Establishing Situational Awareness for Securing Healthcare Patient Records

Aaron Boddy, William Hurst, Michael Mackay, Abdennour El Rhalibi
Department of Computer Science
Liverpool John Moores University
James Parsons Building, Byrom Street
Liverpool, UK, L3 3AF
Email: A.Boddy@2011.ljmu.ac.uk; {W.Hurst, M.I.Mackay, A.Elrhalibi}@ljmu.ac.uk

***Abstract*— The healthcare sector is an appealing target to attackers due to the high value of patient data on the black market. Patient data can be profitable to illegal actors either through direct sale or extortion by ransom. Additionally, employees present a persistent threat as they are able to access the data of almost any patient without reprimand. Without proactive monitoring of audit records, data breaches go undetected and employee behaviour is not deterred. In 2016, 450 data breaches occurred affecting more than 27 million patient records. 26.8% of these breaches were due to hacking and ransomware. In May 2017, a global ransomware campaign adversely affected approximately 48 UK hospitals. Response to this attack, named WannaCry, resulted in hospital networks being taken offline, and non-emergency patients being refused care. Hospitals must maintain patient trust and ensure that the information security principles of Integrity, Availability and Confidentiality are applied to Electronic Patient Record EPR data. With over 83% of hospitals adopting EPRs, access to healthcare data needs to be monitored proactively for malicious activity. Therefore, this paper presents research towards a system that uses advanced data analytics techniques to profile user's behaviour in order to identify patterns and anomalies. Visualisation techniques are then applied to highlight these anomalies to aid the situational awareness of patient privacy officers within healthcare infrastructures.**

Keywords-*Electronic Patient Records; Patient Privacy; Patient Confidentiality; Information Security; Data Analysis; Visualisation; Healthcare Infrastructures.*

I. INTRODUCTION

Healthcare infrastructures are complacent towards the risks of patient privacy violations [1]. Reduced information visibility due to data complexity, fragmentation, interoperability and lack of specialisation undermine the security of these organisations [1]. Visualisation techniques can provide both awareness and modelling capabilities for critical infrastructures [2]. Applying these techniques to aids the understanding of how patient data is accessed within healthcare infrastructures. The goal of security engineers is to develop tools capable of detecting malicious, multi-stage intrusion attacks. These tools should weight the individual attacks, and compare them against the enormous and disparate database of attacks within the network [3]. An intruder's objectives should be determined based on the analysis of the entire dataset of attacks as a whole, rather than just an individual attack [3].

Electronic Patient Record (EPR) systems are vulnerable to both insider and outsider threats [4]. A potential insider threat refers to a legitimate user looking at data when it is not appropriate to do so; such as looking at the record of a celebrity. An external threat is comprised of the theft of a legitimate user's credentials, allowing the attacker uninhibited access to EPR data. This is known as an Active Persistent Threat (APT). It is, therefore, a challenge to mitigate both types of threats.

Current Rules-Based solutions to these issues are effective at detecting predictable insider threats. They can process the large quantities of audit data, and can process rules against that data. For example, a rule can be set to inform Information Security if anyone other than a set list of clinicians accesses the patient record of a celebrity or famous individual. Any violation of this will then be reported automatically to the Information Security team. However, this cannot detect the threat of an attacker who has acquired the logon credentials of a clinician; which is achieved through either phishing or social engineering techniques and enables EPR data exfiltration. Additionally, rules are also be set to detect if a user is looking at the record of a patient with the same surname as them to identify potential patient confidentiality violations. Similarly, if an attacker has unauthorised login credentials (and is surveying a patient's record) the rules set would not make provisions for the detection.

Therefore, this paper proposes an advanced data analytics and visualisation-based approach to patient privacy violation detection within EPR systems. Advanced data analytics algorithms have the capability to learn patterns of data and profile users' behaviour, which can then be represented visually. Advanced data analytics detect when a user's behaviour has changed, by comparing behaviours, such as the type of actions being taken and the patients they are viewing.

It is unfeasible to detect fully all illegitimate access within EPR systems, but it is feasible to eliminate legitimate access. In doing so, it becomes possible to focus the attention of information security analysts to where it is needed, within the comprehensive EPR audit datasets.

