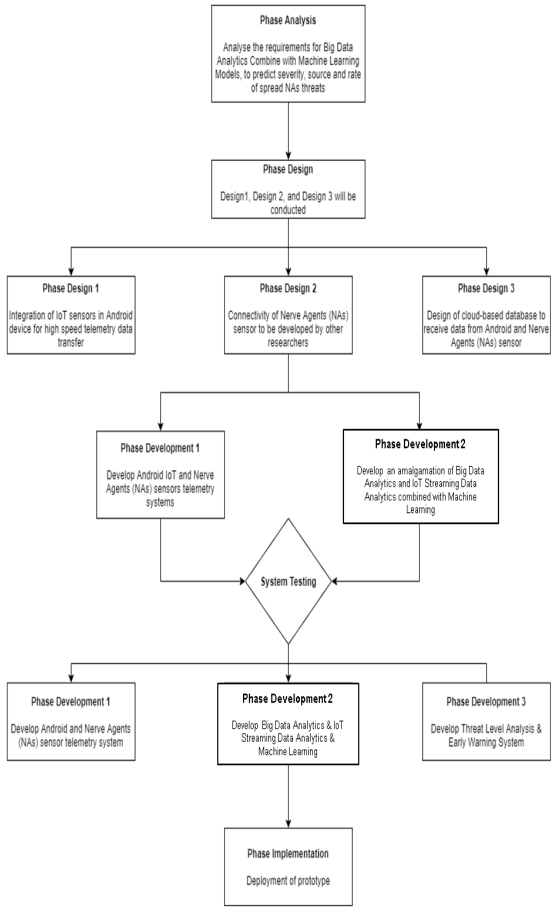

Fig. 2. System Development Diagram.

A. Research Activities

HCN detection data combined with time, location, altitude, pressure and barometer readings of the sensors of the Android device will be transferred to a central server running MySQL database. For a prediction to work, base data needs to be collected for comparison for six months. Correlating the various data along with the weather pattern and temperature of the surrounding area, a more accurate prediction of an impending disaster can be made.

IV. Discussion

In this research, Big Data Analytics and IoT Streaming Data Analytics are combined with a machine learning model to predict the severity and early warning of NAs threat. The product of this research can be used by relevant authorities such as the Malaysian Fire and Rescue Department, Ministry of Science, Technology, and Innovation (MOSTI), Ministry of Health (MOH) and National Disaster Administration Agency (NADMA).

Kim Kim River in Pasir Gudang, Johor, is one of NA's pollution examples. In March 2019, an incident occurred owing to the unlawful disposal of 2.43 tons of chemical waste, which released benzene, acrolein, acrylonitrile, hydrogen chloride, methane, toluene, xylene, ethylbenzene, and d-limonene [20]. The pollution of the Kim Kim River has spread 5 kilometers away from the dumping site by air vapor.

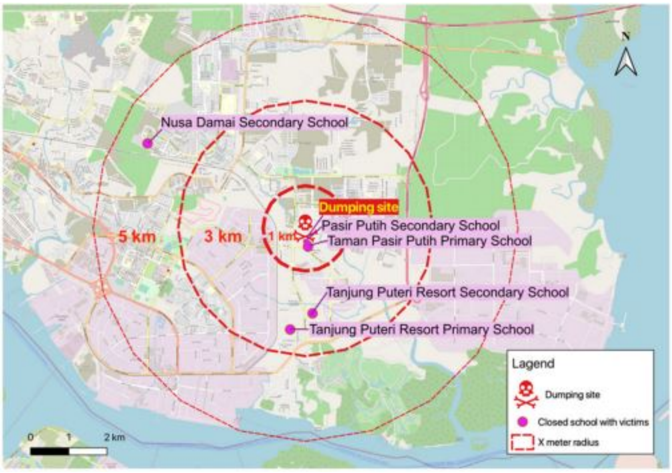

Fig. 3. Radius of affected area [20].

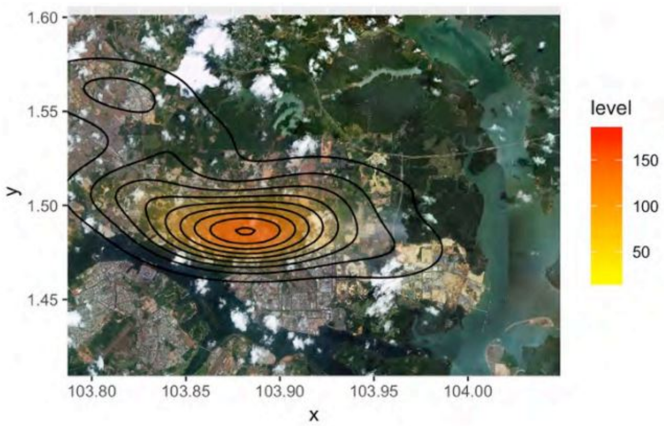

Fig. 4. Level of severity of affected area [21].

