

Intrusion detection systems for internet of thing based big data: a review

Imane Laassar, Moulay Youssef Hadi

Department of computer science, Faculty of science, Ibn Tofail University, Morocco

Article Info

Article history:

Received Apr 8, 2021

Revised Jul 27, 2021

Accepted Sep 29, 2021

Keywords:

Big data

Cloud computing

Internet of thing

Intrusion detection systems

Machine learning technique

ABSTRACT

Network security is one of the foremost anxieties of the modern time. Over the previous years, numerous studies have been accompanied on the intrusion detection system. However, network security is one of the foremost apprehensions of the modern era this is due to the speedy development and substantial usage of altered technologies over the past period. The vulnerabilities of these technologies security have become a main dispute intrusion detection system is used to classify unapproved access and unusual attacks over the secured networks. For the implementation of intrusion detection system different approaches are used machine learning technique is one of them. In order to comprehend the present station of application of machine learning techniques for solving the intrusion discovery anomalies in internet of thing (IoT) based big data this review paper conducted. Total 55 papers are summarized from 2010 and 2021 which were centering on the manner of the single, hybrid and collaborative classifier design. This review paper also includes some of the basic information like IoT, big data, and machine learning approaches are discussed.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Imane Laassar

Department of computer science, Faculty of science, Ibn Tofail University

Kénitra, Morocco

Email: Imane.laassar@gmail.com

1. INTRODUCTION

Over the last years there has remained an appropriate growth in the usage of communication technologies and data handling method such as internet of thing, cloud computing and big data. In addition to the natural increase of the requirement of people, governments, and business on computer systems are endorsing new threats and leveraging the impacts and probabilities of information security breaches in this new and complex context. With the growth of internet of thing (IoT) source of big data also generate which make the data security risks increase exponentially [1], [2]. It is compulsory to recognize all the susceptibilities and threats that could occur that are premeditated obviously for IoT data assembly. To decrease likely threats, it is ostensible that the need for more studies that focus on the knowledge of threats becomes a fact for that context and that encounters in their security, such as privacy and privacy, have been documented and must be lectured and avoided. IoT and big data have two main relationships on one hand IoT is of the main producer or source of big data and therefore it is an imperative goal for big data analytics to improve the service of big data [3], [4]. IoT data are dissimilar from then overall big data because they have some different appearances usual then other data like large-scale running data, heterogeneity, time and space correlation, and high noise data. One of the main issue in big data is the security of data this is due to the volume of the data increase after the applying the encryption technique therefore different researcher apply different encryption algorithm and trying to reduce the volume of the data size [5]. The data of real

time processing time increasing day by day which make big issue for data security approaches. Data query processing is also one of the big issues in big data encrypted this is due to the both unstructured and structured encrypted data need decryption of the data first [6], [7]. Due to huge quantities of data this can take momentous volumes of time and inquiry dispensation can take substantial time. Intrusion detection systems are purposely situated on a network to distinguish threats and display packets [8]. The intrusion detection system (IDS) realizes this by congregation data from dissimilar systems and network foundations and investigative the data for possible threats [9], [10]. In this article, we have presented different machine learning (ML) algorithms used for intrusion detection system for IoT based big data security from 2010 to 2021. We have also discussed about IDS system, IoT and big data. Finally, we performed a statistical analysis of the review and selected ML to resolve the causes of various issues for IoT based big data security. For the relief of readers, we delivered a list of the supreme regularly used abbreviations in the paper are discussion in Table 1.

Table 1. Abbreviations list

Acronyms	Meaning	Acronyms	Meaning
IoT	Internet of thing	CC	Cloud computing
IDS	Intrusion detection systems	NN	Neural networks
ML	Machine learning algorithm	ML	Machine learning technique
SL	Supervised learning	SVM	Support vector machine
CCS	Cloud computing security	SC	Self-configuration

The main offerings of this review paper are shortened are as: i) present details information about IoT base big data, ii) extant facts information about machine learning technique, iii) contemporary facts information about machine leaning technique on IDS, iv) summarizations of main contribution in intrusion detection systems for IoT based big data using machine leaning method from 2010 to 2021, and v) present future research direction about intrusion detection systems for IoT based big data using machine learning technique.

The statistics connected to our review were mined from about 55 published papers. This group of papers consumes been amassed by accessing numerous peer-reviewed data sources (Table 2). These papers focus the developments complete intrusion detection systems for IoT based big data by the machine learning technique from 2010 to 2021. The frequency of publication of this work per year for the last ten years was intended to imagine the progress of research on this promising thematic of the machine learning technique for IoT based big data using machine learning technique which present in Figure 1.

