A MACHINE LEARNING FRAMEWORK FOR SECURING PATIENT RECORDS

AARON JOHN BODDY

A thesis submitted in partial fulfilment of the requirements of Liverpool John Moores University for the degree of Doctor of Philosophy

May 2019

Acknowledgements

I would like to thank the following people for their contribution and support throughout this journey.

Dr. William Hurst for being my Director of Studies and giving me consistent and valuable support throughout the project. I would never have pursued a PhD if it wasn't for Will. I wasn't an obvious candidate and it took a few tries before I was accepted but Will was a firm believer that my enthusiasm could carry me through, even at times when my own confidence in this was waning. I genuinely cannot see myself having actually completed a PhD with anyone but Will and I can't thank him enough for his tireless efforts and his unwavering belief in me.

Dr. Michael Mackay and **Prof. Abdennour El Rhalibi** for being my supervisors and for their valuable contributions and insights as this work progressed.

Allison Boddy and **Paul Boddy** for being my Mum and Dad and for supporting me throughout all these years.

Danny Boddy and **Charlotte Boddy** for being my Brother and Sister and telling people that their big brother is doing a PhD in something to do with computers.

John Cooper for helping me understand the more unwieldy side of Excel and for helping me think things through when it all got too much. I'd also like to thank you for the bus ride where you encouraged me to pursue this PhD when I didn't think I could alongside a full-time job.

And finally, **Mary Jayne Cooper**. In my undergraduate dissertation I thanked you for being my best friend. Now I'd like to thank you for being my fiancée.

Table of Contents

Table of Figures

Table of Tables

Publications resulting from this Thesis

Journals:

1. **Boddy, A.**, Hurst, W., Mackay, M., & El Rhalibi, A., *Density-Based Outlier Detection for Safeguarding Electronic Patient Record Systems*, IEEE Access Open Source Journal, vol7 No1, March 2019 - Published

2. **Boddy, A.**, Hurst, W., Mackay, M., El Rhalibi, A., Baker, T., & Curbelo Montañez, C., *An Investigation into Healthcare-Data Patterns*, Special Issue on Smart Systems for Healthcare, Future Internet Open Access Journal, MDPI, 2019 – Published

3. Curbelo Montañez, C., Hurst, W., Chalmers, C., Mackay, M., **Boddy, A.**, *An Investigation into Density-based Anomaly Detection Analysis for Smart Meter Data,* Elsevier Array Open Access Journal – Under Submission

Conference Papers:

1. **Boddy, A.**, Hurst, W., Mackay, M., & El Rhalibi, A., *Establishing Situational Awareness for Securing Healthcare Patient Records*, HEALTHINFO 2019 The Fourth International Conference on Informatics and Assistive Technologies for Health-Care, Medical Support and Wellbeing, Valencia, Spain, 24th–28th November 2019

2. **Boddy, A.**, Hurst, W., Mackay, M., & El Rhalibi, A., *A Hybrid Density-Based Outlier Detection Model for Privacy in Electronic Patient Record Systems*, ICIM 2019 The 5th International Conference on Information Management, Cambridge, UK, 24th–27th March 2019

3. **Boddy, A.**, Hurst, W., Mackay, M., El Rhalibi, A. & Mwansa, M., *Data Analysis Techniques to Visualise Accesses to Patient Records in Healthcare Infrastructures*, CLOUD COMPUTING 2018 The Ninth International Conference on Cloud Computing, GRIDs, and Virtualization, Barcelona, Spain, 18th–22nd February 2018

4. Mwansa, M., Hurst, W., Chalmers, C., Shen, Y. & **Boddy, A.**, *A Study into Smart Grid Consumer-User Profiling for Security Applications*, CLOUD COMPUTING 2018 The Ninth International Conference on Cloud Computing, GRIDs, and Virtualization, Barcelona, Spain, , 18th–22nd February 2018

5. **Boddy, A.**, Hurst, W., Mackay, M., & El Rhalibi, A., *A Study into Data Analysis and Visualisation to increase the Cyber-Resilience of Healthcare Infrastructures*, ACM International Conference on Internet of Things and Machine Learning (IML 2017), Liverpool, England, 17th-18th October 2017

6. **Boddy, A.**, Hurst, W., Mackay, M., & El Rhalibi, A., *A Study into Detecting Anomalous Behaviours within HealthCare Infrastructures*, Developments in eSystems

Engineering (DeSE) 2016, Liverpool/Leeds, England, 31st September-2nd October 2016.

7. **Boddy, A.**, Hurst, W., Mackay, M., & El Rhalibi, A., *Securing Health Care Information Systems using Visualisation Techniques*, UK Academy for Information Systems, Oxford, England, 11th-13th April 2016.

Grants and Awards:

1. **Boddy, A.**, Hurst, W., Mackay, M., & El Rhalibi, A., Aintree University Hospital – A System for Improving the Cyber Resilience of Healthcare Networks – £2,680.
2. **Boddy, A.**, LJMU Chancellors Student Reception attendee – 16th Nov 2017.

Thesis Acronyms

ABAC – Attribute-Based Access Control

AD – Active Directory

AI^2 – Artificial Intelligence x Analyst Intuition

AMM – Access Matrix Model

API – Application Programming Interface

APT – Active Persistent Threat

ASA – Adaptive Security Appliance

BAN – Body Area Network

BYOD – Bring Your Own Device

CADS – Community-based Anomaly Detection System

CareCERT – Care Computer Emergency Response Team

CD – Cardiac Device

CoIN – Community of Interest Network

CP-IS – Child Protection - Information Sharing project

CPU – Central Processing Unit

DAG – Directed Acyclic Graph

DBSCAN - Density-Based Spatial Clustering of Applications with Noise

DDoS – Distributed Denial of Service

DMZ – De-Militarized Zone

EBAM – Experience Based Access Management

ECN – Electronic Case Notes

EDMS – Electronic Document Management System

ESR – Electronic Staff Record

EPMA – Electronic Prescribing and Medications Administration

EPR – Electronic Patient Record

eRS – E-Referral System

FDA – Food and Drug Administration

GBML - Genetics-based machine learning

GP – General Practitioner

HILML – Human-in-the-Loop Machine Learning

HIPAA – Health Insurance Portability and Accountability Act of 1996

HSCIC – Health and Social Care Information Centre (now NHS Digital)

HSCN – Health & Social Care Network

ICO – Information Commissioner’s Office

IDF – Inverse Document Frequency

IDS – Intrusion Detection System

IP – Internet Protocol

IT – Information Technology

LAN – Local Area Network

LDC – Linear Discriminant Classifier

LIMS – Laboratory Information Management System

LOF – Local Outlier Factor

LR – Logistic Regression

LRD – Local Reachability Distance

LSASS – Local Security Authority Subsystem Service

MAP1/MAP2 – Monitoring Access Pattern phase 1/2

MAUDE - Manufacturer and User Facility Device Experience

MCPS – Medical Cyber Physical System

MCPS – Medical Cyber-Physical System

MESH – Message Exchange for Social Care and Health

NHS – National Health Service

PACS – Picture Archiving and Communication System

PARISS – Patient Record Intelligent Security System

PAS – Patient Administration System

PC – Personal Computer

PDS – Personal Demographics Service

PIX – Private Internet eXchange

PKI – Public Key Infrastructure

QDC – Quadratic Discriminant Classifier

RAS – Remote Access Server

RBAC – Role-Based Access Control

RIS – Radiology Information System

SIEM – Security Information and Event Management

SMB – Server Message Block

OS – Operating System

OPTICS – Ordering Points to Identify the Clustering Structure

SQL – Structured Query Language

SVM – Support Vector Machine

TBAC – Task-based Access Control

TCP – Transmission Control Protocol

TeBAC – Team-based Access Control

TF-IDF – Term Frequency-Inverse Document Frequency

TIE – Trust Integration Engine

TMG – Threat Management Gateway

UDP – User Datagram Protocol

VIP – Very Important Person

VLAN – Virtual Local Area Network

VPN – Virtual Private Network

WAF – Web Application Firewall

WAN – Wide Area Network

Abstract

This research concerns the detection of abnormal data usage and unauthorised access in large-scale critical networks, specifically healthcare infrastructures. The focus of this research is safeguarding Electronic Patient Record (EPR) systems in particular. Privacy is a primary concern amongst patients due to the rising adoption of EPR systems. There is growing evidence to suggest that patients may withhold information from healthcare providers due to lack of Trust in the security of EPRs. Yet, patient record data must be available to healthcare providers at the point of care. Roles within healthcare organisations are dynamic and relying on access control is not sufficient. Access to EPR is often heavily audited within healthcare infrastructures. However, this data is regularly left untouched in a data silo and only ever accessed on an ad hoc basis. In addition, external threats need to be identified, such as phishing or social engineering techniques to acquire a clinician's logon credentials. Without proactive monitoring of audit records, data breaches may go undetected. This thesis proposes a novel machine learning framework using a density-based local outlier detection model, in addition to employing a Human-in-the-Loop Machine Learning (HILML) approach. The density-based outlier detection model enables patterns in EPR data to be extracted to profile user behaviour and device interactions in order to detect and visualise anomalous activities. Employing a HILML model ensures that inappropriate activity is investigated and the data analytics is continuously improving. The novel framework is able to detect 156 anomalous behaviours in an unlabelled dataset of 1,007,727 audit logs.

1. Introduction

Data behaviour within healthcare infrastructures needs to be monitored for malicious, irregular or unusual activity. For example in May 2017, a well-documented global ransomware campaign, referred to as WannaCry, targeted Windows operating systems worldwide and in doing so, adversely affected approximately 60 NHS trusts, 595 General Practices (GPs) and thousands of patients [1]. Response to the WannaCry cyber-attack resulted in many hospital networks being taken offline, and non-emergency patients being refused care. However, there is still a perceived lack of threat within healthcare organisations with regards to cyber-security. Hospitals must maintain patient trust and ensure that the information security principles of Integrity, Availability and Confidentiality are applied to EPR data. Hospital infrastructures present a unique threat vector, with a dependence on legacy software, medical devices, and bespoke software. Additionally, many PCs are shared by a number of users, all of whom use a variety of disparate IT systems. Hospitals in the UK are now connecting their traditionally isolated equipment on a large scale to Internet-enabled networks to enable remote data access. With a push for all hospitals in the UK to be paperless by 2020 [2], access to this healthcare data needs to be proactively monitored for malicious activity. This step-change makes sensitive data accessible to a broader spectrum of users. Every healthcare infrastructure configuration is unique and a one-size-fits-all security solution cannot be applied to healthcare. Existing cyber-security technology within hospital infrastructures is typically perimeter-focused [3], so once a malicious user has compromised the boundary through a backdoor, there is a lack of security architecture monitoring active potential threats inside the network.

Modern Information Technology (IT) systems are crucial to clinical care service providers. They are relied upon to collect and store sensitive patient data, govern human life-support devices and enable communication for archiving and information sharing [4]. Disabling or disrupting any of these systems would have far reaching consequences within healthcare infrastructures. Relying on traditional security models to safeguard these systems has proven to be ineffective; particularly in relation to the emergence of new technology such as mobility, cloud, social media and Bring Your Own Device (BYOD) [4].

Digitised data in healthcare is growing. Data is now processed simultaneously from internal and external sources, including mobile devices, wearable sensor devices, EPRs, Radiology Images, Videos, clinical notes, social media, blogs and remote health monitoring systems [5]. Terabytes of data that is generated from medical sensors is also used to increase the likelihood of reliable health diagnosis through accurate and detailed real-time data analysis [6]. However, this volume of data is growing beyond the capacity of health care infrastructures and is expected to increase further in the coming years [5]. The datasets produced are often unstructured, existing in formats which are isolated, disparate or incompatible [7]. There is often a lack of processing capabilities within healthcare networks to load and query the data effectively [5].

Additionally, the boundaries for healthcare systems are evolving, with many patients having the option of accessing their healthcare data from home PCs and mobile devices. This increases the attack surface significantly. Medical data must be private, with data misuse and violation detected in order to release and share data with authorised parties and public institutions [8]. A lack of security for healthcare devices leads to both a loss of patients' privacy and potential physical harm to the patient. There is also a risk that erroneous data is introduced or legitimate data is modified or suppressed by adversaries [9]. The security implications mean that bespoke systems need to be put in place to safeguard and protect data. However, the reliance on legacy software and bespoke systems results in an increased vulnerability to cyber-attacks [10]. The following successful hospital security breaches are testament to this:

- In May 2017, the WannaCry ransomware campaign exploited a Windows Server Message Block (SMB) vulnerability on TCP Port 445. SMB is a legacy protocol used to share files and printers over local networks. The exploit enabled the malware to use worm-like network propagation, encrypting files and demanding ransom payment, unless the system had been patched by Microsoft security bulletin MS17-010. The attack resulted in network downtime for 60 NHS hospitals [1], with 6 suffering disruption lasting several days.

- In October 2016, a UK Hospital in Lincolnshire was taken offline for four days due to a variant of the Globe2 ransomware [11]. All planned operations, outpatient appointments and diagnostic procedures were cancelled, with patients turned away and 2,800 patient appointments cancelled as a result of the disruption.
- In 2008, three London hospitals (St. Bartholomew's, the Royal London Hospital and the London Chest Hospital) lost all network connectivity due to several malware infections by the MyTob worm [10].

The UK government also invests heavily into cyber-security schemes, such as the £1.9billion investment into the national cyber security strategy [12], aiming to make the UK one of the safest places in the world to do business. Yet within healthcare infrastructures, privacy and security are still seen as a secondary consideration [13]; though the importance to establish data access regulations is imminent due to the geographical requirements for healthcare data being stored. Compliance with NHS guidelines, the Information Governance Toolkit, internal audit processes and information security standards (e.g. ISO27001 and ISO27002) is an additional concern to adhere to [14].

1.1. Background

The health sector receives consistently the highest number of reported data security incidents [15]; For example, in 2016, 450 data breaches occurred affecting more than 27 million patient records; 26.8% of these breaches were due to hacking and ransomware [15]. The remaining percentage of data breaches were due to non-cyber breaches relating to human error, such as posting/faxing/emailing personal data to the incorrect recipient. EPRs are valuable due to the wealth of sensitive and valuable detailed personal information held within, and the potential to commit identity fraud as a result. This data is often sold on the black market, where patient data can be profitable to illegal actors either through direct sale or extortion by ransom [16]. At the time of writing this thesis, patient privacy within EPR systems is enforced typically through corrective mechanisms, managed through role-based access [17]. However, once a user has been authenticated, they are essentially granted unhindered access. Additionally, employees present a persistent threat as they are able to access the data of almost any patient without control or reprimand. Without proactive monitoring of audit records, data breaches go undetected and employee behaviour is not deterred. With a requirement for all hospitals in the UK to be paperless by 2020 [18], access to healthcare data needs to be monitored proactively for malicious activity.

EPR systems are vulnerable to both insider and outsider threats [19]. 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 [13]. 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 Advanced Persistent Threat (APT) [20]. It is, therefore, a challenge to mitigate both types of threats. Confidentiality and patient privacy within EPR systems is typically managed through an agreed and signed code of practice between the organisation and its users. A healthcare organisation that collects, analyses, publishes or disseminates confidential patient data must commit to ensuring that the data is only accessed by relevant personnel and only when it is appropriate to do so [21]. However, in many cases, measures are not taken to detect and prevent patient privacy violations, any breaches of confidentiality are only brought to light once an investigation is launched, which is often too late. EPR systems are audited; however, the quantity of EPR audit data is significant and a challenge for regular analysis by an Information Security Analyst. Only a big data-capable solution is able to proactively monitor data for patient privacy violations.

To detect abnormal data behaviours, visualisation techniques provide both situational awareness and modelling capabilities for the benefit of computing in critical infrastructures. Situational awareness is defined in this thesis as extracting knowledge from existing information, enabling decision making by improving perception and comprehension of a situation within the EPR environment [22]. This allows an analyst to understand data correlations and identify anomalies for investigation through the shape or colour of data patterns [23]. Current procedural solutions to these issues are effective at detecting predictable insider threats [24]. They can process the large quantities of audit data, and can process procedures against that data. The three primary procedures processed against the data are 1) a staff member is looking at their own patient record, 2) a staff member is looking at a patient record with the same surname and 3) a staff member is looking at a patient record who lives in close proximity to their own address (taken from Human Resources 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 is 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 by cybercriminals.

1.2. Motivation

To address the lack of information about the rise in cyber-attacks on EPRs [25], this research investigates the growing concern of cyber-security for the critical health care infrastructure. The project examines the recent increase in attacks on the NHS in particular; and proposes a methodology for visualising real-world EPR audit logs in order to better understand cyber-attacks on EPRs. Therefore, we propose 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.

We present a framework for a novel anomaly detection system which integrates density-based outlier detection, human-in-the-loop machine learning and visualisation techniques to ensure patient privacy within EPR systems. The system visualises relationships between users and patients in a novel and interactive way. Outlier detection algorithms have the capability to explore complex datasets, detect hidden patterns and anomalies within them, and learn from analyst feedback. Additionally, they can identify when a user's behaviour has changed, by comparing behaviours such as the type of actions being taken and the patients they are viewing. HILML models employ active learning techniques to leverage human expertise and iterate training the machine learning model. Visualisation techniques are used to represent dense data, to augment the interpretation process. In this way, potentially illegitimate access to patient records can be highlighted and investigated. The results demonstrate the potential for data analysis and visualisation techniques to aid the situational awareness of patient privacy officers within healthcare infrastructures.

1.3. Aim and Objectives

The aim and objectives of the project focus on addressing the issue of internal and external threats to EPRs, through proactive monitoring of EPR audit logs.

1.3.1. Research Question

The following research questions are considered in the context of this research. How can inappropriate access of patient records be proactively detected without explicitly defining what constitutes inappropriate access? Once this has been achieved, how can this be ranked for prioritisation? And how can feedback be provided if the access is, in fact, appropriate?