The hazardous wastes have been identified as benzene, acrolein, acrylonitrile, hydrogen chloride, methane, toluene, xylene, ethylbenzene, and d-limonene. Exposure to any of these chemicals may cause harmful health effects. The details of hazard and dose–response assessment is summarized in Table 1.

Table 1. List of chemicals found at the dumping site and the health effects [20,21].

Chemical	Exposure Route	Symptoms	Lowest Level for Nonlethal Irreversible at 10 to 30 min Exposure Level	Cancer Risk (IARC)	Inhalation Unit Risk for Cancer
Benzene	Inhalation, skin absorption, ingestion, skin, and/or eye contact	Irritation eyes, skin, and nose; respiratory difficulty; dizziness; headache; nausea; staggered gait; anorexia; lassitude (weakness and exhaustion); dermatitis; bone marrow depression	1100–2000 ppm	Carcinogenic to humans. May cause leukemia	2.2 × 10−6 per µg/m3
Acrylonitrile	Inhalation, skin absorption, ingestion, skin and/or eye contact	Irritation eyes, skin; asphyxia; headache; sneezing; nausea; vomiting; lassitude (weakness, exhaustion); dizziness; skin vesiculation; scaling dermatitis	3.2–8.6 ppm	Possibly carcinogenic to humans. Increase risk of brain tumor, lung and bowel cancer	6.8 × 10−5 per µg/m3
Acrolein	Inhalation, ingestion, skin and/or eye contact	Irritation eyes, skin, and mucous membrane; decreased pulmonary function; delayed pulmonary edema; chronic respiratory disease	0.18–0.44 ppm	The agent is not classifiable as to its carcinogenicity to humans	-
Hydrogen Chloride	Inhalation, ingestion, skin and/or eye contact	Irritation nose, throat, and eye; cough; choking; dermatitis; skin burns; contact with	43–100 ppm	The agent is not classifiable as to its carcinogenicity to humans	-

		refrigerated liquid may cause frostbite			
Methane	Inhalation, ingestion, skin and/or eye contact	Breathing difficulties (i.e., suffocation and increased breathing rate); nausea and vomiting; loss of consciousness; weakness; headaches and dizziness; loss of coordination	NA	The agent is probably not carcinogenic to humans	-
Toluene	Inhalation, ingestion, skin and/or eye contact	Headaches; dizziness; loss of consciousness; loss of coordination; sleepiness	760–1400 ppm	Not classifiable as to carcinogenicity to humans	-
Xylene	Inhalation, ingestion, skin and/or eye contact	Irritation of eyes and throat; headaches; dizziness; sleepiness; trembling; lack of coordination	1300–2500 ppm	Not classifiable as to carcinogenicity to humans	-
Ethylbenzene	Inhalation, ingestion, skin and/or eye contact	Irritation of the eyes and/or throat; chest constriction; dizziness	1600–2900 ppm P	Possibly carcinogenic to humans	NA
d-limonene	Inhalation, ingestion, skin and/or eye contact	Breathing difficulties; skin irritation	NA	Not classifiable as to carcinogenicity to humans	-

NA=Not Available

This research hopes to prevent incidents such as the Kim Kim River toxic pollution that occurred on 7 March 2019 caused by illegal chemical waste dumping at the river in Pasir Gudang, Johor, Malaysia. Fortunately, it was a non-fatal incident. About 6,000 people were affected. A total of 2,775 victims received immediate treatment while eight were treated in the Intensive Care Unit (ICU). The authorities were criticized as being too slow in reacting / responding to the incident. If such incidents are to occur in the future, this system will give an early warning to the relevant authorities so that they can promptly and effectively take the necessary action to minimize any adverse effects.

V. Conclusion

In this research, NAs sensors with high sensitivity and a threat response measure will be developed. It is important for NAs to be identified and the authorities to be alerted quickly in any hazardous material incidents or terrorist strikes to protect first responders and treat victims effectively. The study includes data processing, data manipulating and data cleansing from the NAs sensor and IoT to a Big Data Analytics system for prediction of any possibility of NAs threats in indoor cases. It will also include a predictive analysis system that imposes a machine learning model or algorithms to predict how to measure the severity of NAs threats. More work will be done by conducting more experiments under more complicated conditions for sensor detection and system improvement.

Acknowledgment

The authors gratefully acknowledge the Computer Science Department of the National Defence University of Malaysia (NDUM) and Research Centre Management of NDUM for their support for this study under the UPNM Short Term Research Grant under code (UPNM/2020/GPJP/ICT/7).