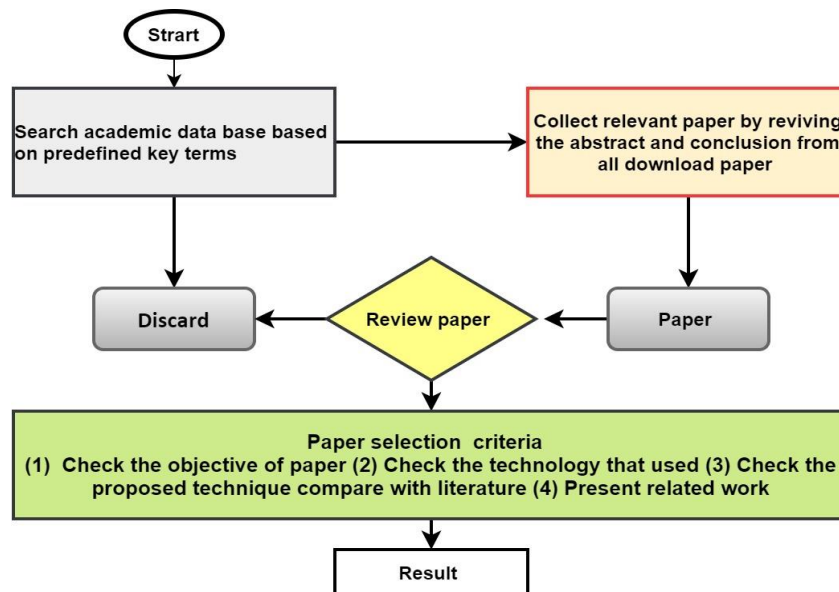


Figure 1. Paper selection procedure

Table 2. Database source and source URL

Google Scholar	https://scholar.google.com/
ACM Digital Library	http://dl.acm.org/
DBLP URL	http://dblp.uni-trier.de/
Springer	www.springer.com
Taylor & Francis	http://taylorandfrancis.com
Wiley Online Library	http://onlinelibrary.wiley.com
IEEE Explore	http://ieeexplore.ieee.org/

Table 2 demonstration the statistics basis by these search engines changed paper are download from 2010 to 2021. The current review presents detailed information about the main contributions in intrusion detection systems for IoT based big data using machine learning technique. The results of recent studies were summarized and discussed. Some recommendation's that could be used as a guide for implementing a future research vision for machine learning techniques for intrusion detection system for IoT based big data are reported, Table 2, and Figure 1 present the paper collection method of this paper.

2. BACKGROUND STUDY

In this section we define all those technologies and parameters which are used for establishment process of intrusion detection systems for IoT based big data using machine learning technique. Different researcher paper are also mention in this section along their advantage and dis advantages. IoT contains of self-formation node they are connected with dynamic and with global network infrastructure. It comprises of small thing with limited storage and processing system internet of thing refers a broad vision. Thing such ways that every day object is place environment are organized with each other with the help of internet. These technology different devices are connecting with other device and used for sharing information or data transfer device to device. The term internet of thing used by Kevin in 1999, it becomes popular due to auto-id entre. In IoT all the device are connected with each other and system architecture should support IoT like as bridge between physical and virtual world. For design process of IoT need to check many factor such as communication, process and commercial models along with security [11].

2.1. Internet of thing

International telecommunication union (ITU) suggested four layer concepts in IoT. These are application layer, network layer, perception layer and middle layer. Application layer consist of several application they offer different service. This is most upper layer and visible for user. No universal standard rule for developing application layer it can be design due to it service. Application layer protocols are distributed at multiple users they can use any information with the help of these protocols [12]. Network layer delivers network broadcast and evidence security and distributes universal entrée atmosphere to the perception layer, that deliver data program and storage consciousness. The network layer comprises mobile strategies, cloud computing, and the internet. Perception layer this layer involves in collection of information and it interconnected network layer. This layer consists of all sensor nodes it means all sensing technology and controlling are data acquired include perception layer are divided in to sub layer [13]. Main element of IoT: internet of thing provides benefit and facilities to the user these services are provide in the form of IoT element which are. Identification process are used for identify each object in the network two main element are used in identification process which are naming and addressing. Naming mentions as tag of the thing and addressing castoff for documentation purpose with precise object both are changed from each other. May be the name of device are same but the addressing not same of object because the method used for addressing with unique code and it assigned with the help of IPv6 [14]. Sensing is the process used for collection of information from different object. Different media are used for storage of information and different sensing device are used for collection of information like actuators, RFID tags, smart sensors, wearable sensing devices. Communication is one of the main elements of IoT in which different device are connected and communicate with each other. In this process different device sent and receive message or different files. Different technologies are used for communication purpose [15].

2.2. Big data concept

Big data conations different source of digital elements like sensor, video, email, numerical modeling, and social resource they data store in these elements are type of text, type of video and graphic. Big data relate the big data or it store big data source but we can get our required data after the analyzing, and visualizing these big data [16]. Big data are generating from different source like online transaction (i-e) email, online management system, and health system, online banking system and networking. Due to large number of data size and information processing they affect the storage and visualization. Last few decades

the data size increasing from different area and types according to the report from international data corporation (IDC) the size of data was (1021TB) in 2012 which increase now 20 times more than it. This was occurred in the data size due to the improve occur in the technologies and its usage [17].

2.3. Big data challenges

Where big data provide lot of opportunities however research and professionals facing several challenges like that, they want to extract big data in to useful knowledge and information. Big data management; data scientist and researcher are facing big issue when they are dealing with big data like big data extraction, storage; integrate when facing less hardware, software as compare data set. Big data management is one of the main issues because collected data from different source and then mange and make them useful for user and organization to use them without any error and duplication. Big data management goal is to ensure reliable data that is easily accessible, manageable, properly stored and secured [18]. Big data cleaning; normally tradition data system consists of cleaning, aggregation, and encoding, storage and access system. When different data are collected from different source however, data sources may contain noises, errors or incomplete data therefore different technique and method are used to cleaning of big data and make them useable for user and organization. Big data aggregation means synchronize outside data in to any organization data and make them useable. Different network are collected to big data it means different data set are generate from outside of IoT technology it need security and make them access able form that user can get the information any time this process is known as big data aggregation [19]. Imbalanced system capacities it is important issue because network architecture is important in the network for access of network. As we know that IoT consist of different network or technology suppose in a same network consist of different device the expectation of each device effect the network if some device are not work properly then it effect the system therefor proper network architecture need and not balance system for good network [20]. Imbalanced big data another challenge is classifying imbalanced dataset its important technique for proper big data system. Normally data set are classified in to two group which are positive and negative now a day they are father divided into sub group. Modern technologies are used to remove the imbalanced data and make them accurate for useable. They make balanced using these types of data class [21].