The remainder of this paper is as follows. Section II presents a literature review of the background research on patient privacy within EPR systems. Section III outlines the systematic approach and presents our results and a sample of test data. Section IV discusses conclusions and the future work to be done.

II. BACKGROUND

Authorised users can access EPR data from virtually anywhere; allowing increased productivity compared with paper-only records and allowing clinicians to make informed decisions towards improving healthcare quality for patients [5]. The management of patient data in electronic form decreases healthcare administration costs, strengthens care provider productivity and increases patient safety [6].

The proliferation of technology within healthcare has brought the advantages of improved efficiency of record keeping, easier detection and prevention of fraud, waste and abuse, and an improvement on the overall quality of care [7]. However, with the added benefits of technology in healthcare, the potential for unauthorised and illegal access to patient information has increased [8]. Users may abuse their privileges for personal reasons, such as viewing records of relatives, friends, neighbours, co-workers or celebrities [5]. Therefore, patients are becoming increasingly concerned regarding the privacy and security of their health data [9]. The cost to a healthcare organisation caused by a security breach is one of the highest of any industry and leads to the loss of trust of patients [10].

A. EPR Audit Logs

When there is reason to suspect that unauthorised accesses have occurred, a review of the audit logs is undertaken by a security expert. This is inefficient, as it requires the information to be collated and reviewed by a security expert. It is a process that is purely retrospective [5]. Therefore, there is a motivation to automate and alleviate the burdens associated with this process [7]. The fundamental limitations in privacy officers manually reviewing audit logs for potentially suspicious accesses are threefold [5]. Firstly, the volume of audit records means that audit logs are only practically useful as adjuncts to investigate suspected breaches, rather than a tool that can be utilised to proactively find inappropriate accesses. Secondly, audit records can only provide data regarding the access itself, and contains no situational or relationship information or knowledge regarding the access. Thirdly, the process is labour-intensive, without guidance of where to look for potential breaches, inappropriate accesses are buried amongst the audits of appropriate accesses.

B. Access Control

Healthcare systems typically employ access control solutions [11], where once an individual has been authenticated, they are allowed unhindered access inside the perimeter [5]. It is a challenge to impose an access control policy on employees in a healthcare setting due to the dynamic and unpredictable care patterns of hospital care [7].

The Access Matrix Model (AMM) is a conceptual framework that specifies each user's permissions for each object in the system [12]. Although it allows for a thorough mapping of access rights, it does not scale well, and lacks the ability to support dynamic changes of access rights, which makes it difficult to apply to EPRs [13].

Role-Based Access Control (RBAC) maps users to roles and maps permissions to the roles [14]. Job positions within the enterprise and tasks the employees need to perform are identified, and privileges are assigned to these positions to enable the employees to accomplish their tasks [15]. Whilst more computationally tractable, RBAC roles tend to be static and inflexible, and therefore not responsive to the shifting nature of roles [16].

Attribute-Based Access Control (ABAC) provides flexible, context-aware access control through evaluating the attributes of entities, it's subject and object, the operation, and the contextual environment of the request [17]. Boolean logic can then be applied to the operational request to determine access rights. ABAC therefore allows for a higher number of discrete inputs and provides a larger, more definitive set of rules to express policies than RBAC.

Experience Based Access Management (EBAM) emphasises the accountability and use of audit data to detect illegitimate access [15]. EBAM enterprises often manually review the audit logs of VIPs to determine inappropriate accesses [18]. Break The Glass (BTG) is a policy, which allows users to override access controls in necessary instances [19]. EBAM enterprises would manually review the audit logs every time a user broke the glass [15].

Task-based Access Control (TBAC) extends the user-object relationship though the inclusion of task-based and contextual information [20]. However TBAC is limited to contexts that relate to tasks, or workflow progress and EPRs cannot always be easily portioned into tasks [13].

Team-based Access Control (TeBAC) groups users in an organisation and associates a collaboration context with the activity to be performed [21]. However, these models have not been fully developed or implemented and it remains unclear how to implement them within a dynamic framework [13].