References

[1] M. Strohbach, H. Ziekow, V. Gazis, and N. Akiva, “Towards a big data analytics framework for IoT and smart city applications,” in Modeling and processing for next-generation big-data technologies, Springer, 2015, pp. 257–282.Author, F., Author, S.: Title of a proceedings paper. In: Editor, F., Editor, S. (eds.) CONFERENCE 2016, LNCS, vol. 9999, pp. 1–13. Springer, Heidelberg (2016).

[2] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, “Internet of things: A survey on enabling technologies, protocols, and applications,” IEEE communications surveys & tutorials, vol. 17, no. 4, pp. 2347–2376, 2015.

[3] E. Brynjolfsson and T. Mitchell, “What can machine learning do? Workforce implications,” Science, vol. 358, no. 6370, pp. 1530–1534, 2017.

[4] M. Chen, S. Mao, Y. Zhang, and V. C. M. Leung, Big data: related technologies, challenges and future prospects, vol. 96. Springer, 2014.

[5] M. Mohammadi and A. Al-Fuqaha, “Enabling cognitive smart cities using big data and machine learning: Approaches and challenges,” IEEE Communications Magazine, vol. 56, no. 2, pp. 94–101, 2018.

[6] Q.-L. Zeng et al., “Effect of convalescent plasma therapy on viral shedding and survival in patients with coronavirus disease 2019,” The Journal of infectious diseases, vol. 222, no. 1, pp. 38–43, 2020.

[7] J. B. A. Gomes, J. J. P. C. Rodrigues, R. A. L. Rabêlo, S. Tanwar, J. Al-Muhtadi, and S. Kozlov, “A novel Internet of things-based plug-and-play multigas sensor for environmental monitoring,” Transactions on Emerging Telecommunications Technologies, p. e3967, 2020.

[8] L. Bottou, “Large-scale machine learning with stochastic gradient descent,” in Proceedings of COMPSTAT’2010, Springer, 2010, pp. 177–186.

[9] H. Valpola, “From neural PCA to deep unsupervised learning,” in Advances in independent component analysis and learning machines, Elsevier, 2015, pp. 143–171.

[10] H. Zou, H. Jiang, X. Lu, and L. Xie, “An online sequential extreme learning machine approach to WiFi based indoor positioning,” in 2014 IEEE World Forum on Internet of Things (WF-IoT), 2014, pp. 111–116.

[11] A. Arora, A. Candel, J. Lanford, E. LeDell, and V. Parmar, “Deep learning with h2o,” H2O. ai, Mountain View, vol. 587, 2015.

[12] S. bin Qaisar and M. Usman, “Fog networking for machine health prognosis: A deep learning perspective,” in International Conference on Computational Science and Its Applications, 2017, pp. 212–219.

[13] P. V. Klaine, M. A. Imran, O. Onireti, and R. D. Souza, “A survey of machine learning techniques applied to self-organizing cellular networks,” IEEE Communications Surveys & Tutorials, vol. 19, no. 4, pp. 2392–2431, 2017.

[14] R. E. Sutanto, L. Pribadi, and S. Lee, “3D integral imaging based augmented reality with deep learning implemented by faster R-CNN,” in International Conference on Mobile and Wireless Technology, 2017, pp. 241–247.

[15] L. Atzori, A. Iera, and G. Morabito, “The internet of things: A survey,” Computer networks, vol. 54, no. 15, pp. 2787–2805, 2010.

[16] V. Gazis, M. Strohbach, N. Akiva, and M. Walther, “A unified view on data path aspects for sensing applications at a smart city scale,” in 2013 27th International Conference on Advanced Information Networking and Applications Workshops, 2013, pp. 1283–1288.

[17] A. Bifet and E. Frank, “Sentiment knowledge discovery in twitter streaming data,” in International conference on discovery science, 2010, pp. 1–15.

[18] Syahaneim Marzukhi, Nursyafeera Azeera Rahim, Syed Nasir Alsagoff Syed Zakaria. 2018. A Model for Mission Critical Multiple Clients And Single Server Operation For Failure Detection and Recovery System. The 2018 International Conference on Information and Communication Technology (ICICTM'18).

[19] Norshahriah Abdul Wahab, Amalina Farhi Ahmad Fazdlah, Mohd Afif, Mohd 'Afizi Mohd Shukran, Syed Nasir Alsagoff. 2019. Mo deling Critical Successfulness Factors of Mobile Game Applications for Military Trainings. International Scholarly Symposium. ISBN 978-967-16241-2-8, WOS

[20] Ibrahim, M. F., Hod, R., Sahani, M., Mohammed Nawi, A., Idris, I. B., Mohd Yusoff, H., & Toha, H. R. (2021). The Impacts of Illegal Toxic Waste Dumping on Children’s Health: A Review and Case Study from Pasir Gudang, Malaysia. International journal of environmental research and public health, 18(5), 2221.

[21] ASM REPORT LESSONS FROM SUNGAI KIM KIM, PASIR GUDANG. (n.d.).