2.4. Intrusion detection system

An intrusion detection system (IDS) is a network security knowledge originally constructed intended for recognizing susceptibility activities beside a board solicitation or network. IPS extended IDS elucidations through count the capability to block burdens in adding to detecting them besides has convert the leading placement selection for IDS/IPS tools. Intrusion detection structures can be hardware system or software system that are mechanically displays and classify the bout or intrusion and make alert the network or knowledge. This watchful report benefits the superintendent or worker to discovery and fortitude the paleness present in the system or system [22]. Intrusion detection systems are purposely positioned on a network in direction to detect threats and screen network traffic. The IDS take either network or host based approaches for recognizing attacks. The IDS achieve this mission by collecting data from systems and network foundations and achieve investigation on it for conceivable intimidations. Mixture based detection system is the grouping of anomaly-based intrusion detection and signature-based intrusion detection. Maximum of the IDSs use any one of the intrusion detections namely difference or signature. Since both intrusion detections have their own drawbacks, hybrid IDS can be used. Based on their action intrusion detection is confidential into different types. IDSs are envisioned to expose intrusions, before they can disclose the secured system possessions [23]. IDSs remain continuously measured as an additional partition of protection from the sanctuary point of view. IDSs are Infobahn corresponding of the intruder alarms that are being used in corporeal security organization nowadays. Different approaches of intrusion detection usually dissertation, IDS has two main types, and they are: network-based IDS (NIDS) and host-based IDS (HIDS). The methods prospect alienated into two groups which are anomaly intrusion detection and misuse intrusion detection [24].

2.5. Machine learning technique

As everybody knew, machine learning has become more and more interesting, used mainly to train machines to manage data more efficiently. Sometimes after examining the data, we cannot interpret the model or extract the necessary information from the data in this case, we apply machine learning technique. With a large number of data sets available, the demand for machine learning is increasing [25]. Machine learning has been applied in many industries, from the military to medicine, to extract relevant information the goal of machine learning is to learn from the data. Several approaches have been developed and created, by programmers and mathematicians to teach machines to learn themselves. Machine learning is made up of several types of learning that have been classified into some popular families [26].

3. BIG DATA AND MACHINE LEARNING

Big data carry diverse issue for upgrading of these issue new method and methodologies are obligatory. The encounters to big data include performance, data federation, data laxative, safety and time to worth. The supreme shared presentations on the data set spending IoT based big data are seizing data, loading data, data investigation, informing, enquiring meditation technology, and data confidentiality and safety [27]. The leading problems retiring for big data are shapeless and unusual databases, since the out-of-date conducts for big data is not sufficient for loading the data and they requisite detailed behaviors for their desires when they grasp pet bytes or zettabytes. Unlike commencing like Google, Yahoo, Facebook, and other enthusiastic startups do not use Oracle tools for exemption big data sources [28]. In its place the attitude is based on cloud; countless open foundations like Hadoop and dispersed systems. In future, it is indispensable to circumvent the costs upsurge exponentially and storage supplies when new IoT data world may generate. Security of IoT based big data is big issues because the big data associated IoT application are great advantage for society, changed corporations and many other huge and small scale industries. Due to the use of these big data submission sanctuary is imperative development [29]. The foremost contests of IoT grounded big data embrace capture of data, length, and stowage of data along with handover and key administration. The difficulties with big data in trade with IoT submissions are, durable key supervision many officialdoms have functional encryption for data security for IoT big data; they habitually supervise collect indistinctness in key society, admittance control, and tending data admittance. If encryption keys shaped are not effusively dwindling and accomplished, they are disposed to theft by altered hackers [30].

4. RESEARCH METHOD

To study briefly about contribution in intrusion detection systems for IoT based big data using machine leaning technique from 2010 to 2021. Figure 2 show the group of paper where these papers goes to different phase once that has done, we nominated 55 for summarization at final stage. The main roles of summarization of these are that we collect information about contribution in intrusion detection systems for IoT based big data using machine leaning technique from 2010 to 2021. After the implementation of these machine leaning technique the different parameter is used to check the performance.

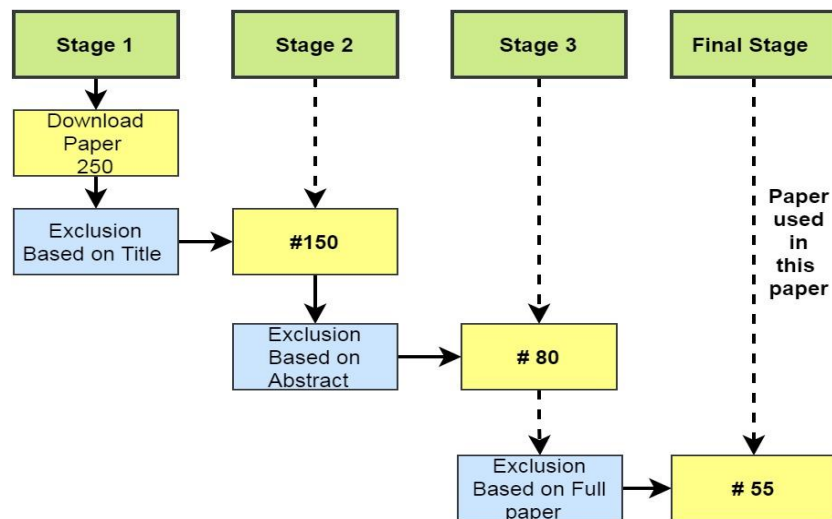


Figure 2. The paper gathering phases for this paper

4.1. Parameters for evaluation

Different parameters are used to check the comprehensively detection effect of different machine learning algorithm for simultaneously in IDS research which are mention in:

- $Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} * 100.$
- Sensitivity is the process which shows the positive fraction or confirms that diagnostic test is positive and the test result for which process has find and it can be written as given below.
- $Sensitivity/rec\ call /true\ positive\ rate = Sensitivity = \frac{TP}{(TP + FN)} * 100$ [31].

- d. Specificity is diagnostic test is negative and person is healthy and can be present as: $\frac{TN}{(TN + FP)} * 100$.
- e. Precision can be: $\frac{TP}{(TP + FN)} * 100$.
- f. True Positive rate = $TP / (TP + FN)$.
- g. False Positive rate = $FP / (FP + TN)$ [32].

5. RESULT AND DISCUSSION

Different technology are involve intrusion detection systems for IoT based big data therefore receiving broadly devotion owing to it energetic wildlife and tractability due to these assets unlike association and researcher are receiving curiosity in this equipment. Different organization and researcher implement by evaluating different data for different experimental purpose therefore used different simulation tools are used for security purpose for big data [33]. In this unit we extant the summary of 28 papers which are used for improvement in about intrusion detection systems for IoT based big data using machine learning technique from 2010 and 2021. The summary of these article contains of method name, process, year, benefit, weakness and references. Table 3 expressions the summary of those articles which report the problems in intrusion detection systems for IoT based big data using machine learning technique taking different parameter with the help of machine learning technique.

Table 3. Summary of related work

Technique	Problems addressed	Improvement	Weakness	Section	Ref
ML Technique	Network intrusion detection	Anomaly detection	Data analytics	Accuracy	[34]
ML Technique	Stream processing	Intrusion prevention systems	Big data in intrusion detection systems	Signature-based detection	[35]
ML Technique	Real-time intrusion detection system	Attack data	Data analytics	Accuracy	[36]
ML Technique	Intrusion detection	Big data processing	Data analytics	Accuracy rate	[37]
ML Technique	Feature selection	Intrusion detection	Big data analytics	Accuracy, detection rate	[38]
Hybrid MLT	Intrusion detection system	Positive detection rate	Classification performance,	Improve detection rate	[39]
Spark-Chi-SVM	Training time	Security	Data analysis	Feature selection	[40]
DNN Technique	Intrusion detection systems	Accuracy	Classification	Evaluate features section	[41]
ASCH-IDS Algorithm	Vulnerabilities	Accuracy and precision recall rates	Data analytics technique	Identification	[42]
Technique	Training section	Accuracy	Data analysis	Data analysis	[43]
Deep Learning	Intrusion detection systems	Accuracy	Training time a	Data analysis	[44]
Learning Algorithm	Auto-update	Packet analysis	Data analysis	Feature selection	[45]
ML algorithm	Context-aware intrusion detection	Detection rate of anomaly signs	Intellectualization based	Analyzing threats	[46]
MLT	Confusion matrices	Traffic dynamically	Data analysis	Feature selection	[47]
ML Techniques	Paramount aspect	Change control identifiers	Big data analytics	Features is selected section	[48]
ML T	Stream processing	Intrusion prevention systems	Big data in intrusion detection systems	Signature-based detection	[49]
MLT	Monitoring anomaly detection	Intrusion detection	Big data analytics	Multivariate big data analysis	[50]
DNN	Intrusion detection	Change control identifiers	Big data analytics	Accuracy, detection rate	[51]
ML technique	Feature selection	Intrusion detection system (IDS)	Big data analytics	Accuracy, detection rate	[52]
ML models	Feature selection	Intrusion Detection	IoT data analytics	Accuracy.	[53]
ML techniques	Cyber-attacks	Intrusion detection system	Big data analytics	Effectiveness	[54]
ML, T	Real-time intrusion detection system	Attack data	Data analytics	Accuracy	[55]
AL, T	Intrusion detection system	Attack data	Data analytics	Effectiveness	[56]
ML, T	Intrusion detection system	Attack data	Dimensional visualization	Accuracy	[57]
ML, Algorithms	Network intrusion detection systems	Ensemble-based	Streaming data	Accuracy	[58]

Table 3 present the result of all summarized paper based on those papers we define different approaches are used intrusion detection systems for IoT based big data and what the advantage and what are the issue still exit is the approached. Once it derives to intrusion detection for IoT based big data using machine approaches it seem that it produce more effective cybercrime security in different area like big data, machine learning technique and network site. From the review paper we become able to discuss about big data issue and machine learning technique facing different issue which are able to solve using intrusion detection system so now we summarized those issues. We fire about that best of the artice are round the detection of denial of service (DoS) attacks. Least facts of the papers are afit the detection of other sorts of attacks.

6. ISSUES AND SUGGESTION

6.1. Deficiency of datasets

There are rare dataset which are certain problems on these datasets thus new more data sets are essential. However, producing new datasets is contingent on expert knowledge, and the labour cost is great. In addition, the eroticism of the internet condition embroiders the dataset deficiency [59].