C. Detection Approaches

The following section examines several related common detection approaches to anomaly detection in large datasets:

- Signature detection is a rules-based algorithm that constructs a set of rules based on historic breaches and can detect correctly known patterns whilst being interpretable [22]. However, it cannot detect unseen patterns and cannot assign risk scores [23].
- Anomaly detection compares incoming instances to previously built profiles and can detect novel patterns, although it requires a large quantity of historic data [24]. Additionally, the output is known to be problematic to interpret and the technique produces false positives [25].
- Clustering is invoked to integrate similar data instances into groups [26]. Clustering evaluates each instance with respect to the cluster it belongs to, while nearest neighbour analyses each instance with respect to its own local neighbourhood [13].
- Spectral Projection estimates the principal components from the covariance of the training data of normal events [27]. The testing phase compares each point with the components and assigns an anomaly score based on the points distance from the principal components [13].

- Classifier detection determines a classification function based on a labelled training set [28] and can be fast, accurate and assign risk scores to all events [29]. However, acquiring class labelled data is expensive and scoring unlabelled events is important in large scale data mining, as human validation is limited and costly [30].

D. Related Work

Machine learning models are trained on historical access data to classify future data access patterns [7]. Supervised machine learning models, such as Support Vector Machines (SVMs), linear regression and logistic regression have been applied successfully to the challenge of detecting inappropriate access within Electronic Patient Record systems [5][7][10].

For example, Community-based Anomaly Detection System (CADS) is an unsupervised learning framework to detect insider threats based on information recorded in audit logs of collaborative environments [13]. It is based on the observation that typical users tend to form community structures, so users with a low connection, to such communities, are indicative of anomalous behaviour. The model consists of two primary components. Firstly, relational pattern extraction, which infers community structures from access logs and subsequently derives communities, which serve as the CADS core. Secondly, potentially illicit behaviour, where CADS uses a formal statistical model to measure the deviation of users from the inferred communities to predict which users are anomalies [13]. CADS does not implement supervised learning techniques to further classify the data with feedback from patient privacy officers.

AI^2 is another example of a cyber-security machine learning system, which improves its accuracy over time through feedback from security analysts [31]. AI^2 is composed of the following four components. Firstly, a Big Data Processing System, which quantifies the behaviours and features of raw data. Secondly, an Outlier Detection System, which learns a descriptive model of data features extracted via unsupervised learning, using either density, matrix decomposition, or replicator neural networks. Thirdly, a Feedback Mechanism and Continuous Learning, which incorporates analyst input through a user interface. The system highlights the top *k* outlier events or entities and tasks the analyst with identifying whether they are malicious; the feedback is then input back into the supervised learning module. Fourthly, a Supervised Learning Model, which predicts whether a new incoming event is normal or malicious, and uses analysts feedback to refines the model. Raw data is input into AI^2 that computes features describing the entities of the data set. Using these features, an unsupervised machine learning module identifies extreme and rare events in the data. These events are then ranked based upon a predefined metric and presented to the analyst, who ranks the behaviours as normal or malicious (and as pertaining to a particular attack type). Finally, these labels are input to the supervised learning module. The novelty of the system proposed in this paper, to that of AI^2, is the addition of visualisation techniques to aid the analyst to understand and explore the data. There is also a specific focus on EPR data, which differ from other enterprise infrastructures due to their reliance on insecure medical devices, legacy systems, and bespoke software.

The use of statistical and machine learning techniques have previously been used to detect fraud in financial reporting [32], to detect fraud in credit card transaction data [33], to construct a spam email detector [34], and to solve a fraud detection problem at a car insurance company [35].

III. Approach

As the background demonstrates, there is a clear need to address the issue of lack of situational awareness on the part of information security professionals within healthcare infrastructures. In this section, an approach is put forward for analysing data within healthcare infrastructures, processing it to eliminate low-risk data points and visualising it in such a way that data anomalies become apparent. Our research to date has focused on the development of a system for modelling data flow within healthcare infrastructures [36][37]. The system assists information security officers, within healthcare organisations, to improve the situational awareness of patient data confidentiality risks.

A. Approach

The novel contribution presented in [36][37] involves the use of advanced data analytics techniques, in addition to an analyst-in-the-loop and the use of visualised attack events. Low-risk data is analysed, processed and pre-filtered using advanced data analytics techniques before the visualisation of the data. This is then visualised and presented to an analyst. The analyst then classifies events within the presented visualisation, which provides feedback to the system. Through the use of the analyst-in-the-loop both models are used to continuously defend the healthcare infrastructure against current attack vectors. The aim is to collect, process, and filter big data sets to provide users of an overall understanding of system behaviour in order to detect security breaches and general anomalies.