1.3.2. Aim

The aim of this project is to develop a novel system framework to enable situational awareness of anomalous data behaviour within EPRs in order to secure patient privacy, particularly within the NHS critical infrastructure network.

1.3.3. Objectives

In order to fulfil the aim of this research, the following objectives are considered.

- Perform a literature review of healthcare infrastructures, including medical devices, hospital networks and hospital systems

- Review related work, including machine learning and visualisation techniques for outlier detection, in addition to related applications.

- Define and develop a novel framework for a system which concerns the use outlier detection algorithms, human-in-the-loop machine learning and visualisation techniques to detect potential patient privacy violations within an EPR.

- Validate the framework through using a real-world dataset and flag potential fraudulent access to patient medical records within the partner healthcare organisation.
- Disseminate findings and results to conferences and journals

1.3.4. Research Scope

The focus of this project is to design a novel framework for a machine learning system capable of identifying patterns and anomalies od data behaviour to identify potentially illegitimate accesses to patient records. The framework will combine a density-based outlier detection machine learning algorithm, human-in-the-loop machine learning, and visualisation techniques. The framework can be used in all healthcare infrastructures. The proposed system is limited through a focus on the NHS healthcare network and the use of a single real-world dataset, rather than multiple datasets.

1.4. Novelties

The research project has the following novel contributions:

- The development of a system framework that is able to analyse automatically EPR audit data and present it as a visualisation is novel. The system framework provides the operator valuable insights into the flow of data within the EPR using the machine learning algorithm Local Outlier Factor (LOF) [26].
- The system framework is bespoke to the healthcare infrastructure due to its use of a density-based clustering approach rather than following a procedure-based analytics approach. In doing so, the system framework can understand the unique characteristics of each user's activity rather than a one size fits all approach to appropriate and inappropriate access to EPR data. The use of machine learning algorithms enables hidden patterns of data to be detected which current procedure-based solutions cannot detect.
- The system framework flags up potential patient privacy violations for review to an analyst, and takes feedback from users to continually refine alerts. This aids in preventing alert fatigue. To our knowledge this is the only patient privacy detection

system that allows the user to interact with the system and provide feedback in order to tailor results to their specific EPR solution.

- Through combining existing research areas, (e.g. Cyber Security, Patient Privacy Analytics, Big Data Analysis, Visualisation Techniques and Machine Learning) the level of existing research is driven forward through the completion of this project.

The contributions and novelties of this work have been published in 6 conference papers (UKAIS 2016 [27], DeSE 2016 [28], IML 2017 [29], CLOUD COMPUTING 2018 [30], ICM 2019 [31], HEALTHINFO 2019 [32]) and 2 journal articles (MDPI Open Access 2019 [33] and IEEE Open Access 2019 [34]).

1.5. Research Methodology

The methodology firstly includes the assessment of real-world data sets collected from within a healthcare infrastructure. The analysis process has the goal of identifying how attacks are performed and modelling infrastructure behaviour in high detail.

The second phase includes the engagement of data visualisation techniques to envision both the generation and effects of data exfiltration on the EPR, and the changes occurring in a system when an attack has taken place. The output is used to educate and create awareness between stakeholders and communicate knowledge of existing threats in order to improve health care infrastructure security.

Thirdly, the development of a novel framework for a system capable of autonomously generating visualisations of cyber-attacks and critical infrastructure behaviours. As cyber-attacks on EPRs in the Healthcare Infrastructures are increasing [35], the research is timely and directly applicable to a real-world environment. Machine Learning algorithms will allow the system to identify patterns and trends in the data without being explicitly taught them. This is advantageous to Procedure-Based Analytics, which does not support learning, only deduction. 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 user’s 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. In the case that this activity is legitimate, such as a member of staff working an irregular shift, this activity can be recorded by the analyst using HILML.

The seriousness of critical information infrastructures' protection is a key issue. Their vulnerability to the growing cyber-threat enforces this further. Using the system framework to identify system threats within EPRs, a layer of security is added to the defence-in-depth approach currently in place [36]. Healthcare EPRs are the most integral component of the healthcare critical infrastructure and must be monitored and protected [37]. Additionally, it is well known that the dangers of cyber-crime increase exponentially with the number of interconnected computers and devices [38]. Therefore, the increasing reliance on EPRs for data capture and transmission is an increasing reliance on devices vulnerable to attacks from the cyber domain.

Health care data is an extremely attractive target for a cyber-criminal, the compromise of this data could lead to severe loss of patient privacy or the tampering and malicious falsifying of data could even lead to patient death. Therefore, unique analysis and reporting tools need to be developed to combat the increasing risk of data compromise and a focus on cyber security innovations is required to maintain patient trust in digital health care innovations. A novel framework to increase situational awareness for cyber security experts within health care infrastructures is proposed. The system framework creates a visualisation of data flow of the information systems used in the health care infrastructure. Through doing so, the operator can more accurately predict potential cyber vulnerabilities within its systems. The system framework furthers knowledge and understanding of EPRs and prevents data compromise from within them.

1.6. Thesis Structure

The structure of the thesis is as follows.

- Chapter 2 details hospital infrastructures, including hospital networks, medical systems and EPRs. The hospital networks section includes a topology detailing the network infrastructure. Hospital network security challenges are then discussed. An exploration of hospital network data using a real-world dataset is presented, demonstrating the complexity of hospital networks and dataflows. The medical systems section presents a topology of the number of medical systems within a hospital infrastructure and the exchange of patient data. This section also includes a detailed flow chart of patient journeys within the complex adaptive network of a hospital. Finally, EPR systems are discussed in detail, presenting related privacy challenges, the use of access control and its limitations, before discussing audit logs and their potential use in detecting confidentiality breaches.

- Chapter 3 details a literature review of related work and techniques. Visualisation techniques are presented as a tool for providing situational awareness of information security breaches. Machine learning is then discussed as an approach for anomaly detection within big data sets, including related algorithms. Finally, related work is discussed, in both academic and commercial fields.
- Chapter 4 demonstrates the detailed system design of the proposed framework. The approach is discussed including where the system is located within the medical systems topology. Then a detailed breakdown of the framework architecture is presented with each component of the framework including figures of the structure, functions, behaviours and interface with accompanying discussions.
- Chapter 5 presents a case study of real-world EPR data. The chapter discusses the dataset in detail, including the fields it is comprised of, how it was obtained and tokenised in order to protect patient privacy. The IDs within the dataset are then extracted, profiled and explored in order to determine anomalous data within the dataset.
- Chapter 6 presents the results of the system framework process on the EPR data. In the first section, the results are presented for one month of data and compares LOF against Density-Based Spatial Clustering of Applications with Noise (DBSCAN). In the second section six months of data is presented. In the third section eighteen months of data is presented. Finally, there is a discussion of the comparisons between each of the results.
- Chapter 7 presents the Thesis conclusion including the contribution to knowledge and a summary of the Thesis; the results and an evaluation of the framework, and a discussion on future work to be considered.

1.7. Summary

In this chapter, the background and motivation for this work are presented, outlining the incentive for the thesis. A research question, aim, objectives and research scope are detailed, contextualising the work within the wider research. The novelties of the work are described, along with the research methodology detailing how this will be achieved. Finally, an outline of the thesis structure is presented. In chapter 2, healthcare infrastructures are investigated in detail, describing hospital networks with a case study using real-world data, medical system interactivity and a focus on EPR systems.

2. *Healthcare Infrastructures*

2.1. Introduction

Hospital infrastructures are classified as mission-critical infrastructures [39], where damage to network communications and the loss of patient data would have a detrimental impact on the healthcare services they provide. In addition, mobile devices are being increasingly deployed within these networks, to support applications ranging from biomonitoring to materials handling and transportation [40]. Healthcare is an essential part of the national critical infrastructure network [41] and one of the four critical infrastructure groups (safety, mission, business and security). Damage to EPRs and the loss of patient data would have a detrimental impact on the health provision, potentially resulting i

n patient death or theft of sensitive data [42]. Yet, healthcare infrastructures are complacent towards the risks of patient privacy violations [43].

This project is timely due to *i)* a fundamental switch from paper to technology being used by beneficiaries within health care infrastructures; [44] *ii)* the increased need for 24-hour data access; *iii)* GPs increasingly using Virtual Private Networks (VPN) and 3G connections [45]; *iv)* Most UK hospitals have/are upgrading to the EPR system EMIS-web [46], *v)* more patient remote monitoring is taking security outside hospitals. Such trends reduce security levels and increase access to hospital networks and exposed Application Programming Interfaces (APIs).

2.2. Hospital Networks

In this section an overview of the hospital networks that host healthcare systems is presented. Figure 1 presents a 3-Layer Network Topology overview of how a typical hospital network is configured. This was provided by the network manager at a UK hospital. The diagram demonstrates the principle of the network configuration; as in reality it is replicated several times over and therefore too substantial to fit on a single diagram. The diagram shows the network layers involved between the server layer and the access layer, and demonstrates that they are duplicated throughout in order to provide network resilience. The servers connect to the server switch, providing the server access layer of the network topology, also known as the application layer. This then connects to the distribution layer, also known the transport layer. The transport layer connects to the core layer, or the network layer, providing access to routing, wireless and the firewall. This then again connects to another distribution/transport layer. This connects to the access layer where networked devices connect.

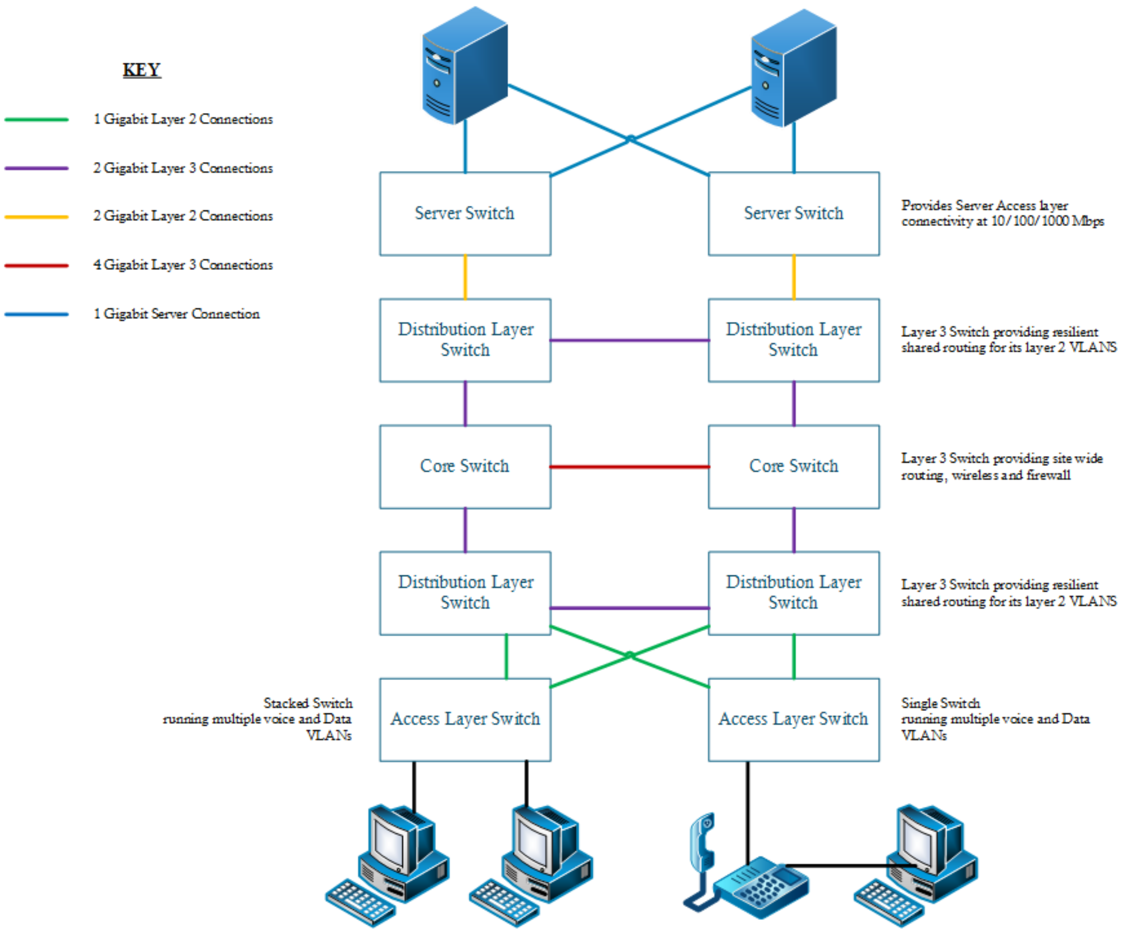

KEY
1 Gigabit Layer 2 Connections
2 Gigabit Layer 3 Connections
2 Gigabit Layer 2 Connections
4 Gigabit Layer 3 Connections
1 Gigabit Server Connection
Server Switch
Server Switch
Provides Server Access layer connectivity at 10/100/1000 Mbps
Distribution Layer Switch
Distribution Layer Switch
Layer 3 Switch providing resilient shared routing for its layer 2 VLANS
Core Switch
Core Switch
Layer 3 Switch providing site wide routing, wireless and firewall
Distribution Layer Switch
Distribution Layer Switch
Layer 3 Switch providing resilient shared routing for its layer 2 VLANS
Stacked Switch running multiple voice and Data VLANs
Access Layer Switch
Access Layer Switch
Single Switch running multiple voice and Data VLANs

Figure 1. 3-Layer Network Topology

Further to this, Figure 2 presents an overview of a typical healthcare network infrastructure for enabling remote access within a hospital. The layout enables staff the ability to work and provide on-call services remotely. Figure 2 demonstrates the locations of the LANs and VLANs within a typical hospital network infrastructure. This was also provided by the network manager at a UK hospital. In addition, the firewall placements, in relation to the Internet, are depicted.

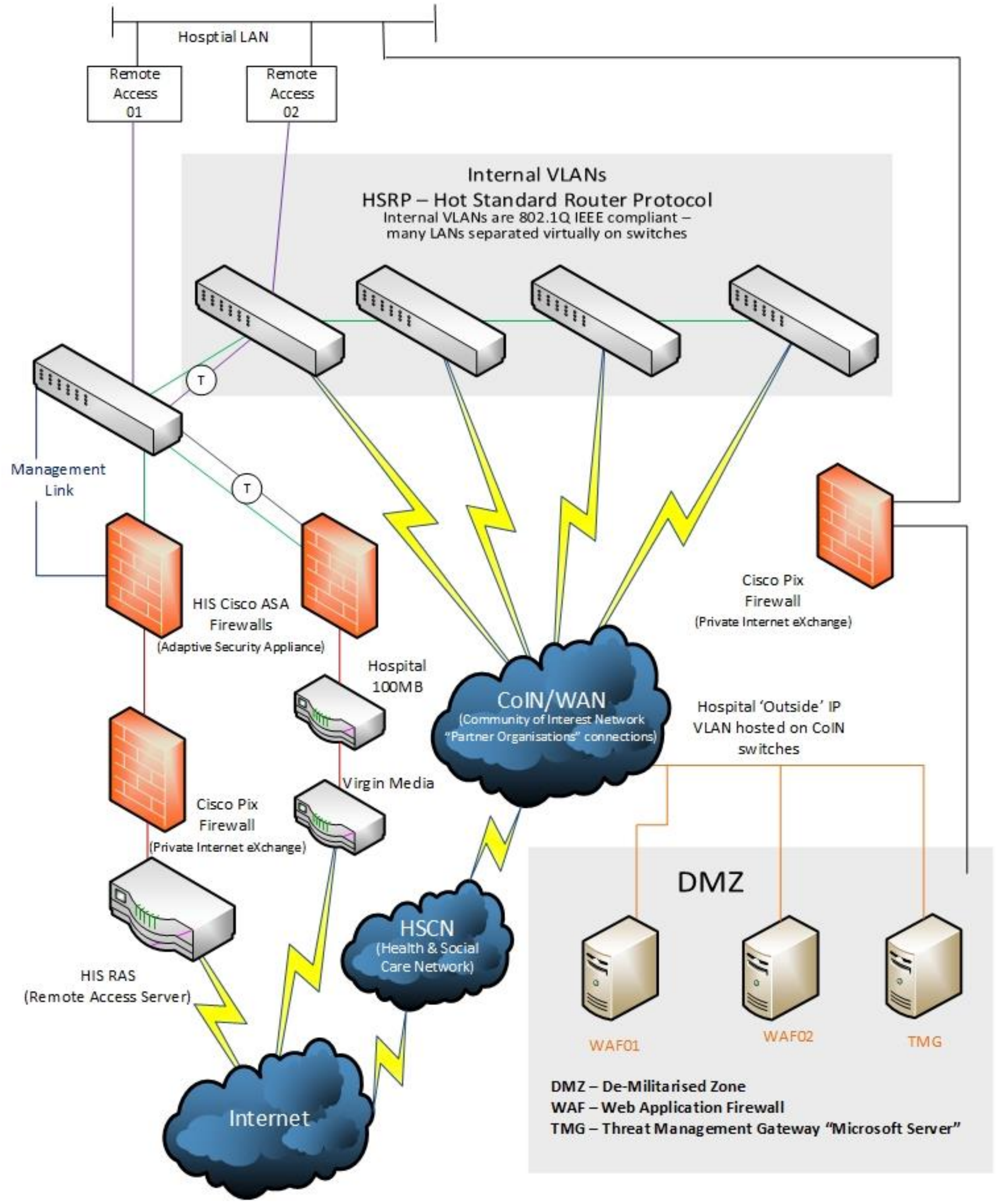

Hosptial LAN
Remote Access 01
Remote Access 02
Internal VLANs
HSRP – Hot Standard Router Protocol
Internal VLANs are 802.1Q IEEE compliant – many LANs separated virtually on switches
T
T
Management Link
HIS Cisco ASA Firewalls
(Adaptive Security Appliance)
Hospital 100MB
Virgin Media
Cisco Pix Firewall
(Private Internet eXchange)
HIS RAS
(Remote Access Server)
Internet
HSCN
(Health & Social Care Network)
CoIN/WAN
(Community of Interest Network "Partner Organisations" connections)
Cisco Pix Firewall
(Private Internet eXchange)
Hospital 'Outside' IP VLAN hosted on CoIN switches
DMZ
WAF01
WAF02
TMG
DMZ – De-Militarised Zone
WAF – Web Application Firewall
TMG – Threat Management Gateway "Microsoft Server"

Figure 2. Network Infrastructure for Remote Access

Figure 2 also displays the relationship between a typical hospital's typical 'Community of Interest Network' (CoIN)/WAN and the HSCN (a WAN used to connect many sites across the UK National Health Service). The system layout leaves a vulnerability to attackers being able to eavesdrop on traffic. From there messages can be injected, replay attacks can be performed, such as maliciously replaying or delaying a valid data transmission, and spoof messages can be generated [10]. In doing so, it is possible to compromise the integrity of the device operation [9]. If successful, patient privacy would be invaded and legitimate data supressed. This compromises patient privacy whilst attempting not to interfere with medical device operation.