6.2. Inferior detection accuracy rates

Machine learning attitudes have confident volume to detect intrusions, but they often do not achieve well on lastly uninformed data. Maximum the prevailing revisions were supplemented by labeled datasets. Therefore, when the dataset does not shelter all archetypal real-world samples, good performance in actual settings is not assured-even if the copies achieve high accurateness on test sets [60].

6.3. System environment

The response of IoT during real-world submissions, such as home computerization, industrial automation and city mechanization resulted in a plethora of micro multiplication campaigns and dynamism-operative announcement machineries, provisions, and protocols. IoT structures have been expansively laboring in requests of martial, agriculture, power organizations, education, and commerce [61].

6.4. Contests and future research instructions

Gigantic records of research works have been issued correlated about IDSs for IoT data security. However, there are still a large number of open research challenges and issues, frequently in the use of ML methods for incongruity and imposition detection in IoT for big data sanctuary resolution and these problems are still exits which need to sloved. We can say the datasets not comprise all material or successfully on real data and gratifies all shareholders' necessities [62].

6.5. Statistical significance tests

Form the education of connected work it seems that diverse machine learning method are used for Intrusion detection organizations security of IoT based big data. As we recognize that multiple ML algorithms used over multiple datasets indispensable issues. An algorithm may show better exhibition over one dataset whereas may fail to realize similar result over another dataset this is due to the article circulation or algorithm characteristics [63].

7. CONCLUSION

This analysis paper delivers a summary about IoT based big data for security intrusion detection system captivating machining learning systems are accessible in specifics. Intrusion detection system is a guaranteed investigation field in the cyber security due to the speedy expansion of the different paraded like IoT, cloud calculating and big data. Newly, machine learning algorithms are applied in IDS in order to identify and categorize security threats of IoT grounded data. This paper discovers the qualified study of several ML algorithms used in IDS for numerous solicitations of IoT based bid data and their erection recommend. The consequence of this review will help in empathetic the challenges of big data due to IoT in NIDS. Final section of paper fixed the forthcoming research.

REFERENCES

- [1] S. Umar, S. Baseer, and Arifullah, "Perception of cloud computing in universities of Peshawar, Pakistan," *2016 6th International Conference on Innovative Computing Technology, INTECH 2016*, pp. 87–91, 2017, doi: 10.1109/INTECH.2016.7845046.
- [2] S. Sagioglu and D. Sinanc, "Big data: A review," in *2013 International Conference on Collaboration Technologies and Systems (CTS)*, May 2013, pp. 42–47. doi: 10.1109/CTS.2013.6567202.