The system provides contextual awareness to detect anomalous behaviour within EPR audit activity. The main challenge of the work involves big data analytics to process datasets generated by healthcare infrastructures.

The system put forward in this paper combines several related data sets and presents them, in such a way, as to identify relationships between them. EPR audit data and behavioural patterns are understood, in order to assist end users in finding the potential vulnerabilities within the health care infrastructure. The data analysis techniques involve interpreting dataset patterns and identify potential on-going patient privacy violations.

The visualisations cluster together salient points and use size to indicate potential threat levels. This gives the analyst a broad overview of the current EPR security at a glance. From here, the visualisation can be interacted with, explored by the analyst to investigate the data points and find in-depth technical information about each data point. Additionally, the analyst can provide feedback to the system and rank the

highlighted data points as either safe, or as pertaining to a patient privacy violation.

B. *Case Study*

In this section, a case study of the EPR audit data is presented. This rich dataset contains 1,007,727 rows of audit logs of every user and their EPR activity in a UK hospital over a period of 18 months (28-02-16 – 21-08-17). Each User UID, Patient UID and Device name is tokenised through isolating the unique entries and assigning each value an incrementing number. There are 1,515 unique User UIDs, 72,878 unique Patient UIDs and 2,270 unique Devices within the dataset.

The dataset consist of the following fields:

- *Date* - The date the patient record was accessed
- *Time* - The time the patient record was accessed
- *Device (Tokenised)* - The name of the device the patient record was accessed
- *User UID (Tokenised)* - A tokenised representation of the User who accessed the patient record
- *Routine* - The routine performed whilst accessing the patient record (was the record updated, was a letter printed etc.)
- *Patient UID (Tokenised)* - A tokenised representation of the patient record that was accessed
- *Duration* - The number of seconds the patient record was accessed (this number counts for as long as the record is on the screen, so may not always be an accurate reflection of how long the User was actively interacting with the data)
- *Latest Adm Date* - The date the patient was last admitted to the hospital
- *Latest Dis Date* - The date the patient was last discharged from the hospital

A snapshot of the first 10 rows in the dataset is presented in Table 1.

TABLE I. EPR AUDIT SAMPLE DATA

Date	Time	Device	User U	Routine	Patient	Durat	LocationLa	Latest Dis
28-02-16	00:00	362	865	PHA.ORD	58991	54	28-02-16	29-02-16
28-02-16	00:02	103	677	ASF	4786	13	22-07-08	22-07-08
28-02-16	00:02	103	677	ASF	4786	54	22-07-08	22-07-08
28-02-16	00:02	923	199	REC REC:	17278	77	15-02-16	15-02-16
28-02-16	00:04	103	677	ASF VH	14067	39	28-09-04	28-09-04
28-02-16	00:04	845	1489	PHA.ORD	49304	22	23-01-02	23-01-02
28-02-16	00:04	923	199	REC UK.C	62121	147	08-02-16	08-02-16
28-02-16	00:06	923	199	REC REC:	60948	165	08-01-16	08-01-16
28-02-16	00:08	775	568	NOTE	32826	75	25-01-12	25-01-12
28-02-16	00:10	393	1361	PHA.ORD	28106	49	16-08-06	16-08-06

In Figure 1, a heatmap is presented of the dataset comparing User UID to the duration of the patient record access. The graph shows that there is consistent point density of up to 47,341 in the first row of the matrix, indicating that most patient records are only accessed for fewer than 300 seconds (5 minutes). This would represent *normal* behaviour within the hospital. Representing the data as a heatmap highlights clear anomalies in the data.