With healthcare organisations using electronic records, cyber-based transactions (such as ordering diagnostic tests and e-prescribing) and mobile electronics, the risk of a data breach is an increasing concern. Healthcare data is intrinsically valuable; the repercussions of data compromise within healthcare infrastructures can range from loss of patient privacy and fraud, to patient injury or potentially death. Therefore, protecting intrinsically private patient data and preventing data compromise is critically important. Visualisation tools can be used by cyber security officers within healthcare organisations to increase their situational awareness of data flow and actively address this issue. Additionally, visualisation tools allow system operators to be proactive about cyber security within healthcare organisations, through highlighting unusual activity for investigation. This is in contrast to the accepted and fundamentally flawed approach of reactivity to cyber security attacks, which does not attempt to address the underlying security flaws within healthcare organisations [47].

With healthcare networks, devices (medical, clinical and personal) are connected to global networks for convenient access [48]. However, modern healthcare networks are complex systems, with hospitals each having their own unique structure [49]. Healthcare organisations differ from other enterprise networks through their use of Medical Cyber Physical Systems (MCPSs). MCPSs are inexpensive personal monitoring devices that can record and transmit multiple physiological signals [50]. Encryption of this data is required for secure storage, secure transmission, and secure computation. MCPSs consist of four layers, which need to be considered and secured. Firstly, the Data Acquisition layer consists of a Body Area Network (BAN), which are wearable sensors that facilitate the collection of patient medical information. Secondly, the Data Concentration/Aggregation Layer, consisting of transmitting the gathered information to a gateway server through short range wireless, such as Bluetooth, due to the low computational power of sensors with a BAN. Thirdly, the Cloud Processing and Storage Layer consists of the long-term secure storage, processing and analytics of medical information. Finally, the Action Layer consists of either active or passive usage of the data. Active usage consists of an actuator using the data and the algorithms used to perform data analytics to be directly influenced by the data, such as through the use of a robotic arm in robot-assisted surgery. Passive action visualises the data in order to provide decision support to medical professionals [50]. Intrusion Detection System (IDS) techniques for MCPSs are still in their infancy [51]. Attacks are detected through recording state information for both local nodes and peers. This information is then updated to a state machine, which models the subject device. This state machine generates a detection when state information becomes malicious [51]. In practice, this means that if a heart rate component is reporting normal cardiac function, but the Cardiac Device (CD) is in defibrillator mode, then the IDS should report a detection instead. Wireless communication security is handled by contemporary secret key technology, such as Public Key Infrastructure (PKI). This provides authentication and prevents man-in-the-middle attacks [14].

2.2.1. Hospital Networks Security Challenges

With increasing requirements for valuable and accurate information, patients need to be confident that their data is being stored safely and securely. Procedures are employed 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. However, it is impractical to set procedures for every single user and patient with an EPR [49] as 1) Information Security teams are typically under resourced and 2) even if the resource is available to do this, it often would not provide meaningful information due to the unpredictable nature of healthcare workflows. Procedure-based solutions cannot detect violations (such as APTs) in these contexts as criteria which constitutes appropriate behaviour cannot be comprehensively defined pre-emptively [4]. Additionally, it is unfeasible to detect fully all illegitimate access within EPR systems [49], 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.

Healthcare organisations have sensitive data spread across a number of devices. Mobile devices, such as laptops or tablets are all used for inputting medical data. Furthermore, specialised devices, such as the Draeger patient monitoring systems [52], require network connectivity. Draeger systems, are medical devices for use in patient bedside care, which operate using a custom OS Shell running Windows XP/7. The system is currently installed in a number of hospitals throughout the UK [52].

Figure 3 - Draeger Technology

This custom Windows shell is known as the Infinity Explorer [52]. For system security, the Draeger Omega system (see Figure 3) uses Infinity Explorer Security, which functions as both Virus and Intrusion Protection, in addition to a Firewall. At present the majority of hospital systems are outward facing, meaning that they are resistant to external attack sources, yet vulnerable to insider attacks. Draeger medical devices, including the Omega system, employ the use of a touch screen interface and customisable user interface for designed patient/diagnosis specific layouts. The intention of Draeger devices, as with most medical devices, is to provide the highest quality of patient care, by providing Data Accessibility, Integrity and Security. The Draeger Omega system runs Windows XP, which is inherently vulnerable since it is no longer supported by Microsoft and its vulnerabilities are well known to attackers.

This risk is further exacerbated by the BYOD revolution. This is a term referring to the technologies enabling employees to access and utilise internal corporate IT resources, with their personal devices [53]. BYOD policies have numerous benefits including reduced costs and improved productivity, convenience and efficiency of work. The practice also means that users from visiting healthcare organisations have access to hospital networks and the Internet. However, BYOD also carries numerous risks including data loss/leakage or theft, application security, network availability, legal liability and regulatory compliance and loss of brand identity, posing various challenges for IT departments who support and secure them [53]. BYOD is a large contributor to the era of ubiquitous computing, also known as pervasive computing, where various technologies interact with one another, with potentially sensitive data being transmitted over less secure means of communication, such as wirelessly [54]. With an increased use of technology there is a corresponding risk of increased exposure to cyber security threats. These wireless vulnerabilities pose a particular risk to the healthcare network, in that potentially insecure devices are granted access to hospital infrastructure and confidential data. An attacker is able to use this to their advantage by hacking a BYOD in order to gain back-door access onto a hospital network.

With regards to the aforementioned tendency for organisational complacency towards the risks of cyber security [43], issues of reduced information visibility due to data complexity, fragmentation, interoperability and lack of specialisation, all undermine the security of these organizations [43]. Organisations need to bridge the gap between cyber operations, resilience and the priorities of the business. In addition to this, the decision makers need to be able to synthesize highly disparate data into a coherent and concise narrative [43]. The goal of security engineers is to develop tools capable of detecting malicious, multistage intrusion attacks, weighting the individual attacks, and comparing them against the universe of attacks within the network [10]. This is a plain recognition problem and an intruder's objectives should be determined based on the analysis of the entire dataset of attacks, rather than just individual attacks [10]. Healthcare network security challenges can be summarised into four categories [55]: 1) system structure, 2) mobile device, 3) medical equipment and 4) user-based challenges. Where, system structure refers to the vulnerabilities in the physical system layout; mobile device challenges relate to issues surrounding on-demand network access; medical equipment challenges concern low security in medical and biomedical devices. Finally, user-based challenges refer to the general staff-awareness of existing cyber-threats or good cyber-security practices. For example, a hospital staff member may click unsuspectingly on a link in an email, which contains a Trojan virus and creates a back door into the hospital network.

Recent attacks, such as the WannaCry campaign [29], have further reduced the levels of public trust in security leading to widespread concern about the health sector's ability to maintain the privacy of patient data. Bell-LaPadula [56], and FairWarning [57], are the staple access control systems employed but are i) inflexible, presenting issues when considering the dynamic boundaries of many modern healthcare networks and ii) do not consider an attacker who has acquired the logon credentials of an approved clinician (e.g. through phishing or social engineering) [58]. This has been a challenge for security experts for many years; referred to as a plain recognition problem [10], Information Security Officers and IT Managers need to interpret disparate data behaviours to preserve privacy and safeguard EPR data [7]. They constantly balance privacy with a need for more intuitive security solutions. Therefore, confidentiality and patient privacy within EPR systems is typically managed through an agreed and signed code of practice between the organisation and its users [21].

2.2.1.1. Healthcare Data Breaches

Healthcare regularly appears in the top three industries for data breaches [4]. Traditional approaches to endpoint security are no longer viable in the modern cyber-threat landscape [4]. Attack models have changed from attacks on single PCs to large scale attacks on entities through APTs. For example, zero day exploits, spear phishing, watering hole models and encrypted side channel methods are being used increasingly to infect critical systems [14]. In addition, modern malwares have adopted and evolved Evasion Techniques, such as malware packing, obfuscation and polymorphism [4]. As such, the discussion in this section is focused on the specific threats and attack vectors facing healthcare infrastructures.

Information Security is based on three key concepts: confidentiality, integrity and availability [59]. Confidentiality and integrity ensure that only an authorised user can access and edit protected data [59]. Confidentiality ensures the inaccessibility of private medical information to unauthorised users [6]. BYODs are vulnerable due to the lack of confidentiality, isolation and compliance on BYODs [60]. Confidentiality refers to sensitive data being stored on a personal device and the difficulty in monitoring unauthorized and illegal access to that data.

Within healthcare infrastructures, MCPSs are personal monitoring devices that can record and transmit multiple physiological signals [50]. For safety-critical MCPSs the ability to detect attackers, whilst limiting false alarms in order to protect the well-being of patients, is of critical importance [51]. However, threats to MCPS components are increasing with the malicious users aiming to cause node compromise [51]. This process can be initiated through over-the-air software updates, stack overflow exploits or 'logic bombs' through third party developers [51]. Security is a concern especially for small medical devices attached to a patient [61]. Compromise of data storage could potentially result in patient death. Similarly, attacks on pharmacy systems could result in the wrong medication being prescribed, due to compromise of patient information leading healthcare providers to make decisions based on incorrect data, leading to long term health concerns for the patient [61]. MCPSs need to prevent the disclosure of information to unauthorised individuals [61]. Particularly, in regards to healthcare, patient personal data needs to be transmitted confidentially though the use of encryption techniques [61]. Additionally, information generated from medical devices is required to only be accessible to authorized users [9].

The most frequent outcome of a cyber-attack on a system is the unavailability of patient care due to computer outages [62]. Other common attacks, which are a challenge to healthcare security systems, are outlined as follows:

- Scanning attacks involve adversaries gathering meaningful information in order to launch a sophisticated attack upon an infrastructure [63]. These scans commonly include, IP address scanning, port scanning and version scanning. With regards to Healthcare Infrastructures, an adversary can carry out segment scanning on Health Level Seven (HL7) information, which is the standard framework for the exchange and integration of healthcare data [64]. In doing so, they can learn personal identifiers, order numbers or patient visit information [63].
- Spoofing attacks involve malicious users masquerading as legitimate [63]. Masquerading is a passive spoofing attack where attackers acquire legitimate

account credentials and then log in. Impersonation is an active spoofing attack, sometimes known as a replay attack, wherein attackers capture authentication traffic and replay the traffic in order to gain access to the healthcare infrastructure.

- Injection attacks involve exploiting vulnerabilities of SQL, JavaScript and other computer programs in order to successfully insert untrusted data [63]. In doing so, attackers may gain access to healthcare databases, attack web users and propagate viruses. Additionally, they may inject malicious segments commands or responses in order to reduce the security of healthcare infrastructures.
- Broken Authentication and Session Management involves attackers exploiting vulnerabilities in authentication mechanisms in order to assume the identities of legitimate users [63]. A brute force attack is an example of this kind of attack, taking advantage of weak passwords and small encryption keys, ultimately allowing a malicious user to perform all the functions available to a legitimate user.
- DDoS involve exhausting system and network resources in order to make them unavailable [63]. For example, a flooding-based DDoS sends a large number of packets to a web server; in doing so, legitimate requests are blocked as the server CPU is overwhelmed.

Healthcare organisations should have processes in place to identify data loss and wipe data remotely. The challenge is that the volume of data and number of connected network devices makes detecting cyber-intrusions a considerable difficulty. Security threats within healthcare have farther reaching implications than other industries. Due to the increased adoption of trends such as the use of mobile, social and cloud, the ease for hackers to infiltrate the healthcare infrastructure massively increases [4]. Medical data must be private, with data misuse and violation detected in order for authorised users to release and share data to authorised parties and public institutions [8].

2.2.2. Hospital Network Data

In this sub-section, further investigations into hospital networks are presented through the visualisation of real-world netstat hospital data on three servers. This serves as a background demonstration of the data within a hospital network setting. Through these investigations further insights into hospital networks are gained. The investigation highlights the complexity of the network data and the structure of network packets from the medical systems within the hospital infrastructure.

Transmission Control Protocol (TCP) socket connections to three different servers offering an Electronic Prescribing and Medications Administration (EPMA) System, a Patient Administration System (PAS) and an Active Directory Domain Controller (AD). Data from a Liverpool-based hospital is analysed. TCP data was captured as it allows for two hosts to establish a connection and exchange data streams. Specifically, three visualisation techniques were investigated, 1) Force-Directed Visualisation algorithms, 2) Logarithmic Heatmaps and 3) Nonparametric Statistical Graphics to present network data in such a way as to highlight anomalies and identify relationships between data points.

This TCP data employed in the visualisations was captured using the netstat command with the command line utility. In addition to running a netstat command [65] a number of additional parameters were included in the command (-nab). These parameters are defined as follows:

- The netstat command (without any additional parameters) displays the Protocol, Local Address, Foreign Address and State of the TCP connections.
- The netstat–n command displays active TCP connections numerically and no attempt is made to determine names, in order to facilitate the dataset analysis.
- The netstat–a command displays all active connections of the TCP and UDP ports, on which the computer is listening.
- The netstat–b command displays the executable program name associated with the creation of the connection or listening port.

For each of the three datasets, the four separate netstat commands were executed. Firstly, without any parameters, secondly with the parameters –nab, thirdly with the parameter –b, and finally with the parameter –n. The first command returns a dataset which attempts to determine the names of the Foreign Address, so that the devices which the servers are connected to are known. The –nab command displays all TCP connections numerically alongside the executable involved in creating the connection. The –b command displays TCP connections with named Foreign Address values. Finally, the –n command produces a clean dataset. Specifically, the netstat commands are executed on these servers:

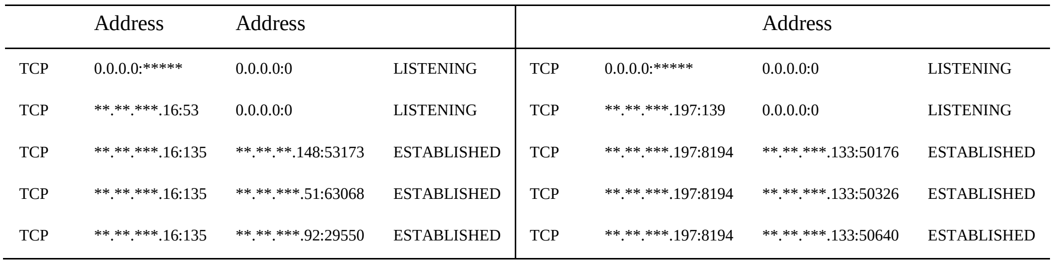
- The AD server, which manages directory-based services for the hospital.
- The PAS server, which manages core functionality, such as patient administration, across the hospital.
- Finally, the EPMA server, which generates, transmits and files prescriptions across the hospital.

Additionally, for each dataset, the most recurring items for both local and foreign addresses are shown. The Liverpool-based hospital has currently 274 servers, a combination of both physical and virtual servers, providing specialist applications and functions across the hospital network. Of these, there are 5 Domain Controller servers, 4 Patient Administration System servers and 2 Electronic Prescribing servers. The three server types used for the purpose of this are chosen for two reasons; *1)* they are the most active servers on the network, and *2)* there potential value to a malicious attacker. For example, if the attacker is able to infiltrate the Active Directory Domain Controller, then they have access to an authentication certificate and access to the wider hospital network. By accessing the Electronic Prescribing System, the attacker has access to large quantities of pharmaceutical drugs. Additionally, by accessing the Patient Administration System, intrinsically valuable and confidential medical data would be reachable. If an attacker were to, instead, attempt to shut down any of these servers through a DDoS attack, it could limit the ability of legitimate medical professionals to provide appropriate patient care and could potentially lead to patient harm.

In Table 1, a sample of the analysed netstat data is shown displaying *(i)* the connection type, *(ii)* the IP source connecting to the DC, *(iii)* the target of the IP address (the DC server), and *(iv)* the state of the connection. All data presented is anonymised. The data is a single snapshot of the domain controller server and comprises of 590 established connections of 5,688 total ports.