- [3] L. Gong, T. Xu, W. Zhang, X. Li, X. Wang, and W. Pan, "Approach research on the techniques for network intrusion detection based on data mining," in *Proceedings of the 2015 International conference on Applied Science and Engineering Innovation*, 2015, vol. 12. doi: 10.2991/asei-15.2015.418.
- [4] T. Hoppe *et al.*, "Corporate semantic web-applications, technology, methodology," *Informatik-Spektrum*, vol. 39, no. 1, pp. 57–63, Feb. 2016, doi: 10.1007/s00287-015-0939-0.
- [5] M. I. Pramanik, R. Y. K. Lau, M. A. K. Azad, M. S. Hossain, M. K. H. Chowdhury, and B. K. Karmaker, "Healthcare informatics and analytics in big data," *Expert Systems with Applications*, vol. 152, Aug. 2020, doi: 10.1016/j.eswa.2020.113388.
- [6] H. N. Dai, R. C. W. Wong, H. Wang, Z. Zheng, and A. V. Vasilakos, "Big data analytics for large-scale wireless networks: Challenges and opportunities," *ACM Computing Surveys*, vol. 52, no. 5, pp. 1–36, Sep. 2019, doi: 10.1145/3337065.
- [7] Y. Wei *et al.*, "A review of data-driven approaches for prediction and classification of building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 1027–1047, Feb. 2018, doi: 10.1016/j.rser.2017.09.108.
- [8] W. Meng, E. W. Tischhauser, Q. Wang, Y. Wang, and J. Han, "When intrusion detection meets blockchain technology: a review," *IEEE Access*, vol. 6, pp. 10179–10188, 2018, doi: 10.1109/ACCESS.2018.2799854.
- [9] H. Artail, H. Safa, M. Sraj, I. Kuwatly, and Z. Al-Masri, "A hybrid honeypot framework for improving intrusion detection systems in protecting organizational networks," *Computers & Security*, vol. 25, no. 4, pp. 274–288, Jun. 2006, doi: 10.1016/j.cose.2006.02.009.
- [10] D. Chaboya, R. Raines, R. Baldwin, and B. Mullins, "Network intrusion detection: automated and manual methods prone to attack and evasion," *IEEE Security and Privacy Magazine*, vol. 4, no. 6, pp. 36–43, Nov. 2006, doi: 10.1109/MSP.2006.159.
- [11] A. Carcano, I. N. Fovino, M. Masera, and A. Trombetta, "State-based network intrusion detection systems for SCADA protocols: a proof of concept," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6027 LNCS, 2010, pp. 138–150. doi: 10.1007/978-3-642-14379-3_12.
- [12] S. Akcay, A. Atapour-Abarghouei, and T. P. Breckon, "Ganomaly: semi-supervised anomaly detection via adversarial trainin," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11363 LNCS, 2019, pp. 622–637. doi: 10.1007/978-3-030-20893-6_39.
- [13] Y. Zhu, N. Zabarar, P.-S. Koutsourelakis, and P. Perdikaris, "Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data," *Journal of Computational Physics*, vol. 394, pp. 56–81, Oct. 2019, doi: 10.1016/j.jcp.2019.05.024.
- [14] A. Prieto *et al.*, "Neural networks: An overview of early research, current frameworks and new challenges," *Neurocomputing*, vol. 214, pp. 242–268, Nov. 2016, doi: 10.1016/j.neucom.2016.06.014.
- [15] M. F. Mushtaq, U. Akram, I. Khan, S. N. Khan, A. Shahzad, and A. Ullah, "Cloud computing environment and security challenges: a review," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 10, 2017, doi: 10.14569/IJACSA.2017.081025.
- [16] A. Ullah, N. M. Nawari, A. Arifianto, I. Ahmed, M. Aamir, and S. N. Khan, "Real-Time Wheat Classification System for Selective Herbicides Using Broad Wheat Estimation in Deep Neural Network," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 9, no. 1, p. 153, Jan. 2019, doi: 10.18517/ijaseit.9.1.5031.
- [17] I. A. T. Hashem, I. Yaqoob, N. B. Anuar, S. Mokhtar, A. Gani, and S. U. Khan, "The rise of 'big data' on cloud computing: Review and open research issues," *Information Systems*, vol. 47, pp. 98–115, Jan. 2015, doi: 10.1016/j.is.2014.07.006.
- [18] I. Yaqoob *et al.*, "Big data: from beginning to future," *International Journal of Information Management*, vol. 36, no. 6, pp. 1231–1247, Dec. 2016, doi: 10.1016/j.ijinfomgt.2016.07.009.
- [19] O. Y. Al-Jarrah, A. Siddiqui, M. Elsalamouny, P. D. Yoo, S. Muhaidat, and K. Kim, "Machine-learning-based feature selection techniques for large-scale network intrusion detection," in *2014 IEEE 34th International Conference on Distributed Computing Systems Workshops*, Jun. 2014, vol. 30-June-20, pp. 177–181. doi: 10.1109/ICDCSW.2014.14.
- [20] S. M. Othman, F. M. Ba-Alwi, N. T. Alsohybe, and A. Y. Al-Hashida, "Intrusion detection model using machine learning algorithm on Big Data environment," *Journal of Big Data*, vol. 5, no. 1, p. 34, Dec. 2018, doi: 10.1186/s40537-018-0145-4.
- [21] M. M. Hassan, A. Gumaei, A. Alsanad, M. Alrubaiyan, and G. Fortino, "A hybrid deep learning model for efficient intrusion detection in big data environment," *Information Sciences*, vol. 513, pp. 386–396, Mar. 2020, doi: 10.1016/j.ins.2019.10.069.
- [22] R. Patgiri, U. Varshney, T. Akutota, and R. Kunde, "An investigation on intrusion detection system using machine learning," in *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, Nov. 2018, pp. 1684–1691. doi: 10.1109/SSCI.2018.8628676.
- [23] J. Camacho, J. M. García-Giménez, N. M. Fuentes-García, and G. Maciá-Fernández, "Multivariate big data analysis for intrusion detection: 5 steps from the haystack to the needle," *Computers & Security*, vol. 87, Nov. 2019, doi: 10.1016/j.cose.2019.101603.
- [24] K. A. P. da Costa, J. P. Papa, C. O. Lisboa, R. Munoz, and V. H. C. de Albuquerque, "Internet of Things: A survey on machine learning-based intrusion detection approaches," *Computer Networks*, vol. 151, pp. 147–157, 2019, doi: 10.1016/j.comnet.2019.01.023.
- [25] H. Alqahtani, I. H. Sarker, A. Kalim, S. M. M. Hossain, S. Ikhlaiq, and S. Hossain, "Cyber intrusion detection using machine learning classification techniques," *Communications in Computer and Information Science*, vol. 1235 CCIS, pp. 121–131, 2020, doi: 10.1007/978-981-15-6648-6_10.
- [26] P. Sangkatsanee, N. Wattanapongsakorn, and C. Charnsripinyo, "Practical real-time intrusion detection using machine learning approaches," *Computer Communications*, vol. 34, no. 18, pp. 2227–2235, Dec. 2011, doi: 10.1016/j.comcom.2011.07.001.
- [27] H. M. Tahir *et al.*, "Hybrid machine learning technique for intrusion detection system," *Hybrid Machine Learning Technique For Intrusion Detection System*, no. 209, pp. 464–472, 2015.
- [28] C. Ieracitano *et al.*, "Statistical analysis driven optimized deep learning system for intrusion detection," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10989 LNAI, 2018, pp. 759–769. doi: 10.1007/978-3-030-00563-4_74.
- [29] R. Abdulhammed, H. Musaffer, A. Alessa, M. Faezipour, and A. Abuzneid, "Features dimensionality reduction approaches for machine learning based network intrusion detection," *Electronics (Switzerland)*, vol. 8, no. 3, Mar. 2019, doi: 10.3390/electronics8030322.
- [30] M. Salman, D. Husna, S. G. Apriliani, and J. G. Pinem, "Anomaly based detection analysis for intrusion detection system using big data technique with learning vector quantization (LVQ) and principal component analysis (PCA)," in *Proceedings of the 2018 International Conference on Artificial Intelligence and Virtual Reality - AIVR 2018*, 2018, pp. 20–23. doi: 10.1145/3293663.3293683.
- [31] S. A. Shah, D. Z. Seker, M. M. Rathore, S. Hameed, S. Ben Yahia, and D. Draheim, "Towards disaster resilient smart cities: Can internet of things and big data analytics be the game changers?," *IEEE Access*, vol. 7, pp. 91885–91903, 2019, doi: 10.1109/ACCESS.2019.2928233.