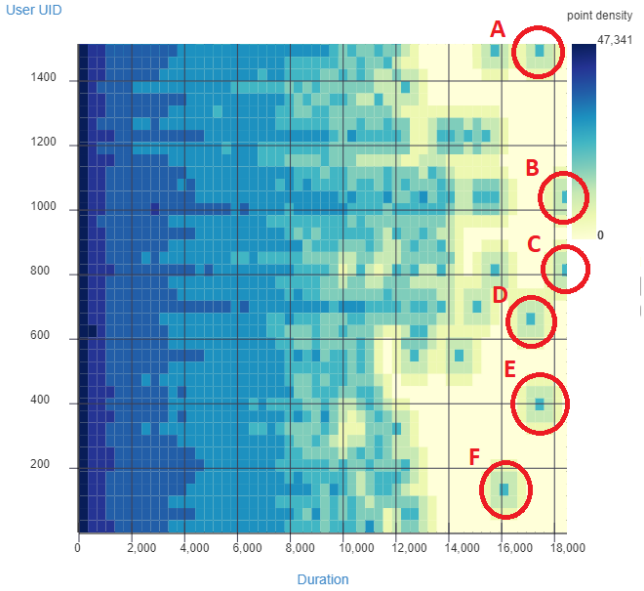

Figure 1 - Heatmap - User UID and Duration

Notably, as displayed in Figure 1, users B and C are identified spending over 18,000 seconds (over 5 hours) accessing patient records. Additionally users, A, D, E and F all spent over 16,000 seconds (almost 4.5 hours) accessing records. These anomalies can be investigated by an analyst indicating potentially illegitimate access to EPR data.

C. *Discussion*

These initial results display only preliminary explorations of the dataset and demonstrate the potential insights the dataset holds. Once feature selection and pre-processing work has been completed on the data, machine learning models will be used to explore the data further, with a particular emphasis on unsupervised learning such as clustering. This will allow initial patterns within the data to be identified to understand the data and identify illegitimate access to patient records within this real world EPR dataset. Extracting features from this data (such as mean, median, mode and range of duration), will be used to train classifiers to autonomously learn normal and abnormal patterns through supervised learning techniques. This process will occur once the data has been clustered through the use of unsupervised learning algorithms. In combing both unsupervised and supervised machine learning techniques, the system will aid privacy officers in their situational awareness of access to patient records and identify outliers for investigation.

IV. CONCLUSION AND FUTURE WORK

Electronic Patient Record (EPR) data is both sensitive and valuable. Patients need to be assured of three crucial security principles regarding their healthcare data. Firstly, patients need to be assured that the data stored is trustworthy and

accurate. Secondly, patients need to be assured that the data can be reliably accessed by healthcare professionals when needed. And thirdly, patients need to be assured that only authorised healthcare professionals have access to the data, and only access it when it is appropriate to do so. It is therefore of utmost importance that the Information Security principles of Integrity, Availability and Confidentiality are applied to EPR data. Therefore, this paper presents research towards a system, which can detect unusual data behaviour through the use of advanced data analytics and visualisation techniques. Machine learning algorithms have the capability to learn patterns of data and profile users' behaviour, which can be represented visually. The proposed system is tailored to healthcare infrastructures by learning typical data behaviours and profiling users. The system adds to the defence-in-depth of the healthcare infrastructure by understanding the unique configuration of the EPR and autonomously analysing user's access.

Future work will build on the visualisation work undertaken in the research case studies presented in this paper. The visualisations will allow the user to explore the data and understand the patterns and trends within the comprehensive EPR audit data sets. Unsupervised machine learning techniques will be implemented to classify this data in future work as there is limited abnormal data and a lack of labelled training data. Feedback from the analysts will inform the machine learning algorithms and refine the results to reduce alert fatigue. Machine Learning algorithms will allow the system to pick up on patterns and trends in the date without being explicitly taught them, as in Rules-Based Analytics. For example, if a user typically only logs into their account on weekdays, then if the account is logged in on a weekend, it may be an indication that the users' username and password has been compromised by an attacker. The attacker could either be illegally accessing hospital records, or searching for further vulnerabilities within the EPR in order to perform a privilege escalation attack.

Additionally, the machine learning algorithms will be automated and tested on "live" real-world data once it has been refined. This will allow the process outlined in this paper to alert information security analysts of illegitimate shortly after they occur. Over time, the analyst will be able to provide feedback to the system through the use of supervised machine learning algorithms, and the algorithms will be refined and tailored to the unique threat landscape and infrastructure of the hospital.

REFERENCES

[1] J. Stoll and R. Z. Bengez, "Visual structures for seeing cyber policy strategies", in *2015 7th International Conference on Cyber Conflict: Architectures in Cyberspace*, pp. 135–152, 2015.

[2] M. Merabti, M. Kennedy, and W. Hurst, "Critical infrastructure protection: A 21st century challenge", in *2011 International Conference on Communications and Information Technology (ICCIT)*, pp. 1–6, 2011.