Table 1. TCP Socket Connections Sample Data (Anonymised)

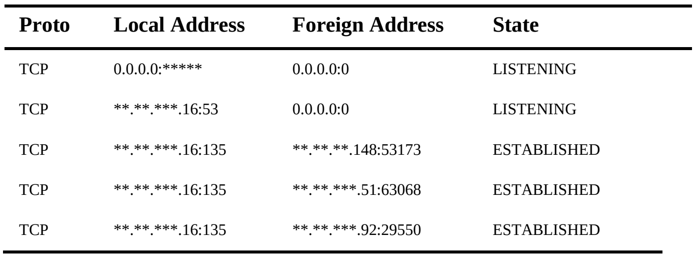
Active Directory Domain Controller				Electronic Prescribing System			
Proto	Local	Foreign	State	Proto	Local Address	Foreign	State



	Address	Address				Address	
TCP	0.0.0.0:*****	0.0.0.0:0	LISTENING	TCP	0.0.0.0:*****	0.0.0.0:0	LISTENING
TCP	**.**.***.16:53	0.0.0.0:0	LISTENING	TCP	**.**.***.197:139	0.0.0.0:0	LISTENING
TCP	**.**.***.16:135	**.**.**.148:53173	ESTABLISHED	TCP	**.**.***.197:8194	**.**.***.133:50176	ESTABLISHED
TCP	**.**.***.16:135	**.**.***.51:63068	ESTABLISHED	TCP	**.**.***.197:8194	**.**.***.133:50326	ESTABLISHED
TCP	**.**.***.16:135	**.**.***.92:29550	ESTABLISHED	TCP	**.**.***.197:8194	**.**.***.133:50640	ESTABLISHED

In Table 2 a sample of the netstat data displays, the connection type, the IP source connecting to the PAS, the target of the IP address (the PAS server) and the connection state. The data is a single snapshot of the domain controller server and comprises 93 established connections of 173 total ports.

Table 2. Patient Administration System – TCP Socket Connections Sample Data (Anonymised)



Proto	Local Address	Foreign Address	State
TCP	0.0.0.0:*****	0.0.0.0:0	LISTENING
TCP	**.**.***.16:53	0.0.0.0:0	LISTENING
TCP	**.**.***.16:135	**.**.**.148:53173	ESTABLISHED
TCP	**.**.***.16:135	**.**.***.51:63068	ESTABLISHED
TCP	**.**.***.16:135	**.**.***.92:29550	ESTABLISHED

In Table 3 a sample of analysed netstat data is shown displaying firstly, the connection type, secondly the local IP address (the AD server), thirdly the local port number, fourthly the foreign address connecting to AD, fifthly the foreign port number, sixthly the state of the connection and finally the process running on the connection. All data presented is anonymised. The data is a single snapshot of the domain controller server and comprises of 590 established connections of 5,688 total ports. Netstat tables without parameters for the AD, EPMA and PAS servers are presented in the Appendix in Table 40, Table 41 and Table 42 respectively.

Table 3: Active Directory Dataset Sample

Proto	**Local Address**	**Local Port**	**Foreign Address**	**Foreign Port**	**State**	**Process**
TCP	0.0.0.0	49288	0.0.0.0	0	LISTENING	dns.exe
TCP	0.0.0.0	49293	0.0.0.0	0	LISTENING	services.exe
TCP	0.0.0.0	49331	0.0.0.0	0	LISTENING	PolicyAgent
TCP	**.***.***.16	53	0.0.0.0	0	LISTENING	dns.exe
TCP	**.***.***.16	135	**.**.**.148	53173	ESTABLISHED	RpcSs
TCP	**.***.***.16	135	**.**.***.51	63068	ESTABLISHED	RpcSs
TCP	**.***.***.16	135	**.**.***.81	62264	ESTABLISHED	RpcSs
TCP	**.***.***.16	135	**.**.***.92	29550	ESTABLISHED	RpcSs
TCP	**.***.***.16	135	**.**.***.135	55335	ESTABLISHED	RpcSs

TCP	**.***.***.16	135	**.**.***.143	50150	ESTABLISHED	RpcSs
TCP	**.***.***.16	135	**.**.***.158	58659	ESTABLISHED	RpcSs

In order to provide an appropriate visualisation, the dataset undertakes a pre-processing phase. To do this, node and edge data are extracted and isolated into individual datasets. Three datasets are then produced for each data capture procedure. Examples of the node dataset and the two edge datasets are found in Table 4.

In Table 4, the Source column refers to a common ID number indicating the TCP connection or listening port. Table 4(a) contains the port process running on the IP connection. Table 4(b) contains the local port number and Table 4(c) contains the foreign port number. In order to provide a proof of concept, the data is manually cleansed of low-risk data. Unknown port processes, or processes running on unfamiliar ports, are left in the dataset, whilst common processes, running on secure and known port mappings are removed.

Table 4: AD Data Preparation – (a) Port Process – (b) Local Port – (c) Foreign Port

Source	Label		Proto	Local Address		Proto	Local Address	
24	dns.exe		24	49288		24	0	
25	services.exe		25	49293		25	0	
26	PolicyAgent		26	49331		26	0	
27	dns.exe		27	53		27	0	
28	RpcSs		28	135		28	53173	
29	RpcSs		29	135		29	63068	
30	RpcSs		30	135		30	62264	
31	RpcSs		31	135		31	29550	
32	RpcSs		32	135		32	55335	
33	RpcSs		33	135		33	50150	
34	RpcSs	(a)	34	135	(b)	34	58659	(c)

2.2.2.1. Network Visualisation Algorithms

ForceAtlas2 [66], Yifan Fu [67] and Fruchterman-Reingold [68], are force directed placement algorithms allowing the display of complex graph structures. This is achieved through sorting and placing nodes into structured topologies satisfying several visual requirements such as even distribution, symmetry, non-overlapping edges and minimising distance between close nodes. ForceAtlas2, for example, is a force directed layout algorithm, simulating a physical system in order to spatialise a network [66]. Nodes repulse one another and edges attract nodes, which ultimately creates a balanced state and allows for a visual representation of data structure. This enables data communities with denser relations to appear as groups of nodes. Yifan Fu is a graph drawing algorithm combining a multilevel approach, to overcome local minimums, and the Barnes-Hut octree technique, to approximate short and long range forces [67]. Fruchterman-Reingold is an algorithm intended to draw undirected graphs through distributing vertices evenly, making edge lengths uniform and reflecting symmetry [68]. Furthermore, two principles are adhered to. Firstly, that vertices connected by an edge should be drawn near each other, and secondly, that vertices should not be drawn too close to each other [68]. In addition to these criteria, the Fruchterman-Reingold algorithm adds even vertex distribution and treats vertices as atomic particles or celestial bodies, exerting attractive and repulsive force from one another [69]. In other words, the algorithm distributes vertices evenly, makes edge lengths uniform and reflects symmetry by placing spring-like attractive forces on each edge, and letting the system stabilise. The algorithm is considered a standard in graph-drawing [64]. Employing an iterative approach for determining the position of all the nodes and the distance between the nodes connected by an edge, using the sum of force vectors to calculate the direction and distance a node should be moved at each step, and updating the forces between them, until stability (i.e. minimum energy) is reached.

The optimal distance *k* between two vertices is defined as [68]:

$$k = C\sqrt{\frac{area}{numberof\ vertices}}$$

(2.1)

where *C* is the constant to be found experimentally and represents the width step a node is moved at each iteration [68].

The Fruchterman-Reingold attractive force is defined as:

$$f_a(d) = d^2/k$$

(2.2)

where f_a is the attractive force, d is the distance between two vertices and k is the radius of the empty area around a vertex [68].

The Fruchterman-Reingold repulsive force is defined as:

$$f_r(d) = -k^2/d \tag{2.3}$$

where f_r is the attractive force, d is the distance between two vertices and k is the radius of the empty area around a vertex [68].

OpenOrd is an open-source graph-drawing algorithm specialised for drawing large-scale real-world graphs incorporating edge-cutting, a multi-level approach, average-link clustering and a parallel implementation of a force-directed method [70]. It is based upon and implementation of the Fruchterman-Reingold algorithm known as the VxOrd [71]. All of these visualisation algorithms are employed and evaluated as part of the data processing case study in section 2.2.2.2.

Yifan Hu, which is a force-directed graph drawing algorithm, is used to model the network data through a system of *bodies*, with forces acting between them [67]. It is the first algorithm to combine both techniques for large scale graph drawing [67]. Typically, this multilevel approach has three phases, as described in Algorithm 1:

- Coarsening: In this phase, a series of graphs are generated, with the aim of encapsulating the information of its parent, while containing fewer vertices and edges. The process continues until the coarsest graph layout is determined.
- Coarsest graph layout: In this stage, the graph is then presented using an algorithmic technique that combines an adaptive step length control scheme.
- Prolongation and Refinement: The layout on the coarsest graphs are then prolonged recursively to the finer graphs. Once this has been carried out, the layout is then refined again using the algorithm used in phase 2. This is an iterative process.

The algorithm uses the repulsive forces on one *node* from a cluster of distant *nodes.* The algorithm calculates both the attraction and repulsions forces to visually demonstrate the component nodes within a hospital network structure. The Yifan Hu repulsion F_r formula is mathematically defined as [67]:

$$F_r = \frac{k}{d^\wedge 2}$$

(2.4)

Here, *d* represents the distance between the two nodes, while the attraction F_a formula is expressed as [67]:

$$F_a = -k \cdot d$$

(2.5)

Algorithm 2.1 – Yifan Hu Multilevel algorithm [67]

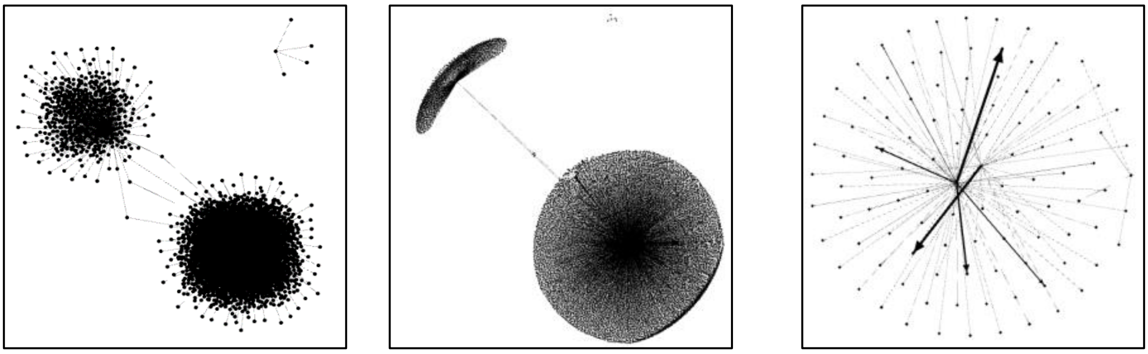
```
1) Coarsest Graph Layout, which is as modelled as follows:
    $if\ (n^{(i+1)} < MinSize\ or\ n^{(i+1)}/n^{i}\ > p)\{$
        $*\ x^{i} := random\ initial\ layout$
        $*\ x^{i} = ForceDirectedAlgorithm(G^{i}, x^{i}, tol)$
        $*\ return\ x^{i}\}$
2) The Coarsening Phase, calculated as outlined:
    $set\ up\ the\ n^{i} \times n^{(i+1)}\ prolongation\ matrix\ P^{i}$
    $G^{(i+1)} = P^{(i^{T})} G^{i}\ P^{i}$
    $x^{(i+1)} = MultilevelLayoutG^{(i+1)}, tol)$
and 3) the Prolongation and Refinement Phase, where prolongation is employed to acquire initial layout:
    $x^{i} = P^{i}\ x^{(i+1)}$
    $refinement: x^{i} = ForceDirectedAlgorithm(G^{i}, x^{i}, tol)$
    $return\ x^{i}$
```

The starting point is the original graph, $G0 = G$ and $ni = |\,Vi\,|$ is the coefficient for the number of vertices in the ith level graph, represented as Gi. xi is defined as the coordinate vector for the vertices in Vi. Gi is represented by a symmetric matrix Gi, where all entries of the matrix act as the edge weights. $Gi + 1$ to Gi is the continuation operator, also represented by a matrix Pi, of dimension $ni * ni + 1$.

2.2.2.2. Network Visualisation Examples and Techniques

Figure 4 depicts visualisations of the raw data connections at the Liverpool hospital. The nodes, depicted by black circles, represent devices connecting to the domain controller.

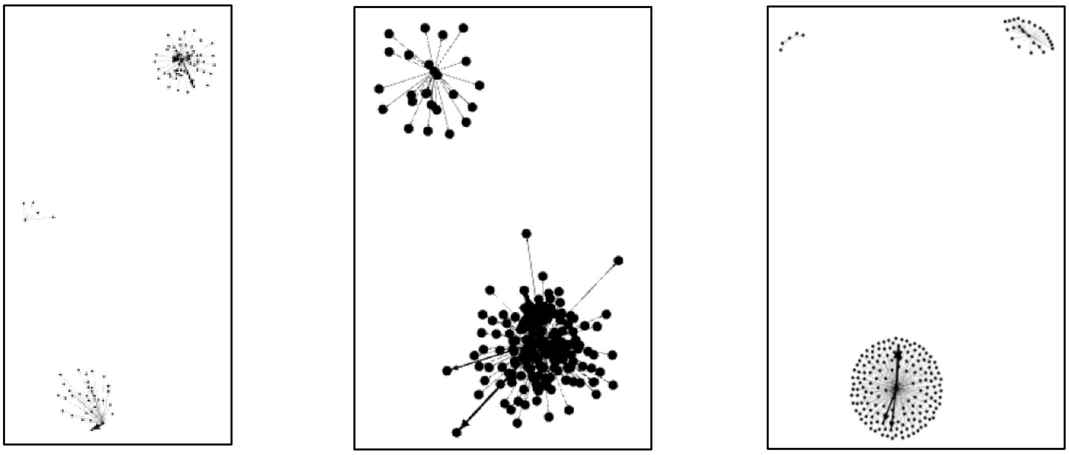
Visualisations of the hospital′s Act*ive Directory Domain Controller*, are presented in Figure 4 (a) and (b). In Figure 4 (a), the algorithm performed on the data is a ForceAtlas2 [66] [72]. The fundamentals of ForceAtlas2 are intended to be simplistic, nodes repulse and edges attract continuously*,* while the layout is running and can be manipulated by the user while running. Figure 4(b) displays the algorithm applied to the data using the Yifan Hu Multilevel layout [67]. *In this case, t*he repulsive forces on a node from a cluster of nodes are approximated by a Barnes-Hut calculation [72], which treats them as a super-node. The repulsive force is global and proportional to the inverse of the (physical) distance between vertices. The attractive force (the spring force) is only between neighbouring vertices [67]. The effect of this is that one node may be attracted to two other nodes, but those nodes may be repelling each other. This leads to a stretching effect. As the graph shows, visualising the raw data provides a challenging insight into security detection. The identification of anomalous behaviour for one device is arduous due to the volume of data present. Additionally, there is no clear differentiation between low-risk data, and medium to high-risk data, which may indicate a potentially malicious device. Therefore, the visualisation of the data without being pre-filtered provides a limited insight into network behaviour; as the data is dense, with no clear indication of anomalous data points to be highlighted for security analysts.



(a) Domain Controller Data with ForceAtlas2 algorithm

(b) Domain Controller Data Yifan Hu Multilevel layout algorithm

(c) Electronic Prescribing Data with Frutcherman-Reingold layout algorithm



(d) Electronic Prescribing Data with Yifan Hu Multilevel layout

(e) Patient Administration System Data with OpenOrd layout algorithm

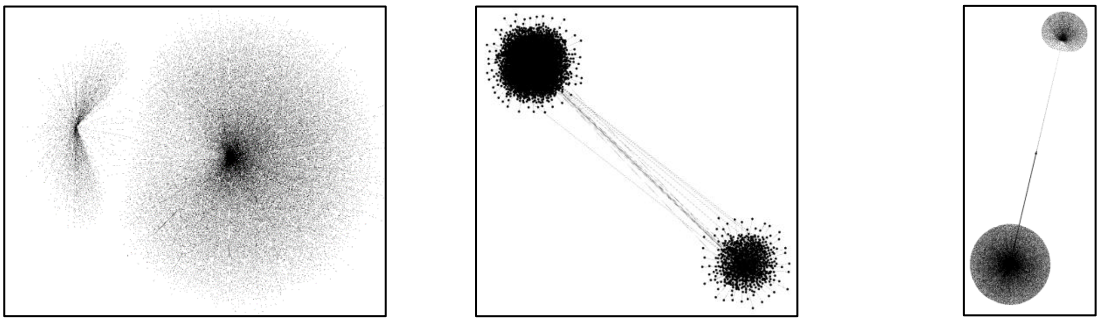
(f) Patient Administration System Data with Yifan Fu Multilevel layout algorithm

Figure 4. Netstat Visualisations

Figure 4 (c) and (d) depict visualisations of data connections for the EP system at the Liverpool hospital. In Figure 4 (c) the data is processed through the Frutcherman-Reingold layout algorithm, which simulates the graph as a system of mass particles [68]. The nodes are mass particles and the edges are springs between the particles. The arrows on the visualisation represent weighted edges, where there are two or more edges between two vertices, with the larger arrows representing more edges. In Figure 4(d) the algorithm applied to the data is the Yifan Hu Multilevel layout [67]. Figure 4 (c) and (d) demonstrates that, at a glance, current industry standard visualisation algorithms cannot present the data in a meaningful way without first pre-processing the data. Advanced data analytics techniques would, therefore, allow the data to be filtered in such a way that the resulting visualisations would highlight the potential security risks more clearly in the data [73].

Figure 4 (e) and (f) presents visualisations of data connections for the PAS system at the Liverpool hospital. Figure 4 (e) displays the PAS data visualised through the 'OpenOrd layout' algorithm, which expects undirected weighted graphs and aims to distinguish clusters of data [70]. It is based on the Frutcherman-Reingold [68] algorithm. In Figure 4 (f), the Yifan Fu layout algorithm is applied. In Figure 4 (e) and (f), even with a much smaller dataset, it is shown the data is still unmanageable. It is unreasonable to expect the user to identify which operators are potentially high-threat, what the data points tell the user about the overall state of the system, and the anomalies within it.

In order to demonstrate the interconnectivity of the three datasets, the data is combined into a single dataset and visualised. Figure 5 presents visualisations of data connections for the DC, EP and PAS systems at the Liverpool hospital.