- [32] F. Amalina *et al.*, "Blending big data analytics: review on challenges and a recent study," *IEEE Access*, vol. 8, pp. 3629–3645, 2020, doi: 10.1109/ACCESS.2019.2923270.
- [33] M. C. Kerman, W. Jiang, A. Blumberg, and S. E. Buttrey, "Event detection challenges, methods, and applications in natural and artificial systems," *The 14th International Command and Control Research and Technology Symposium Washington, DC June 15-17, 2009*, pp. 1–39, 2009.
- [34] S. Ouhame, Y. Hadi, and A. Ullah, "An efficient forecasting approach for resource utilization in cloud data center using CNN-LSTM model," *Neural Computing and Applications*, vol. 33, no. 16, pp. 10043–10055, Aug. 2021, doi: 10.1007/s00521-021-05770-9.
- [35] I. Zografopoulos, J. Ospina, X. Liu, and C. Konstantinou, "Cyber-physical energy systems security: threat modeling, risk assessment, resources, metrics, and case studies," *IEEE Access*, vol. 9, pp. 29775–29818, 2021, doi: 10.1109/ACCESS.2021.3058403.
- [36] D. Picca, "From intelligent to wise machines: why a poem is worth more than 1 million tweets," *Informatik Spektrum*, vol. 43, no. 1, pp. 28–39, Feb. 2020, doi: 10.1007/s00287-020-01245-8.
- [37] A. Ullah and N. M. Nawari, "Enhancing the dynamic load balancing technique for cloud computing using HBATAABC algorithm," *International Journal of Modeling, Simulation, and Scientific Computing*, vol. 11, no. 05, Oct. 2020, doi: 10.1142/S1793962320500415.
- [38] S. Ouhame, Y. Hadi, and A. Arifullah, "A hybrid grey wolf optimizer and artificial bee colony algorithm used for improvement in resource allocation system for cloud technology," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 16, no. 14, p. 4, Nov. 2020, doi: 10.3991/ijoe.v16i14.16623.
- [39] A. Ullah, N. M. Nawari, M. H. Khan, and H. A. Khan, "Rise of big data due to hybrid platform of cloud computing and internet of thing," *Journal of Soft Computing and Data Mining*, vol. 01, no. 01, Mar. 2020, doi: 10.30880/jscdm.2020.01.01.006.
- [40] O. Alkadi, N. Moustafa, and B. Turnbull, "A review of intrusion detection and blockchain applications in the cloud: approaches, challenges and solutions," *IEEE Access*, vol. 8, pp. 104893–104917, 2020, doi: 10.1109/ACCESS.2020.2999715.
- [41] J. D. Kelleher, B. M. Namee, and A. D'Arcy, "Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies," *MIT press*, no. 1, pp. 1–691, 2015.
- [42] A. Ullah, I. Laassar, C. B. Şahin, O. B. Dinler, and H. Aznaoui, "Cloud and internet-of-things secure integration along with security concerns," *International Journal of Informatics and Communication Technology (IJ-ICT)*, vol. 12, no. 1, Apr. 2023, doi: 10.11591/ijict.v12i1.pp62-71.
- [43] A. Ullah, H. Aznaoui, C. B. Şahin, M. Rafie, O. B. Dinler, and L. Imane, "Cloud computing and 5G challenges and open issues," *International Journal of Advances in Applied Sciences*, vol. 11, no. 3, 2022, doi: 10.11591/ijaas.v11.i3.pp187-193.
- [44] A. Ullah and A. Chakir, "Improvement for tasks allocation system in VM for cloud datacenter using modified bat algorithm," *Multimedia Tools and Applications*, vol. 81, no. 20, pp. 29443–29457, Aug. 2022, doi: 10.1007/s11042-022-12904-1.
- [45] A. Ullah, S. A. Khan, T. Alam, S. Luma-Osmani, and M. Sadie, "Heart disease classification using various heuristic algorithms," *International Journal of Advances in Applied Sciences*, vol. 11, no. 2, Jun. 2022, doi: 10.11591/ijaas.v11.i2.pp158-167.
- [46] D. Sebai and A. U. Shah, "Semantic-oriented learning-based image compression by Only-Train-Once quantized autoencoders," *Signal, Image and Video Processing*, 2022, doi: 10.1007/s11760-022-02231-1.
- [47] A. Ullah, A. Salam, H. El Raoui, D. Sebai, and M. Rafie, "Towards more accurate iris recognition system by using hybrid approach for feature extraction along with classifier," *International Journal of Reconfigurable and Embedded Systems*, vol. 11, no. 1, pp. 59–70, 2022, doi: 10.11591/ijres.v11.i1.pp59-70.
- [48] T. Alam, A. Ullah, and M. Benaida, "Deep reinforcement learning approach for computation offloading in blockchain-enabled communications systems," *Journal of Ambient Intelligence and Humanized Computing*, Jan. 2022, doi: 10.1007/s12652-021-03663-2.
- [49] A. Ullah, N. M. Nawari, and S. Ouhame, "Recent advancement in VM task allocation system for cloud computing: review from 2015 to 2021," *Artificial Intelligence Review*, vol. 55, no. 3, pp. 2529–2573, Mar. 2022, doi: 10.1007/s10462-021-10071-7.
- [50] T. Alam, R. Gupta, S. Qamar, and A. Ullah, "Recent applications of artificial intelligence for sustainable development in smart cities," in *Studies in Computational Intelligence*, vol. 1061, 2022, pp. 135–154. doi: 10.1007/978-3-031-14748-7_8.
- [51] Y. Otoum, D. Liu, and A. Nayak, "DL-IDS: a deep learning-based intrusion detection framework for securing IoT," *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 3, Mar. 2022, doi: 10.1002/ett.3803.
- [52] S. Sengan, O. I. Khalaf, P. V. Sagar, D. K. Sharma, L. A. J. Prabhu, and A. A. Hamad, "Secured and privacy-based IDS for healthcare systems on E-medical data using machine learning approach," *International Journal of Reliable and Quality E-Healthcare*, vol. 11, no. 3, pp. 1–11, Jul. 2022, doi: 10.4018/IJRQEH.289175.
- [53] A. K. Balyan *et al.*, "A hybrid intrusion detection model using EGA-PSO and improved random forest method," *Sensors*, vol. 22, no. 16, p. 5986, Aug. 2022, doi: 10.3390/s22165986.
- [54] O. A. Alzubi, "A deep learning-based frechet and dirichlet model for intrusion detection in IWSN," *Journal of Intelligent & Fuzzy Systems*, vol. 42, no. 2, pp. 873–883, Jan. 2022, doi: 10.3233/JIFS-189756.
- [55] A. Abbas, M. A. Khan, S. Latif, M. Ajaz, A. A. Shah, and J. Ahmad, "A new ensemble-based intrusion detection system for internet of things," *Arabian Journal for Science and Engineering*, vol. 47, no. 2, pp. 1805–1819, Feb. 2022, doi: 10.1007/s13369-021-06086-5.
- [56] W. W. Lo, S. Layeghy, M. Sarhan, M. Gallagher, and M. Portmann, "E-GraphSAGE: a graph neural network based intrusion detection system IoT," in *NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium*, Apr. 2022, pp. 1–9. doi: 10.1109/NOMS54207.2022.9789878.
- [57] J. Agle, Y. Xiao, R. Nolan, and L. Golzarri-Arroyo, "Quality control questions on Amazon's Mechanical Turk (MTurk): A randomized trial of impact on the USAUDIT, PHQ-9, and GAD-7," *Behavior Research Methods*, vol. 54, no. 2, pp. 885–897, Apr. 2022, doi: 10.3758/s13428-021-01665-8.
- [58] M. Nagarajan, M. Rajappa, Y. Teekaraman, R. Kuppusamy, and A. R. Thelkar, "Renovated XTEA encoder architecture-based lightweight mutual authentication protocol for RFID and green wireless sensor network applications," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1–12, Mar. 2022, doi: 10.1155/2022/8876096.
- [59] M. A. Harris *et al.*, "Fission stories: using PomBase to understand *Schizosaccharomyces pombe* biology," *Genetics*, vol. 220, no. 4, Apr. 2022, doi: 10.1093/genetics/iyab222.
- [60] M. Alagarsamy, P. Kasinathan, G. Manickam, P. R. Duruvarajan, J. Sakkarai, and K. Suriyan, "IoT based E-vehicle monitoring system using sensors and imaging processing algorithm," *International Journal of Reconfigurable and Embedded Systems*, vol. 11, no. 2, pp. 196–204, Jul. 2022, doi: 10.11591/ijres.v11.i2.pp196-204.
- [61] R. Sudarmani, K. Venusamy, S. Sivaraman, P. Jayaraman, K. Suriyan, and M. Alagarsamy, "Machine to machine communication enabled internet of things: a review," *International Journal of Reconfigurable and Embedded Systems*, vol. 11, no. 2, pp. 126–134, 2022, doi: 10.11591/ijres.v11.i2.pp126-134.