[3] J. J. Walker, T. Jones, and R. Blount, "Visualization, modeling and predictive analysis of cyber security attacks against cyber infrastructure-oriented systems", in *2011 IEEE International Conference on Technologies for Homeland Security (HST)*, pp. 81–85, 2011.

[4] M. Rahman and C. Kreider, "Information Security Principles for Electronic Medical Record (EMR) Systems", *AMCIS 2012 Proc.*, Jul. 2012.

[5] J. Kim *et al.*, "Anomaly and signature filtering improve classifier performance for detection of suspicious access to EHRs.", *AMIA ... Annu. Symp. proceedings. AMIA Symp.*, vol. 2011, pp. 723–31, 2011.

[6] N. Menachemi and R. G. Brooks, "Reviewing the benefits and costs of electronic health records and associated patient safety technologies.", *J. Med. Syst.*, vol. 30, no. 3, pp. 159–68, Jun. 2006.

[7] A. K. Menon, X. Jiang, J. Kim, J. Vaidya, and L. Ohno-Machado, "Detecting Inappropriate Access to Electronic Health Records Using Collaborative Filtering.", *Mach. Learn.*, vol. 95, no. 1, pp. 87–101, Apr. 2014.

[8] O. of T. A. U.S. Congress, "*Electronic Record Systems and Individual Privacy*." (Washington, DC: Federal Government Information Technology, 1986.

[9] D. C. Kaelber, A. K. Jha, D. Johnston, B. Middleton, and D. W. Bates, "A Research Agenda for Personal Health Records (PHRs)", *J. Am. Med. Informatics Assoc.*, vol. 15, no. 6, pp. 729–736, Nov. 2008.

[10] A. A. Boxwala, J. Kim, J. M. Grillo, and L. Ohno-Machado, "Using statistical and machine learning to help institutions detect suspicious access to electronic health records", *J. Am. Med. Informatics Assoc.*, vol. 18, no. 4, pp. 498–505, Jul. 2011.

[11] D. F. Ferraiolo, R. Sandhu, S. Gavrila, D. R. Kuhn, and R. Chandramouli, "Proposed NIST standard for role-based access control", *ACM Trans. Inf. Syst. Secur.*, vol. 4, no. 3, pp. 224–274, Aug. 2001.

[12] K. Sikkel, "A Group-based Authorization Model for Cooperative Systems", in *Proceedings of the Fifth European Conference on Computer Supported Cooperative Work*, Dordrecht: Springer Netherlands, pp. 345–360, 1997.

[13] Y. Chen and B. Malin, "Detection of Anomalous Insiders in Collaborative Environments via Relational Analysis of Access Logs.", *CODASPY Proc. ... ACM Conf. data Appl. Secur. privacy. ACM Conf. Data Appl. Secur. Priv.*, vol. 2011, pp. 63–74, 2011.

[14] G.-J. Ahn, D. Shin, and L. Zhang, "Role-Based Privilege Management Using Attribute Certificates and Delegation", Springer, Berlin, Heidelberg, pp. 100–109, 2004.

[15] W. Zhang, C. A. Gunter, D. Liebovitz, J. Tian, and B. Malin, "Role prediction using Electronic Medical Record system audits.", *AMIA ... Annu. Symp. proceedings. AMIA Symp.*, vol. 2011, pp. 858–67, 2011.

[16] R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman, "Role-based access control models", *Computer (Long. Beach. Calif).*, vol. 29, no. 2, pp. 38–47, 1996.

[17] V. C. Hu, D. R. Kuhn, and D. F. Ferraiolo, "Attribute-Based Access Control", *Computer (Long. Beach. Calif).*, vol. 48, no. 2, pp. 85–88, Feb. 2015.

[18] Z. Zhou and B. J. Liu, "HIPAA compliant auditing system for medical images", *Comput. Med. Imaging Graph.*, vol. 29, no. 2–3, pp. 235–241, Mar. 2005.

[19] A. Ferreira *et al.*, "How to Break Access Control in a Controlled Manner", in *19th IEEE Symposium on Computer-Based Medical Systems (CBMS'06)*, pp. 847–854, 2006.