(a) Full Dataset with Frutcherman-Reingold algorithm

(b) Full Dataset with OpenOrd layout algorithm

(c) Full Dataset with Yifan Hu Multilevel layout algorithm

Figure 5. Full Dataset

In Figure 5 (a), the data is visualised through the Frutcherman-Reingold layout. In Figure 5 (b), the OpenOrd layout algorithm is used to visualise the data. In Figure 5 (c), the Yifan Hu Multilevel layout algorithm is presented. The visualised data presents snapshots from three servers of the hospital networks, representing only a small section of the hospital network infrastructure. With the ultimate aim of the process to capture snapshots at regular intervals on all the hospital servers the visualised data demonstrates that this is problematic due to the sheer quantity of data. Once the data has been visualised, interactions between the user and the visualisation itself are a challenge. For this reason, captured network data needs to be pre-filtered in order to simplify the visualisation and the visualisation process.

In Figure 6, several heatmaps are presented, comparing frequency of Local Address and Foreign Address counts in the datasets. Additionally, in Figure 6, a comparison of IP addresses (converted to integer values) are presented.

Calculating the integer value of an IP address is commonly calculated by breaking the address into 4 octets and using the equation:

$$(first\ octet\ *\ 256^3)\ +\ (second\ octet\ *\ 256^2)\ +\ (third\ octet\ *\ 256)\ +\ (fourth\ octet) \tag{2.6}$$

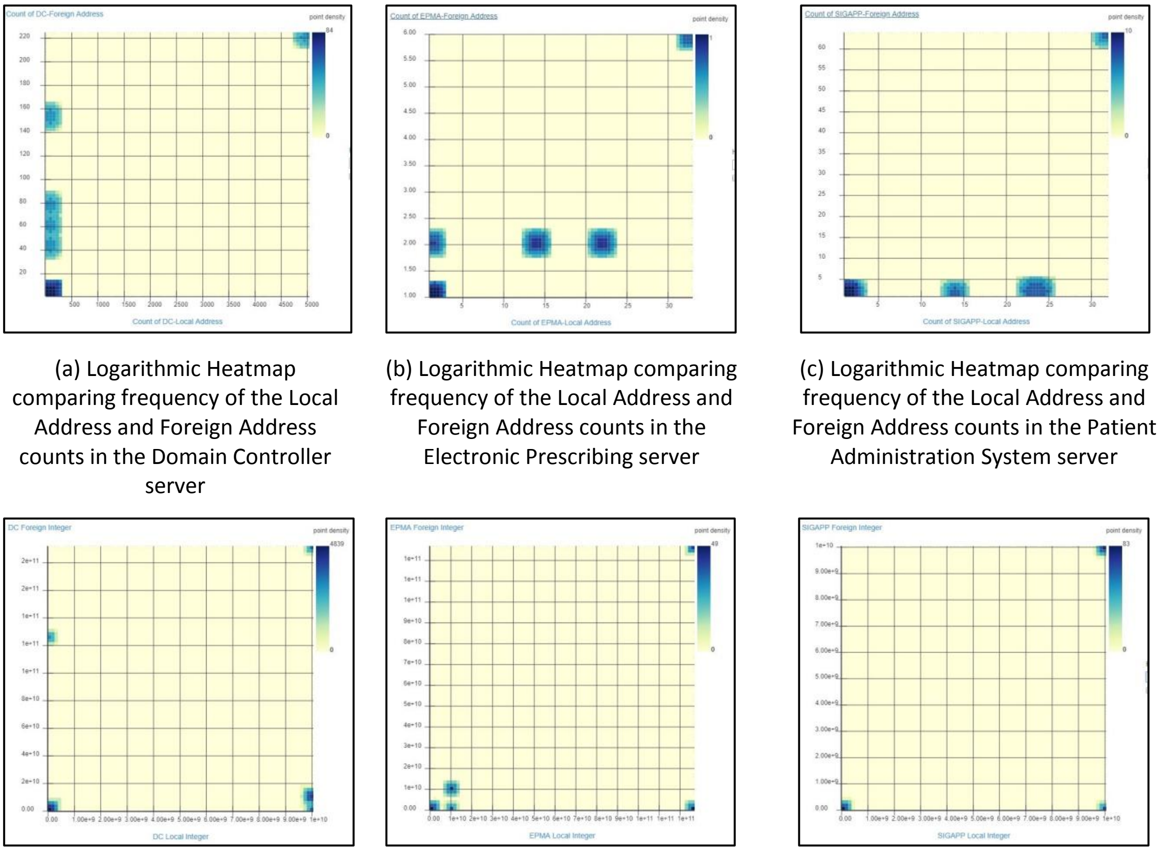
Representing the data as a logarithmic heat-map is an approach for identifying data points of interest. Using a logarithmic scale, lower-scale values are not compressed down into the congested section of the graph where the unique values would be challenging to identify. However, the density of the dataset prohibits valuable insights from being derived. Producing a real-time heatmap would be inefficient, as frequencies would need to be calculated in real time. IP addresses would also need to be converted into integers in real time, and this is a computationally intensive process. Additionally, many of the IP values are within the same range, as they are communicating locally within the hospital. In comparison to Figure 5, individual data points cannot be visualised separately, as they all fall on the same point on the heatmap, with only colour range to represent density. This approach is inefficient for visualising hospital network data.

The visualisation is constructed using a logarithmic algorithm, outlined in *(7)*.

$$f(x) = log_b(x)$$

(2.7)

Where the base *b* logarithm of *x* is equal to *f(x)*. In this sense, a logarithmic heat-map is appropriate as the log scales enable a significant range of coefficients to be displayed. Lower-scale values are not compressed down into the congested section of the graph where the unique values would be challenging to identify. Representing the data as a logarithmic heat-map is a clear approach for identifying data points of interest. However, the density of the dataset prohibits valuable insights from being derived, and a real-time graph would be inefficient. The quantity of data prohibits all the data points from being visualised. In section 2.2.2.3., data normalisation, feature extraction and machine learning algorithms are applied to the dataset to detect abnormal EPR access. Once the dataset has been administered by these algorithms, visualisation techniques are applied. In doing so, the situational awareness of a patient privacy officer is enhanced.



(a) Logarithmic Heatmap comparing frequency of the Local Address and Foreign Address counts in the Domain Controller server

(b) Logarithmic Heatmap comparing frequency of the Local Address and Foreign Address counts in the Electronic Prescribing server

(c) Logarithmic Heatmap comparing frequency of the Local Address and Foreign Address counts in the Patient Administration System server

(d) Logarithmic heatmap comparing IP addresses (converted to Integer values) of the Domain Controller server

(e) Logarithmic heatmap comparing IP addresses (converted to Integer values) of the Electronic Prescribing server

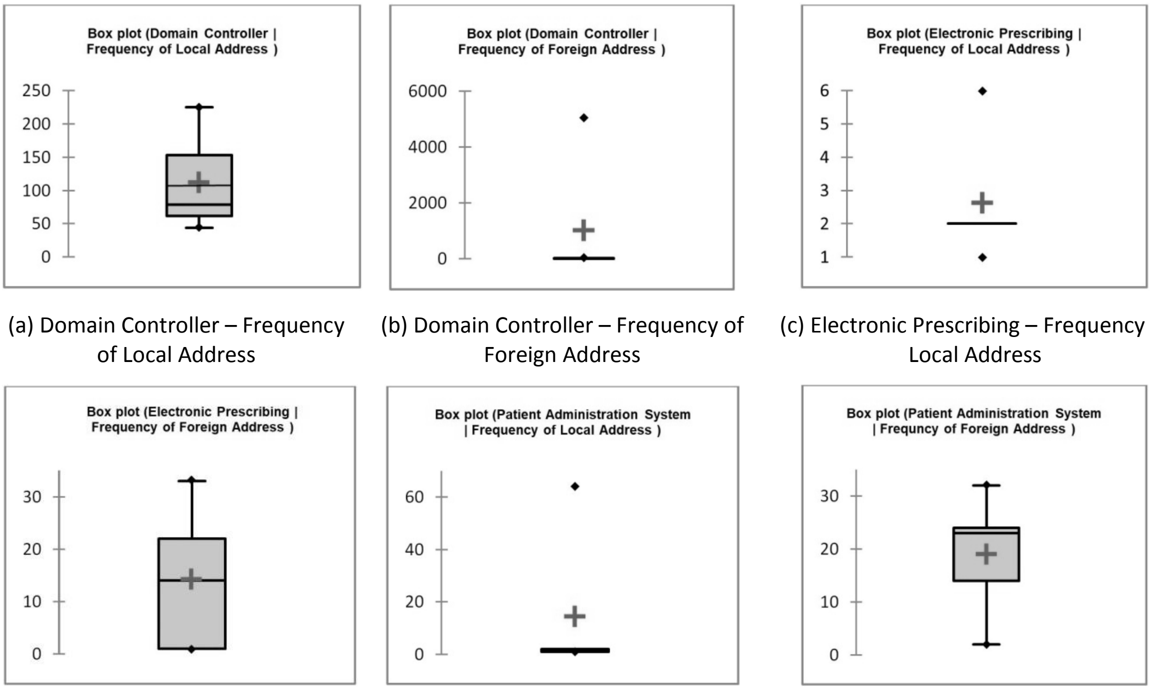
(f) Logarithmic heatmap comparing IP addresses (converted to Integer values) of the Patient Administration System server

Figure 6. Logarithmic Heatmaps

Figure 7 presents histograms of the most frequent items of Local and Foreign Address values in the Domain Controller dataset. Displayed in Figure 7 (a) the frequency for Local Addresses on the Domain Controller is shown. The output is varied, where the most significant IP address item counts comprise around 4% of the local IP address ports. 89.84% of the IP addresses in the dataset are unique values. In Figure 7 (b), the common most value is that of an asterisk due to the port not having been established indicating that at this time the Domain Controller had approximately 11.4% of its ports connected.

These ports are comprised of unique values displaying the port number of the remote computer to which the socket is connected. In the case of this dataset, all port numbers begin with the IP ranges used within the hospital, which indicates that all devices connected to the domain controller are devices on site and on the hospital network.

Figure 7 (c) and (d) presents the most frequent items of Local and Foreign Address values in the Electronic Prescribing dataset. Displayed in Figure 7 (c) the most frequent items for Local Addresses on the Electronic Prescribing is shown. The output is varied with the largest IP address item counts comprising around 7% of the local IP address ports. There are 3 sets of two value counts, with the rest being unique values; however, the Electronic Prescribing dataset is the smallest of the three datasets. Figure 7 (d) depicts the most frequent value is that of an asterisk due to the port not having been established.



(a) Domain Controller – Frequency of Local Address

(b) Domain Controller – Frequency of Foreign Address

(c) Electronic Prescribing – Frequency of Local Address

(d) Electronic Prescribing – Frequency of Foreign Address

(e) Patient Administration System – Frequency of Local Address

(f) Patient Administration System – Frequency of Foreign Address

Figure 7. Boxplots of Local and Foreign Addresses

Similarly, there are comparable data patterns between the data trends in Foreign Address frequencies for the Electronic Prescribing data in Figure 7 (d) and for the Domain Controller displayed in Figure 7 (b). Similarly, open port values such as 0.0.0.0:0 and [::]:0 are available indicating that the Electronic Prescribing had approximately 21.5% of its ports connected. This indicates that in both the Electronic Prescribing and Domain Controller, at any given time, most ports are open and waiting for a connection.

Figure 7 (e) and (f) presents the most frequent items of the Local Address and Foreign Address values in the Patient Administration System (PAS) dataset. Displayed in Figure 7 (e) the most frequent items for Local Addresses on the PAS are shown. Uniquely here, the output is largely comprised of a single port, accounting for almost 37% of the total Local Address connections. This number then quickly decreases and the majority of the IP addresses are unique values much like the Electronic Prescribing dataset. Figure 7 (f) shows the most frequent value is that of the 0.0.0.0:0 port, the [::]:0 port and the asterisk due to the port not having been established indicating that at this time the PAS had approximately 54.5% of its ports connected. Again, similar to the Domain Controller and Electronic Prescribing Datasets, the values for the Foreign Address have more than half the ports open. This indicates that each server has a large number of open ports, waiting for a connection to a device.

2.2.2.3. Situational Awareness Example

Situational awareness of network data enables end users to be able to identify where further cyber security systems need to be put in place. In addition, identifying where best practices and policies can be implemented minimises the risk of a cyber-attack; such as scanning attacks, injection attacks and jamming attacks (as detailed in Section 2.2.1.1), on highly confidential personal data.

Through removing irrelevant data from the dataset, only pertinent and useful data is analysed. For example, each of the datasets used has more than half of their Foreign Address data points ‘Open’, waiting for a connection to be made. When processed, this creates a large amount of unwanted noise in the visualisation which can be removed and potentially replaced with a figure representing the percentage of the dataset whose data points are open. This allows cyber security analysts’ attention to be focused on the more important, and potentially threatening data points.

Advanced data analytics techniques are used to further refine the data, removing low-threat data. This allows cyber security analysts to focus their attention on only the data points that are presenting any clear potential of moderate to high level threat, and differentiating between the two at the top-level visualisation. These algorithms filter the initial datasets in order to remove noise from the visualisation in order to present salient points to the analyst. The analyst then explores the visualisation and marks the highlighted data as either safe, or as malicious and pertaining to a certain attack type. Additionally, through the use of information-rich packet data captured at regular minute intervals, the visualisation draws from a wider database and highlight commonalities between them. In this case the refined visualisation may look more like the visualisation in Figure 9 (a). Rather than simply showing clusters of all data, as the visualisation in Figure 9 (b) does, the data has been processed and cleaned, so that only clusters with moderate to high level threats are present. As such, the visualisations further enable the situational awareness of network activity, over time, to become clearer. The visualisation clearly highlights any unidentified Foreign Address IP, regularly connecting to a Local Address IP or several IP addresses, for further investigation.

Comparing the four netstat command datasets, for each server after initially processing the data, enables it to become clearer. For example, with regards to the Domain Controller dataset, once the data has been processed for most frequent IP values as in Figure 8, this data can be investigated further. This process would identify and remove superfluous low-risk data connections in order to present further, refined data visualisations. Doing so would present the data without unnecessary noise cluttering the visualisation and leaving only the relevant and potentially malicious data connections highlighted for the situational awareness analysts.

For example, with a frequency of 225, representing almost 4% of the Local Address IP values, are a single IP Address open on port 445. This data point therefore, can be isolated in the four separate netstat command datasets to determine which Foreign Address devices these are connecting to and which processes are running, in order to determine their potential risk. Once this data point is isolated it becomes clear that this Local IP Address is connecting to a number of hospital devices with the hospital prefix, and that Foreign Addresses were resolving to a device name with a hospital device name prefix, initially suggesting that this could be ignored. The netstat was unable to obtain ownership information of the binary program involved in the creation of the connection on every value of these IP connections using the netstat –b parameter. This is due to netstat not having the necessary administration privileges to call this information. Through the use of the Hospital’s Resource Monitor, and checking the active TCP Connections against current processes with network activity, the process running on this port can be determined as the Server Message Block (SMB). The SMB operates as an application-layer network protocol, which is used for providing shared access between nodes on a network, such as access to shared files, printers and serial ports. In May 2017, the WannaCry ransomware campaign exploited an SMB vulnerability on Port 445. The exploit enabled the malware to use worm-like network propagation, encrypting files and demanding ransom payment, unless the system had been patched by Microsoft security bulletin MS17-010. The attack resulted in network downtime for 48 UK hospitals, with 6 suffering disruption lasting several days [29]. Through removing noise within the network infrastructure this vulnerability is detected.

Having isolated the Foreign Address values not resolving to a hospital device or server name with a known prefix, there were two discernible data points worth investigating further. The first is connected to a device/server with a notably unusual name. While this may be benign it would still be worthwhile leaving data points like this in the visualisation system, in order to highlight this unusual data connection to a cyber security analyst and the hospital IT team. A noteworthy factor is a Foreign Address value, to which the device name cannot be resolved by netstat; it returns an IP Address. So, this connection is established to be the most frequent IP address value assigned by the Domain Controller. Yet the device name cannot be resolved by netstat, nor can the program which initiated the connection. Through filtering the data in this way, a prominent data point becomes apparent.

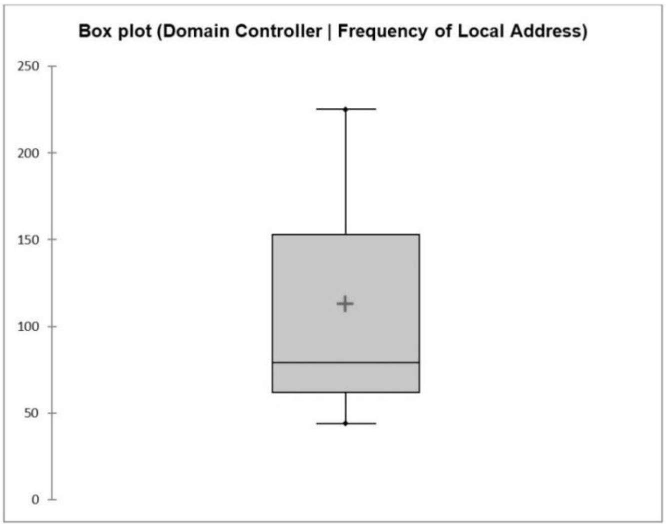


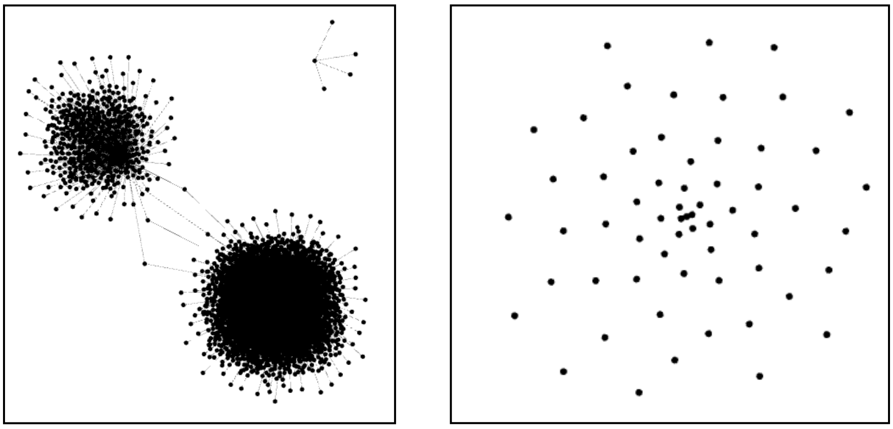
Figure 8. Domain Controller – Frequency of Local Address

The other most common address values present in the Domain Controller dataset, in Figure 8, have value counts of 153, 79, 62 and 44 respectively. With the exception of a value of count 3, all other Local Address IP values in the dataset have counts of 2 or fewer. The Local Address value with the count of 153 are all running the *lsass.exe* process. The *lsass.*exe process is the Local Security Authority Subsystem Service and it verifies the validity of logons to the server, handles password changes and creates access tokens. It is a critical system process but is sometimes targeted by malware because of this. All but two of the IP address are resolving to known hospital devices. So, these can all be categorised as low-risk connections initially and filtered from the dataset and the two connections with unresolved Foreign Address device names can be isolated and visualised for further investigation.