- [62] M. M. Rahman, A. Z. M. T. Kabir, A. M. Mizan, K. M. R. Alvi, N. S. Nabil, and I. Ahmad, "Smart vehicle management by using sensors and an IoT based black box," *International Journal of Reconfigurable and Embedded Systems*, vol. 11, no. 3, pp. 284–294, 2022, doi: 10.11591/ijres.v11.i3.pp284-294.
- [63] D. M. Goldberg, S. Khan, N. Zaman, R. J. Gruss, and A. S. Abrahams, "Text mining approaches for postmarket food safety surveillance using online media," *Risk Analysis*, vol. 42, no. 8, pp. 1749–1768, Aug. 2022, doi: 10.1111/risa.13651.

BIOGRAPHIES OF AUTHORS



Imane Laassar    Working as Research Assistant at Department of Computer Science, Faculty of Science, IbnTofail University, Kenitra, Morocco. Her area of expertise is in Cloud computing, IoT. Areas of interest include software defined networking (SDN), load balancing, switches migration, WSN, E-learning, AI, WSN, and security. She can be contacted at email: imane.laassar@gmail.com.



Moulay Youssef Hadi    Working as Professor at the Department of Computer Science Faculty of Science, IbnTofail University, Kenitra, Morocco. His research interests are distributed computing, parallel computing, software engineering. He can be contacted at email: moulay89@hotmail.com.