[20] T. A. Sandhu, T. A. Sandhu, and R. K. Thomas, "Task-based Authorization Controls (TBAC): A Family of Models for Active and Enterprise-oriented Authorization Management", *Proc. IFIP WG11.3 Work. DATABASE Secur. LAKE TAHOE*, pp. 166--181, 1997.

[21] C. K. Georgiadis, I. Mavridis, G. Pangalos, and R. K. Thomas, "Flexible team-based access control using contexts", in *Proceedings of the sixth ACM symposium on Access control models and technologies - SACMAT '01*, pp. 21–27, 2001.

[22] D. Barbara, *Applications of Data Mining in Computer Security*, vol. 6. Springer US, 2002.

[23] Wenke Lee, S. J. Stolfo, and K. W. Mok, "A data mining framework for building intrusion detection models", in *Proceedings of the 1999 IEEE Symposium on Security and Privacy (Cat. No.99CB36344)*, pp. 120–132, 1999.

[24] K. Das, J. Schneider, and D. B. Neill, "Anomaly pattern detection in categorical datasets", in *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD 08*, p. 169, 2008.

[25] A. Patcha, J. Park, "An overview of anomaly detection techniques: Existing solutions and latest technological trends", *Comput. Networks*, vol. 51, no. 12, pp. 3448–3470, Aug. 2007.

[26] Z. He, X. Xu, and S. Deng, "Discovering cluster-based local outliers", *Pattern Recognit. Lett.*, vol. 24, no. 9–10, pp. 1641–1650, Jun. 2003.

[27] M.-L. Shyu, S.-C. Chen, K. Sarinnapakorn, and L. Chang, "Principal Component-based Anomaly Detection Scheme", in *Foundations and Novel Approaches in Data Mining*, Berlin/Heidelberg: Springer-Verlag, pp. 311–329, 2005.

[28] A. Shen, R. Tong, and Y. Deng, "Application of Classification Models on Credit Card Fraud Detection", in *2007 International Conference on Service Systems and Service Management*, pp. 1–4, 2007.

[29] S. Mukkamala, G. Janoski, and A. Sung, "Intrusion detection using neural networks and support vector machines", in *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No.02CH37290)*, pp. 1702–1707, 2002.

[30] J. Zhu, H. Wang, B. K. Tsou, and M. Ma, "Active Learning With Sampling by Uncertainty and Density for Data Annotations", *IEEE Trans. Audio. Speech. Lang. Processing*, vol. 18, no. 6, pp. 1323–1331, Aug. 2010.

[31] K. Veeramachaneni, I. Arnaldo, V. Korrapati, C. Bassias, and K. Li, "AI2: Training a Big Data Machine to Defend", in *Proceedings - 2nd IEEE International Conference on Big Data Security on Cloud, IEEE BigDataSecurity 2016, 2nd IEEE International Conference on High Performance and Smart Computing, IEEE HPSC 2016 and IEEE International Conference on Intelligent Data and S*, pp. 49–54, 2016

[32] T. B. Bell and J. V. Carcello, "A Decision Aid for Assessing the Likelihood of Fraudulent Financial Reporting", *Audit. A J. Pract. Theory*, vol. 19, no. 1, pp. 169–184, Mar. 2000.

[33] Tao Guo and Gui-Yang Li, "Neural data mining for credit card fraud detection", in *2008 International Conference on Machine Learning and Cybernetics*, pp. 3630–3634, 2008.

[34] K. Yoshida *et al.*, "Density-based spam detector", in *Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '04*, p. 486, 2004.

[35] J. M. Pérez, J. Muguerza, O. Arbelaitz, I. Gurrutxaga, and J. I. Martín, "Consolidated Tree Classifier Learning in a Car Insurance Fraud Detection Domain with Class Imbalance", Springer, Berlin, Heidelberg, pp. 381–389, 2005.

[36] A. Boddy, W. Hurst, M. MacKay, and A. El Rhalibi, "A Study into Detecting Anomalous Behaviours within HealthCare Infrastructures", *9th Int. Conf. Dev. eSystems Eng.*, pp111-117, 2016.

[37] A. Boddy, W. Hurst, M. Mackay, and A. El Rhalibi, "A Study into Data Analysis and Visualisation to increase the Cyber-Resilience of Healthcare Infrastructures", *Internet Things Mach. Learn.*, 2017.