The Manufacturer and User Facility Device Experience (MAUDE) database is a publicly available database managed by the Food and Drug Administration (FDA) [62]. It was established through the Safe Medical Devices Act of 1990 which requires sites in which medical devices are used to report device-related fatalities and serious adverse events to the FDA [74]. The data has been publicly available since 1995 [75]. Attackers can attempt to masquerade malicious software through techniques such as naming their file *Isass.exe* (with a capital I rather than lower case l). If an *lsass.exe* file is located in a folder other than *C:\Windows\system32* it can be considered malware [76]. Anti-malware software is a potential solution for malware security risks such as this; however, in 2016 a medical device monitoring a patient’s physiological data whilst undergoing a heart catheterization procedure shut down and required a reboot. This caused a 5 minute delay to patient care, and was due to an anti-malware software performing hourly scans [77].

In order to remove noise from the data in Figure 9, the data was analysed. The Local Address values with a count of 79 are resolving to known hospital network devices and running the *lsass.exe* process. The Local Address values with a count of 62 are largely resolving to known hospital network devices and running the *lsass.exe* process. In this case there are some more unusual device names being returned suggesting that it is connected to some more niche hospital devices and some hospital servers. There are a further 7 unresolved Foreign IP Address values in this subsection of the dataset. Finally, the Local Address values, with a count of 44, are primarily resolving to known hospital network device names. Albeit similar to the previous count, there are some more unusual device names, and there are a further 6 unresolved Foreign IP Address values. The process involved in creating these connections is the *svchost.exe* process. The *svchost.exe* process is the ‘Service Host’ and is a critical Windows component. It allows a number of services to share a process in order to reduce resource consumption, often with a number working in tandem, to prevent a failure in one causing a full system crash. From our research, comparable to *lsass.exe*, malicious attackers sometimes masquerade their malware to look like the *svchost.exe* process and, if it is located in a folder other than *C:\Windows\system32*, it can be considered malware [76]. In addition to reporting the process *svchost.exe* the netstat –b command also returns in this instance the process *RPCSS* [78]. *RPCSS* are Remote Procedure Call System Services, which is a Service Control Manager for servers. It performs activation requests, object exporter resolution and distributed garbage collection designed to make client/server interaction easier and safer by factoring out these common tasks.

Of the remaining data therefore, there are no further unresolved Foreign Address device names, however these connections were initiated by some interesting and unique processes in the dataset on these connections and as they are running on unique Local IP Addresses, they have been included in the refined dataset for the visualisation in Figure 9. In Figure 9 the data is visualised again, with the data points identified as low risk removed from the dataset to reduce noise, in addition to removing the data points of the Listening UDP ports.



(a) Visualisation of Domain Controller dataset before noise reduction using ForceAtlas2 algorithm

(b) Visualisation of Domain Controller dataset after noise reduction using ForceAtlas2 algorithm

Figure 9. Visualisation of Domain Controller dataset before and after noise reduction

It is worth noting that in the interim between running the netstat –b command (the dataset executables, for investigation purposes), and running the netstat –n command (the numerical dataset, for visualisation purposes), a few of the connections disconnected. Therefore, some anomaly data connections identified in the netstat –b dataset were no longer present in the netstat –n dataset.

As demonstrated, hospital network data is complex. Visualisations of potentially anomalous data behaviour is difficult to achieve, but necessary to form a situational awareness of the hospital infrastructure. The complexity of this problem is amplified with the complexity of medical systems' data structure, specifically EPRs. This section demonstrates that highlighting unusual data points within healthcare infrastructures can focus analyst attention on potentially malicious activity within a dataset.

2.3. Medical System Interactivity

Reduced information visibility due to data complexity, fragmentation, interoperability and lack of specialisation undermine the security of healthcare organisations [43]. Due to the increased need for 24-hour data access the boundaries for healthcare systems is evolving; GPs are progressively using VPNs and 3G connections to remotely access patient data. As a result, the number of access points for hackers is increasing [4] and healthcare organisations should have processes in place to identify data loss and wipe data remotely if necessary. Hospital networks are frequently upgraded with new digital technologies, rather than being replaced, due to the cost involved with new developments. The reliance on legacy software, and small scale bespoke software solutions, results in an increased vulnerability to cyber-attacks from external sources [10].

Healthcare has only recently been digitised [78]. The result is that organisations are still understanding how best to protect their digital assets. As detailed in section 2.2.2.3, attaining situational awareness of data flow within hospital infrastructures is already a considerable challenge. The inclusion of more complex and bespoke medical systems only compounds this problem. One of the main challenges is that a significant number of users require access to patient records to facilitate their occupation [79]. Healthcare infrastructures may also involve temporary employees or visitors from partner organisations. As such, in this section, a background discussion is put forward on the layout of medical infrastructures and the existing healthcare network security challenges. This background aids with the development of an approach to understanding the overall network behaviour.

Figure 10 is the medical system topology of the partner hospital in this research, showing the data flow process currently in place within the network. A single directional arrow indicates data flow in that direction. A bi-directional arrow indicates data flow in both directions. A system that overlaps the EPR is directly integrated, such that it works seamlessly within the EPR, but is still a separate system that can be accessed independently if required.

Each piece of software in the green section represents primary care software, for the aid of GPs. Typically, they use a software called EMIS-Web (Egton Medical Information Systems) [80]. A number of pieces of software are hosted by NHS Digital and interact with both the Health Information Exchange platform for secondary care and the GP System. These sit on a platform named the SPINE which supports the IT infrastructure for health and social care [81]. SPINE software includes:

- Message Exchange for Social Care and Health (MESH), as the primary messaging service across the NHS
- E-Referral Service (eRS) enables GPs and patients to directly book a hospital appointment electronically (rather than sending a referral letter manually)
- Personal Demographics Service (PDS) is the authoritative source of patient demographics data in the NHS. When a GP updates a patient's demographic data it is updated on the PDS. If a patient presents at another NHS service (such as a walk-in centre, or A&E), this information can be pulled down given some key patient data (such as surname, postcode and date of birth).
- The Child Protection Information Sharing (CP-IS) project enables staff to be quickly notified across the NHS if a child known to Social Services as a "Looked After Child" presents in an unscheduled NHS setting (such as A&E). The Social Care systems across the country populate the CP-IS.

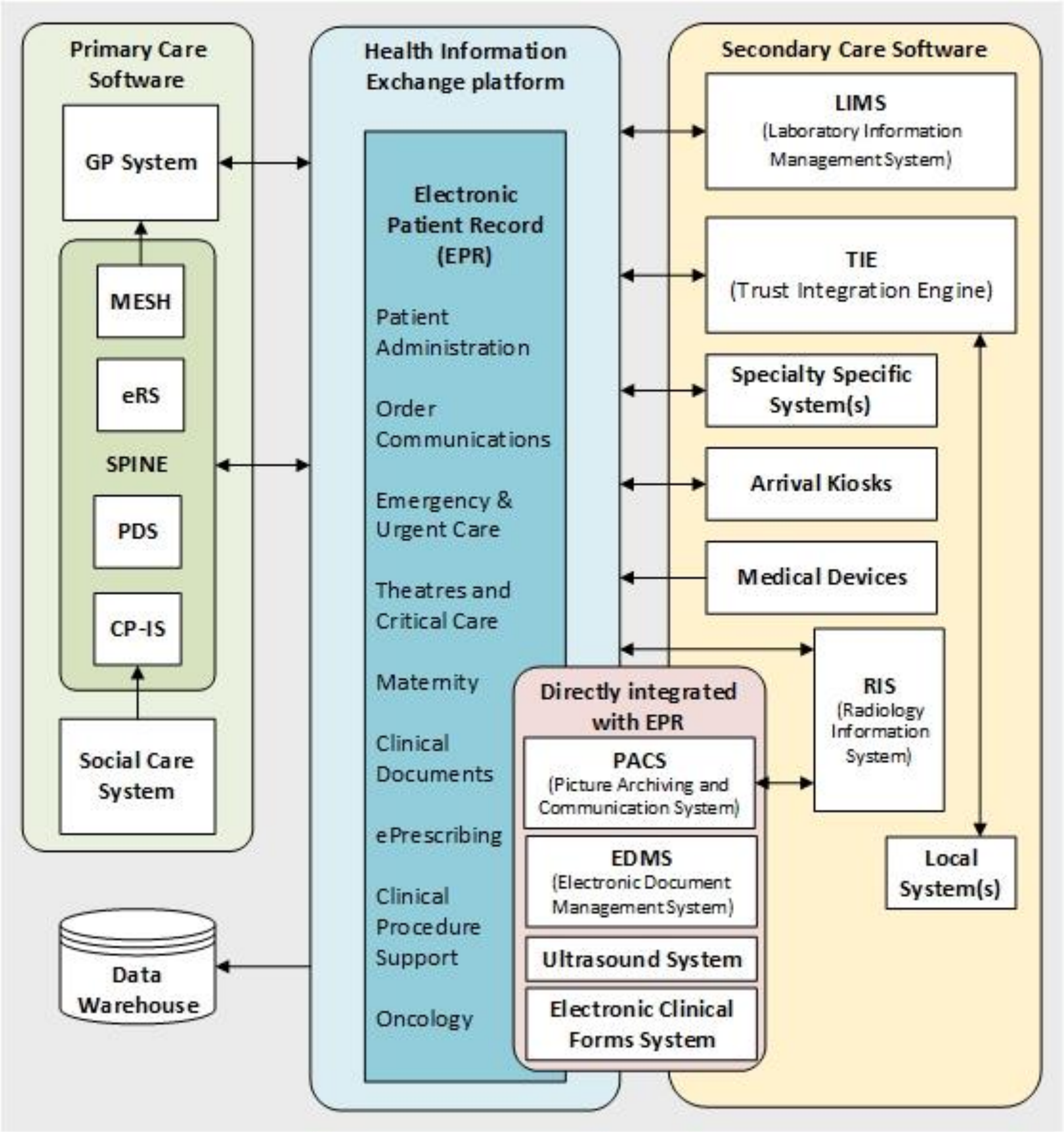

Primary Care Software
GP System
MESH
eRS
SPINE
PDS
CP-IS
Social Care System
Data Warehouse
Health Information Exchange platform
Electronic Patient Record (EPR)
Patient Administration
Order Communications
Emergency & Urgent Care
Theatres and Critical Care
Maternity
Clinical Documents
ePrescribing
Clinical Procedure Support
Oncology
Directly integrated with EPR
PACS (Picture Archiving and Communication System)
EDMS (Electronic Document Management System)
Ultrasound System
Electronic Clinical Forms System
Secondary Care Software
LIMS (Laboratory Information Management System)
TIE (Trust Integration Engine)
Specialty Specific System(s)
Arrival Kiosks
Medical Devices
RIS (Radiology Information System)
Local System(s)

Figure 10 – Medical System Topology

The Health Information Exchange (HIE) platform enables integration of data from a number of different pieces of software with the EPR. The EPR itself sits on a server within the Trust and is the most comprehensive record of clinical information within the hospital. It enables a number of clinical workflows, such as:

a) Patient administration - registering patients, admitting them, managing clinics, booking appointments etc.

b) Order Communications - enabling clinicians to order diagnostic tests and medications)

c) Emergency & Urgent Care, Theatres and Critical Care, Maternity, Oncology - bespoke tools for these settings
d) Clinical Documents – digitised versions of clinical documentation enabling information in one place and across the organisation
e) ePrescribing – prescribing and administering medications to patients
f) Clinical Procedure Support – procedure-based decision support enabling clinicians to provide more standardised healthcare, such as alerting an unfamiliar doctor/nurse that a patient is showing signs of Sepsis/Pneumonia based on unusual vital signs

Software in yellow are used by secondary care such as hospitals. These include the Laboratory Information Management System (LIMS), which manages the flow of the clinical Labs within the healthcare infrastructure. The Trust Integration Engine (TIE), which enables a number of specialty specific local systems to deposit data to be provided to the HIE. A number of Specialty Specific Systems also integrate directly with the HIE, as do some Medical Devices and patient Arrival Kiosks. Additionally, a number of pieces of software are directly integrated with the EPR, though are themselves standalone systems and can be accessed independently. These include 1) the Picture Archiving and Communication System (PACS), which includes X-Ray images, 2) the Radiology Information System (RIS) which operates in conjunction with the PACS for managing radiology tasks, 3) Electronic Document Management System (EDMS), which stores scanned images of paper documents, used to manage the transition from paper to digital, 4) Ultrasound System, allowing integration with the maternity flow, and finally 5) Electronic Clinical Forms system, which is another way of managing digital versions of clinical documents.

Figure 11 is the hospital flow overview process and details the hospital flow of all NHS Trusts. Figure 11 was created with the partner hospital in this research and demonstrates the complex adaptive network of the hospital, and the number of interactions with the Electronic Patient Record involved within a patient’s care.

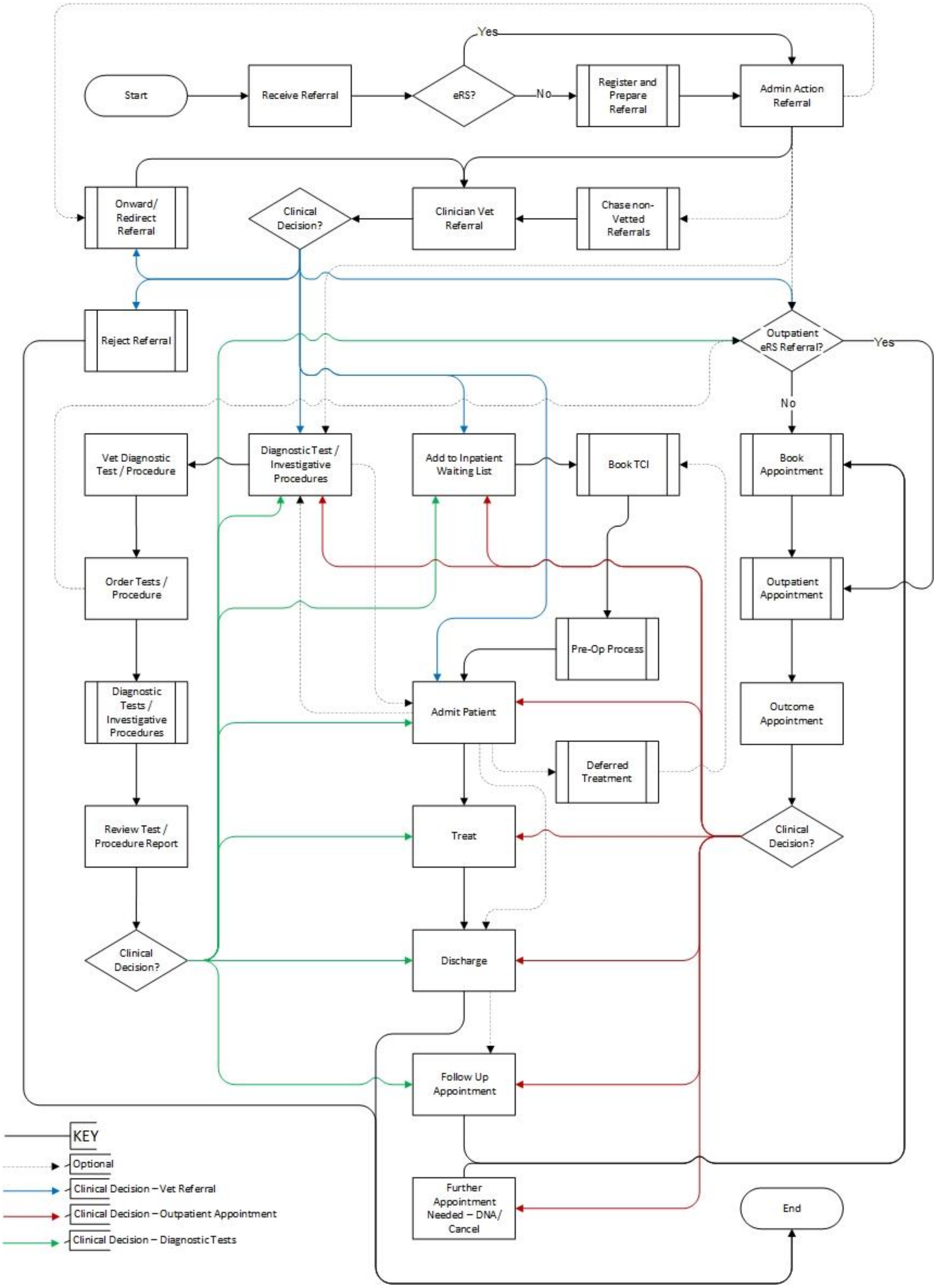

Start
Receive Referral
eRS?
Yes
No
Register and Prepare Referral
Admin Action Referral
Onward/ Redirect Referral
Clinical Decision?
Clinician Vet Referral
Chase non-Vetted Referrals
Reject Referral
Outpatient eRS Referral?
Yes
No
Vet Diagnostic Test / Procedure
Diagnostic Test / Investigative Procedures
Add to Inpatient Waiting List
Book TCI
Book Appointment
Order Tests / Procedure
Outpatient Appointment
Pre-Op Process
Diagnostic Tests / Investigative Procedures
Admit Patient
Outcome Appointment
Deferred Treatment
Review Test / Procedure Report
Treat
Clinical Decision?
Clinical Decision?
Discharge
Follow Up Appointment
Further Appointment Needed – DNA/ Cancel
End
KEY
Optional
Clinical Decision – Vet Referral
Clinical Decision – Outpatient Appointment
Clinical Decision – Diagnostic Tests

Figure 11 - Hospital Flow Overview

The NHS has an 18-week Referral to Treatment (RTT) pathway for all patients [82]. Therefore, each stage of this process is monitored in order to progress patients through their journey. A typical journey is that a patient is referred to the hospital, who then process the referral by admin and a clinician vets the referral as appropriate. They are then brought in for an Outpatients clinic consultation, where they are either sent for diagnostics, given treatment, or admitted (Inpatients). However, a number of clinical decision points affect this flow and a comprehensive understanding and tracking of each patient journey is required in order to maintain patient safety. Using an EPR is typically the most effective way of collating this information and enabling a holistic view of patient care.

2.4. Electronic Patient Records

Electronic Patient Record (EPR) systems support clinical operations within healthcare organisations [83] and improve the safety [84] and efficiency [85] of healthcare delivery, whilst reducing costs [86]. The shift from paper-based to computer-based patient records has improved the availability of patient data without limitations of time or place [87]. Additionally, availability is one of the three principles of information security. However, this shift in making health information accessible and useable by a range of health professionals conflicts with public perception of patient confidentiality and autonomy [88]. To ensure patient privacy in this landscape, there is a requirement for focus on the other two principles of information security, confidentiality and integrity [89]. A continued focus on trustworthy security and privacy mechanisms for health information sharing is necessary due to public concern regarding privacy of EPRs [90].

Patient data privacy, security and confidentiality concerns are validated through numerous reports of patient information being stolen, lost, misplaced, or released without authorisation [91]. Hacking and identity theft is often cited as a cause for concern regarding EPR security, alongside unauthorised access [92]. In particular, patients are often sceptical regarding the ability of the NHS to safeguard medical information and manage large technological projects, due to failed programs such as NPfIT (National Programme for IT) [93]. This view is particularly held amongst those who had worked in the NHS themselves [92].

2.4.1. Patient Privacy Considerations

Patient privacy within EPR systems is typically enforced through corrective mechanisms, such as two factor authentication, training and confidentiality agreements [94][95]. Approaches for detecting illegitimate access to EPRs [96] include i) restricting access control [97], ii) applying patient-user matching algorithms [98], iii) applying scenario-based rule extraction [99], and iv) information gathering from EPR and non-EPR systems using a secure protocol [100]. This is in addition to commonly-used security mechanisms, such as secure networks with firewalls, encrypted devices and messages, strong user passwords, auditing and device timeouts [95]. 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 [96]. The management of patient data in electronic form decreases healthcare administration costs, strengthens care provider productivity and increases patient safety [101].

There are 13 features required for security and privacy in EPRs [102]. These include system and application access control, compliance with security requirements, interoperability, integration and sharing, consent and choice mechanism, policies and regulation, applicability and scalability and cryptography techniques.

Additionally there are 3 primary focuses of HIPAA (Health Insurance Portability and Accountability Act of 1996) regulations for attaining security in an EPR [103].

1. Provide sufficiently anonymous release of information for research purposes.
2. Provide appropriate controls to prevent unauthorised people from gaining access to an organisation's information systems and control of external communications links and access.
3. Provide mechanisms for controlled and user-differentiated access to individual patient records.

Traditional methods for defining security policies within organisations are problematic within the context of healthcare organisations due to their reliance on the knowledge of domain experts, or observations of external specialists [104]. Within healthcare the number of security policies are large, defined in an ad hoc manner and can be revised at a moment's notice [105]. A primary feature of patients' desire for widespread EPR adoption is transparency, with patients enquiring who has the ability to access their medical records, in addition to determining who has viewed them [106].

Due to the risk of unauthorised use, access and disclosure of patient information, patient privacy and confidentiality concerns need to be addressed [107]. The patient privacy perspective is operationalised through using privacy concern as the most common measure [108]. Leakage or modification of patient data can be intentional or unintentional and can derive from both external attackers and internal staff [109]. The intrinsic value of stolen healthcare data on the black market is well recognised [110]. Additionally, the healthcare sector mandates public disclosure of data breaches, increasing public awareness and concern over privacy [111].

Patient Privacy concerns within EPRs is resulting in a loss of trust in healthcare providers by patients [112]. This is evidenced by the following studies:

Table 5 – Patient Privacy concerns by patients

Year	Findings
2015	78.9% of participants would worry about the security of their record if it was part of a national EPR system and 71.3% felt the NHS was unable to guarantee EPR safety [92]. Additionally, 46.9% responded that EPRs would be less secure compared with how their health record was held at the time of the survey [92].
2014	64.5% of patients expressed concerns regarding data breaches when personal health information was being transferred between healthcare professionals electronically [113].
2013	48% of people believed health IT would worsen privacy and security [114].
2012	Approximately 60% of respondents believed that widespread adoption of EPR systems will lead to more personal information being lost or stolen [115].

2012	31% have concerns that the privacy and security of their medical information may be at risk within EPRs [116].
2010	California Healthcare Foundation found that 68% of patients are concerned about the privacy of personal medical records [117].
2009	76% of people thought it was likely that an unauthorised person would get access to an EPR [118].
2008	62% of respondents did not think data stored within an EPR would remain confidential [119].
2006	80% of people were very concerned about identify theft or fraud [120].

Additionally, concerns about patient privacy can lead to patients being selective about the information they provide to healthcare providers, or offering incomplete or misleading information [112]. Withholding information due to privacy concerns among patients is evidenced in the following studies:

Table 6 – Studies evidencing patients withholding information due to privacy concerns

Year	**Findings**
2014	The Office of the National Coordinator for Health Information Technology (ONC) found 7% of patients have withheld information from their healthcare provider due to privacy of security concerns, with this percentage increasing to 33% among those who strongly disagree that there are reasonable protections in place for EPRs [121].
2014	12.3% of patients withhold information out of concern for a data breach, with the likelihood of withholding information higher among respondents who perceived they had little say regarding how their medical records were used [113].
2011	FairWarning in Canada found that 43.2% of patients withhold information based on privacy concerns, and 31.3% would postpone seeking care for a sensitive medical condition [121]. Additionally, 61.9% reported that if there were serious or repeated breaches at a hospital

	where they had treatment it would reduce their confidence in the quality of healthcare at the hospital.
2010	California Healthcare Foundation found that 48% of patients may hide information from their doctor if it was shared through an EPR [117].
2007	Harris Interactive found that 17% of patients would withhold medical data due to worries of data disclosure [122].

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 in the overall quality of care [94]. However, with the added benefits of technology in healthcare, the potential for unauthorised and illegal access to patient information has increased [123]. Users may abuse their privileges for personal reasons, such as viewing records of relatives, friends, neighbours, co-workers or celebrities [96]. Therefore, patients are becoming increasingly concerned regarding the privacy and security of their health data [124]. 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 [95]. Examples of privacy protection within healthcare include encrypted devices, strong passwords, two factor authentication, training and confidentiality agreements [95]. It is difficult to impose an access control policy on employees in a healthcare setting due to the dynamic and unpredictable care patterns of hospital care [94].

2.4.2. Access Control Limitations

Access to EPR systems is often managed through role based access, where once a user has been authenticated, they are allowed unhindered access inside the perimeter [96]. Access to EPR data is audited heavily within healthcare infrastructures. However, it is often left untouched in a data silo and only accessed on an ad hoc basis. When there is reason to suspect that illegitimate accesses have occurred, a review of the audit logs is undertaken by a security expert. VIPs are an exception, for which the audit logs are regularly monitored. Otherwise, only if an official complaint is logged by a patient are audit logs reviewed. The majority of breaches are currently discovered by the person whose confidentiality has been breached. However this is inefficient, as it requires the information to be collated and reviewed by a security expert, is purely retrospective, and the process is only triggered when there is cause to believe that illegitimate access has occurred [96]. Therefore, there is a motivation to automate and alleviate partially the burden of this process [94]. The fundamental limitations in privacy officers manually reviewing audit logs for potentially suspicious accesses are threefold [96]. 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 contain 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.

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 [94]. Access control based approaches are limited due to several factors [95], including:

- In a hospital setting, it is safer to detect anomalous behaviour than prevent it, as preventing access to patient data could lead to patient harm
- Unpredictable and dynamic care patterns, including scheduled and unscheduled inpatient, outpatient and emergency department visits
- Varied workflows, with providers requiring access in unexpected areas
- A mobile workforce, with access required at unexpected locations and times
- The collaborative nature of clinical work and teaching environments
- A large number of users with varied job titles and roles
- Users job titles not directly relating to a list of patients whose records it would be appropriate to access
- Access Control can stifle innovation within a healthcare setting. If anomalous behaviour is detected it can be justified, if it is prevented altogether it inhibits potential improvements

Due to these limitations, access control approaches are insufficient as the sole method of anomaly prevention within EPRs. For example, the Access Matrix Model (AMM) is a conceptual framework that specifies each user’s permissions for each object in the system [125]. 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 [126].

Role-Based Access Control (RBAC), however, maps users to roles and maps permissions to the roles [127]. 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 [128]. Whilst more computationally tractable, RBAC roles tend to be static and inflexible, and therefore not responsive to the shifting nature of roles [129].

Attribute-Based Access Control (ABAC) provides flexible, context-aware access control through evaluating the attributes of entities, their subject and object, the operation, and the contextual environment, such as time and location, of the request [130]. Boolean logic can then be applied to the operational request to determine access rights, such as "IF, THEN" statements regarding the request, the resource and the action. 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 [128]. EBAM enterprises often manually review the audit logs of VIPs to determine inappropriate accesses [131]. Break The Glass (BTG) is a policy, which allows users to override access controls in necessary instances [132]. EBAM enterprises would manually review the audit logs every time a user broke the glass [128].

Task-based Access Control (TBAC) extends the user-object relationship though the inclusion of task-based and contextual information [133]. However TBAC is limited to contexts that relate to tasks, or workflow progress and EPRs cannot always be easily portioned into tasks [126].

Team-based Access Control (TeBAC) groups users in an organisation and associates a collaboration context with the activity to be performed [134]. However, these models have not been fully developed or implemented and it remains unclear how to implement them within a dynamic framework [126].

2.4.3. Audit Logs

Without audit mechanisms, EPR systems are vulnerable to undetected misuse, as users could modify or delete health information without their actions being traceable [135]. Audit Logs are usually recorded and stored for the purposes of access management [136]. However, they can also be used for the benefits of monitoring employee behaviour and system failures [137]. Audit Logs should have at least these elements: 1) Time; 2) Date; 3) Information Accessed and 4) User ID [138].

Thorough and frequent analysis of audit logs have been shown to discourage abuse [139]. Yet, this analysis often consists of manual audit log review [140]. Motives for a breach of confidentiality within an EPR that may be detected through audit log analysis include [139]:

1. Characteristics of the patient or patient record (such as a VIP).
2. A relationship between the user and the patient.
3. A relationship between the user and another person represented in the patient record (such as a spouse or child).

Indicators of confidentiality breaches can also be separated into positive and negative indicators, where positive indicators are evidence of a potential breach and negative indicators are evidence of expected behaviour, typically based on the established provider role [139]. Probability scoring and an indicator weighting mechanism can aid in prioritising possible breaches for further investigation.

To detect accurately a confidentiality breach, Motivational Indicators can be applied at the Patient Access Level. Whilst Behavioural Deviation Indicators can be applied at the Session Level [139]. Motivational Indicators include considerations given to: 1) any relationship between the user and patient that could be a motivator to breach confidentiality, such as a friend or spouse; and 2) characteristics of the patient or of the patient record that could be a motivator to breach confidentiality, such as a high-profile person. Behavioural deviation indicators include considerations given to: 1) session level statistics, such as total session length and number of patient records accessed; and 2) login characteristics such as failed attempts prior to successful login, time of day of login, and location from which login occurs.

Visualising audit log data can provide some initial insights into these datasets. In Figure 12, a heatmap is presented of a dataset comparing User ID to the duration of the patient record access. Figure 12 is extracted from a dataset of 1,515 unique User IDs and 72,878 unique Patient IDs. The graph shows a consistent point density of up to 47,341 patient records in the first row of the matrix, indicating that the majority of patient records are only accessed for fewer than 300 seconds (5 minutes). This would represent normal (expected) behaviour within the hospital (as revealed in consultation with the hospital). Whereas, 6 clusters (A-F) require investigation, as they represent users performing routines for over 16,000 seconds (4.44 hours); which would be classed as abnormal (unexpected) behaviour.

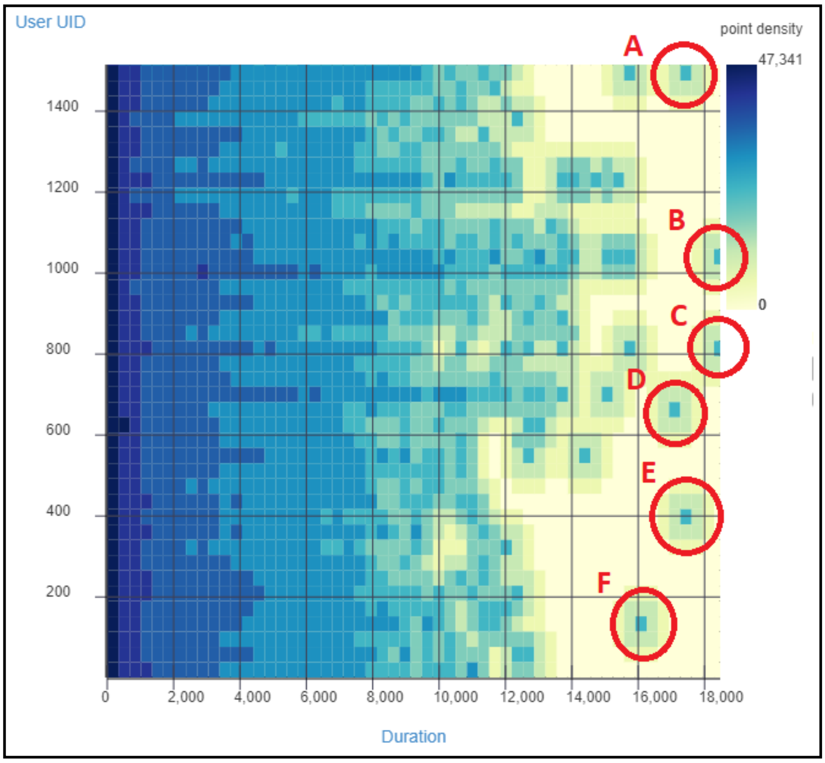


Figure 12 – Heat-maps (logarithmic) comparing 1million rows of ID data to the duration of the patient record access

2.5. Summary

In this chapter, a discussion of hospital infrastructures is presented. Hospital networks are discussed, along with their security challenges. This is highlighted through an investigation incorporating a real-world dataset, discussed in detail in Chapter 5. Through this work the complexity of acquiring situational awareness of data flow is demonstrated. Medical systems and their topology within the hospital infrastructure is then detailed. This further adds to the complexity of maintaining an awareness of data flow, due to the incorporation of more bespoke software’s interacting and a reliance on legacy software. EPRs allow a holistic view of patient activity within a hospital, however the wealth of patient information they hold presents a key challenge for maintaining patient data confidentiality. Therefore, Patient Privacy within EPRs is noted as the focus of this work, along with a discussion on current solutions such as access control methods and reactively monitoring audit logs. In chapter 3, artificial intelligence and machine learning is presented as the area of research to be applied to the context of patient privacy within EPRs. Multiple machine learning algorithms are detailed and investigated, along with a review of related work and applications of machine learning in similar contexts.

3. Artificial Intelligence and Machine Learning

3.1. Introduction

Organisations need to bridge the gap between cyber operations/resilience and the priorities of the business [43]. Organisational decision makers should synthesize highly disparate data into a coherent and concise narrative [43]. However, they face considerable challenges such as data complexity, fragmentation, interoperability issues and lack of spatialisation, which, together, degrade information visibility within organisations. Relating to healthcare security, the challenges are as follows; 1) a lack of labelled data from previous attacks; 2) constantly evolving bespoke attacks and 3) the analyst's limited investigative time and budget [141].

Artificial Intelligence aims to simulate human intelligence processes using machines, broadly making appropriate generalisations based on limited data [142]. The scientific field of Machine learning is an intersection of Computer Science and Statistics and seeks to answer the following question '*How can we build computer systems that automatically improve with experience, and what are the fundamental laws that govern all learning processes?*' [143]. Machine learning systems aim to learn automatically from data [144]; with an emphasis on the design of self-monitoring systems and take advantage of the data flowing through the program, rather than simply processing it [143]. Deep Learning is a subset of machine learning which aims to imitate the human brain in processing data using neural networks [145]. Big Data refers to information assets characterized by such a high volume, velocity and variety to require specific technology and analytical methods for its transformation into Value [146].

Current EPR privacy solutions employ either analyst driven solutions, or unsupervised machine learning solutions [96]. Both of these approaches are inadequate on their own. Analyst driven solutions often lead to a high number of false negatives, due to their reliance on human judgement, in addition to delays between attack detection and the implementation of countermeasures [141]. Similarly, unsupervised machine learning solutions are typically insufficient when unsupported by other classification techniques. Often an unsupervised approach leads to high numbers of false positive alarms, alarm fatigue and distrust by analysts [141].

3.2. Machine Learning

Typically, for an analytic process, machine learning approaches are considered in most modern-day applications [147]. This is because machine learning emphasises the design of self-monitoring systems, which self-diagnose and self-repair [143]. The technique is used commonly in web search algorithms, spam filters, recommender systems, ad placement, credit scoring, fraud detection, stock trading, drug design and many other real-world applications [144].

Machine Learning methods are most valuable in applications where it is too complex for people to manually design the algorithm. Machine learning models are trained on historical access data to classify future data access patterns [94]. 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 [94]–[96]. The use of statistical and machine learning techniques have also been used previously to detect fraud in financial reporting [148], to detect fraud in credit card transaction data [149], to construct a spam email detector [150], and to solve a fraud detection problem at a car insurance company [151].

Machine learning algorithms observe and learn data patterns and profile users' behaviour, which can then be denoted. Combined with cloud infrastructure and data visualisation, the way large datasets are understood is being transformed; allowing extraction of otherwise unobtainable meaning from vast quantities of information. This is now a proven approach for detecting zero day attacks and uncovering unknown threats [141]. There is a large volume of literature concerning big-data-based privacy-preserving machine learning algorithms. Genetics-based machine learning (GBML) [152], clustering fuzzy rule-based classifiers [153] and Linear Support Vector Machines (SVMs) [154] are examples of the general conventional means of choice for researchers.

There are three primary challenges facing information security that can be addressed through the use of machine learning [141]:

1. Lack of labelled data
2. Constantly evolving attacks
3. Limited investigative time and budget

To address these problems, a solution should use analysts' time effectively, detect new and evolving attacks in the early stages, reduce response times between detection and prevention and have a low false positive rate [141].

Machine Learning workflows require iterative experimentation in order to attain a desired accuracy. Through analysis of an existing model, the workflow is modified to improve performance with a developer-in-the-loop during the development cycle. Such iterations include adding/removing features, introducing new data sources, changing the machine learning model, adding ensemble averaging to the model and adding a Human-in-the-Loop Machine Learning (HILML) model.

Ensemble and semi-supervised machine learning techniques involve the combination of both labelled and unlabelled data to change learning behaviour [155]. Through the application of active learning, outlier detection is improved [156]. Due to the lack of labelled data for patient privacy violations within EPRs, semi-supervised learning has been applied for healthcare fraud detection [157].

3.2.1. Machine Learning Workflow

Machine learning techniques principally consist of combinations of three components, Representation, Evaluation and Optimisation [144] where the data is modelled as a set of variables [158]. The following metrics are employed, a particular task *T*, a performance metric *P*, and a type of experience *E*. If a system reliably improves its performance *P* at task *T*, following experience *E*, then it can be said to have '*learned*' [143].

The machine learning workflow is presented in Figure 13. Features and labels are extracted from raw data. If the workflow is supervised, a testing and a training dataset is formed of a subset of the data. In an unsupervised workflow the whole dataset is typically used as the training dataset. The algorithm is then applied to the data. This algorithm is then evaluated and validated, with hyperparameters selected and tweaked in order to achieve a desired model. Finally a model is obtained which can provide predictions [159].

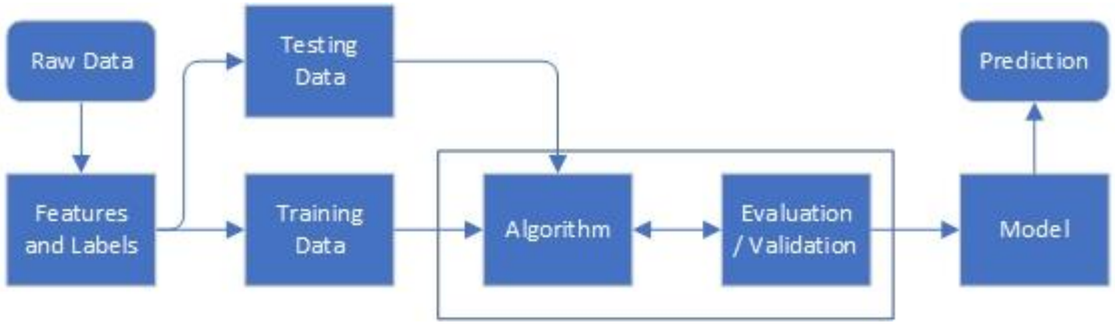


Figure 13 - Machine Learning Workflow

The classifier is trained using the training data set. This must be represented in a formal language to be interpreted by the computer [144]. Additionally, choosing the representation of the learner defines the set of classifiers it can learn. This is known as the *hypothesis space* of the learner. The evaluation function distinguishes accurate classifiers from inaccurate ones [144]. The function used internally by the machine learning algorithm may differ from the external one the classifier is intended to optimise, for ease of optimisation. The optimisation technique provides a method to search among classifiers in the language in order to identify the highest scoring classifier [144]. It is integral to the efficiency of the learner and determines the classifier produced if the evaluation function produces more than one optimum.

Statistical and machine learning techniques have been used with great success to detect fraud in financial reporting [148]. They detect fraud in credit card transaction data [149], construct spam email detectors [150] and solve fraud detection problems [151]. Their success is partly due to the fact that machine learning models can be trained on historical data access behaviours to identify future abnormal patterns [94]. This is known as a supervised learning approach.

3.2.2. Supervised Learning

In supervised learning, the goal is to map input variables $X_1, \ldots X_p$ to output variables Y. A sample is defined as a pair of values $([X_1, \ldots X_p]^T, y)$ of these variables [158]. A pair of matrices is often used to represent the data set, one for input values and one for output values, with each row corresponding to a sample of the dataset, and each column to a variable [158]. In supervised anomaly detection approaches, a set of labelled training instances are provided, typically in the form of *anomal*y and *non-anomaly* [126]. The instances are then trained using a classification model based on their variable features. The resulting models are used to classify new actions. Supervised machine learning models, such as SVMs, linear regression and logistic regression have been successfully applied to the challenge of detecting illegitimate access within EPR systems [94]–[96]. A clearly labelled training dataset, however, is too resource intensive to generate for EPRs, particularly in the context of a dynamic, evolving environment [126].

Classifier detection determines a classification function based on a labelled training set [160] and can be fast, accurate and assign risk scores to all events [154]. 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 [161]. A classifier is a machine learning system which inputs a vector of discrete and/or continuous *feature values* and outputs a single discrete value, the *class* [144]. A *learner* inputs a training set of examples, with the observed input and corresponding output, and outputs a classifier. The purpose of the test is to analyse whether the classifier correlates with the corresponding output, in order to determine the accuracy for future data input, for which the corresponding output is not known [144].

Traffic analysis is a common application of a classifier detection method. Traffic classification techniques aim to identify flow patterns within a data stream, including anomalies sent or received by a host on the network [73]. Online traffic classification serves as the input for practical solutions such as network monitoring, quality-of-service and intrusion detection [162]. Machine learning algorithms have the capability of accurately classifying 99.8% of TCP network traffic performance, such as accuracy, throughput and latency [162]. Machine learning is desirable over port and signature based systems due to the capability of identifying encrypted flows or flows using irregular ports, and identifying previously unknown applications [162]. Network Traffic Flow Classification can utilise visualisation techniques to assist an analyst in discovering the root cause of network malfunction, showing normal and abnormal data activities in parallel and classifying traffic with 92% accuracy [163].

Signature detection are rules-based algorithms that construct a set of rules based on historic breaches and can detect correctly known patterns whilst being interpretable [164]. However, it cannot detect unseen patterns and cannot assign risk scores [165].

3.2.2.1. Linear Discriminant Classifier (LDC)

Within the classification domain, LDC is one of the most rudimentary approaches. The algorithm sorts or divides data into groups based on its characteristics in order to create a classification [166]. It performs an ordered transformation of unknown quantities, which are separated by a linear vector.

LDC consists of statistical properties of data which are calculated for each class [167]. In addition to the assumption of normally distributed classes, the LDC assumes equal class covariance matrices [168]. The LDC is derived as the minimum-error, or Bayes, classifier for normally distributed classes with equal covariance matrices [169]. In other words, for a single input variable, the mean and variance of the variable for each class is calculated. For multiple variables, the mean and variance are calculated over the multivariate Gaussian, which assumes that each variable is shaped like a bell curve when plotted [167]. These statistical properties are estimated from the data and then the LDC equation is performed in order to make predictions.

The Linear Discriminant function can be defined as [170]:

$$\delta_k(x) = x^T \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k^T \Sigma^{-1} \mu_k + \log(\pi_k)$$

(3.1)

3.2.2.2. *Quadratic Discriminant Classifier (QDC)*

Unlike the LDC approach, the Quadratic Discriminant Classifier (QDC) divides data using a quadratic surface, rather than a one-dimensional one, into groups based on its characteristics [166]. QDC assumes that each class has its own covariance matrix and assumes that the changing of two random variables will not be the same [167]. Comparatively to LDC, QDC uses supervised learning to separate data using a curved line. QDC is therefore more flexible than LDC and is a better fit for large training sets. Whereas LDC is a better fit if there are few training observations and reducing variance is an important factor.

The QDC can be defined as [170]:

$$\delta_k(x) = -\frac{1}{2}\log|\Sigma_k| - \frac{1}{2}(x - \mu_k)^T \Sigma_k^{-1}(x - \mu_k) + \log(\pi_k) \tag{3.2}$$

3.2.2.3. *Genetic Algorithms*

Genetic algorithms are evolutionary algorithms intended to obtain more accurate solutions as time progresses [152]. The algorithms encode a set of parameters to define a set of potential solutions to a problem as a binary string, referred to as a chromosome (or genotype), and apply recombination and mutation operators to the structures so as to preserve critical information and improve the utility/objective/fitness function [171]. A number of initial solutions are generated (which act as 'parents'). Crossover and mutation operators are applied and new solutions are then generated, with the stronger solutions remaining and the weaker solutions being eliminated [152]. This process continues until the convergence criteria is fulfilled, which may lead to the "near"-optimum solution, and in some simple problem or exceptionally well-designed GA, the best solution can be found. There is therefore potential for the application of genetic algorithms to the field of anomalous access behaviour detection within healthcare infrastructures.

3.2.3. Unsupervised Learning

In unsupervised anomaly detection approaches, the inherent structure, or patterns in a dataset, are utilised in order to determine when a particular instance is sufficiently different [126]. Unsupervised techniques, such as k-nearest neighbour anomaly detection, are designed to measure the distances between instances using features such as social structures [172]. Anomaly detection compares incoming instances to previously built profiles and can detect novel patterns, although it requires a large quantity of historic data [173]. Additionally, the output is known to be problematic to interpret and the technique produces false positives [174].Clustering is invoked to integrate similar data instances into groups [175].

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 [126].

3.2.3.1. Local Outlier Factor (LOF)

Nearest neighbour anomaly detection techniques are designed to measure the distances between instances using features such as social structures [172]. They determine how similar an instance is to other nearby instances, and if the instance is not sufficiently similar it is classified as an anomaly. The Local Outlier Factor (LOF) process involves five stages [26]:

i) *k*-distance computation: Euclidian distance of the *k*-th nearest object from an object **p** is computed and defined as k-distance, where a user defined parameter *k* is the number of nearest neighbours.

ii) *k*-nearest neighbour set construction for **p**: Set *k*NN(**p**) is constructed by objects within *k-distance* from **p**.

iii) A *reachability distance* computation for **p**: *Reachability distance* of **p** to an object **o** in *k*NN(**p**) is computed as which is defined as follows:

$$reach - distk(\boldsymbol{p}, \boldsymbol{o}) \ = \ max\{k - distance(\boldsymbol{o}), d(\boldsymbol{p}, \boldsymbol{o})\} \tag{3.3}$$

where d(**p**, **o**) is Euclidian distance of **p** to **o**.

iv) *lrd* computation for **p**: Local reachability density (*lrd*) of **p**, defined as follows:

$$lrd_k(\boldsymbol{p}) = \frac{k}{\boldsymbol{\Sigma}_{0 \in kNN(\boldsymbol{p})} reach - distk(\boldsymbol{p}, \boldsymbol{o})}$$

(3.4)

v) LOF computation for **p**: Local Outlier Factor of **p** is computed defined as follows:

$$LOF(\boldsymbol{p}) = \frac{\frac{1}{k}\boldsymbol{\Sigma}_{0 \in kNN(\boldsymbol{p})} lrd_k(o)}{lrd_k(\boldsymbol{p})}$$

(3.5)

The LOF process exposes anomalous data points by measuring the local deviation. In other words, patterns in data that do not conform to the expected behaviour are revealed. In the case of EPR data, employing the LOF process can be effective in that it recognises points, which are outliers from similar/related points in one area of the dataset. Therefore, the algorithm is particularly applicable to a dataset, where multiple job types/roles are present. It considers the relative density of points and can detect data in biased datasets. This means that it is advantageous over proximity-based clustering. LOF employs the relative-density of a coefficient against its neighbours as the indicator of the degree of the object being an outlier [176].

If a global outlier is employed, the detection of irregular behaviours would not be possible without correlating the different hospital roles with each other; adding an extra stage to the detection process – one which might not be possible. This is due to the process that a global outlier detection process undertakes in identifying data points that are far from other points in the dataset.

3.2.3.2. LOOP

Local Outlier Probability (LoOP) is a method that derives a local outlier factor which provides an outlier score in a range of 0 to 1 [177]. The advantage of this method is the outlier is interpretable as a probability of a data point being an outlier, rather than interpreting a threshold for outliers as in LOF.

The Local Outlier Probability (LoOP), indicating the probability that a point $o \in \mathrm{D}$ is an outlier is defined as [178]:

$$LoOP_s(o) \coloneqq max\left\{0, erf\left(\frac{PLOF_{\lambda}, s(o)}{nPLOF \cdot \sqrt{2}}\right)\right\}$$

(3.6)

Where the Probabilistic Local Outlier Factor (PLOF) can be defined as follows:

$$PLOF_{\lambda,\mathrm{S}}(o) \coloneqq \frac{pdist\big(\lambda, \mathrm{o}, \mathrm{S}(\mathrm{o})\big)}{E_{s\epsilon S(o)}\big[pdist\big(\lambda, \mathrm{s}, \mathrm{S}(\mathrm{s})\big)\big]} - 1$$

(3.7)

And a normalised PLOF (nPLOF) can be defined as:

$$nPLOF \coloneqq \lambda \cdot \sqrt{E[(PLOF)^2]}$$

(3.8)

3.2.3.3. Collaborative Filtering

Collaborative filtering is a dyadic prediction method, where the task is to predict a label for the interaction of a pair of entities [94]. Within a hospital setting, these entities would be the system user, and the patient record. Collaborative filtering approaches for detecting unauthorised access to EPR data have been successful in recognising the identity of users and patients involved in patient record access [94]. Through the use of explicit and latent features for staff and patients, the following scenarios can be understood to be more likely to be involved in a future violation 1) a patient, whose record has previously been involved in a violation, or 2) a staff member who has performed a violation in the past [94]. In addition to the use of latent features of a dataset to *fingerprint* a user based on historical access data, collaborative filtering can collate data for reliable parameter estimation and create interaction-specific predictions [94].

A latent feature model of collaborative filtering where a is the suspiciousness of access can be modelled as [178]:

$$\hat{y}(a;\theta) = f(w^T\phi(a) + \alpha_u^T\beta_p + \gamma_u + \delta_p + \mu)$$

(3.9)

For an appropriate link function $f(\cdot)$. Including a global bias μ, user and item-specific biases γ_u and δ_p. a_u and β_p compromise latent features of user and patient. Additionally, $w^T \phi(a)$ leverages information present in explicit features $\phi(a)$.

3.2.3.4. DBSCAN

Density-based spatial clustering of applications with noise (DBSCAN) is a cluster analysis method [179]. DBSCAN divides a dataset into *n* dimensions and forms an *n*-dimensional shape around each datapoint creating data clusters. The clusters are then expanded by including other datapoints within the cluster and adding their *n* dimensions in the cluster. It requires two parameters. 1) ε – the minimum distance between two points to be considered neighbours and 2) *MinPoints* – the minimum number of points which form a dense region. Any datapoints that do not fall within a cluster can be handled as an outlier. DBSCAN is often compared with LOF as an outlier model even with large scale analysis [180].

DBSCAN can be presented as Pseudocode as [179]:

Algorithm 3.1 – DBSCAN Pseudocode

```
DBSCAN (SetOfPoints, Eps, MinPts)
// SetOfPoints is UNCLASSIFIED
   ClusterId := nextId(NOISE);
      FOR i FROM 1 TO SetOfPoints.size DO
          Point := SetOfPoints.get(i);
          IF Point.CiId = UNCLASSIFIED THEN
             IF ExpandCluster(SetOfPoints, Point,
                   ClusterId, Eps, MinPts) THEN
             ClusterId := nextId(ClusterId)
          END IF
      END IF
   END FOR
END; // DBSCAN
```

3.2.3.5. K-*Nearest Neighbour*

k-nearest neighbours (*k*-NN) is a non-parametric algorithm for classification and regression [183]. For classification, an output of class membership is given, where an object is assigned to a class most common among its neighbours. For regression, an output of a property value is given, which is the average of the values of its neighbours. *k*-NN is often used when little is known about the data as the model structure is determined by the data. *k*-NN uses instance-based learning and therefore defers generalisation of the data until classification. *k*-NN is often used as an outlier score for anomaly detection, as the larger the distance to the *k*-NN and the lower the local density, the more likely a datapoint is an outlier [184].

The k-Nearest Neighbour algorithm can be presented as Pseudocode as [179]:

Algorithm 3.3 – k-Nearest Neighbour Pseudocode

```
Where $W = \{x_1, x_2, \cdots, x_n\}$ is a set of n labeled samples.
BEGIN
    Input y, of unknown classification.
    Set $K, 1 \leq K \leq n$.
    Initialise $i = 1$.
    DO UNTIL (K-nearest neighbours found)
        Compute distance from y to $x_i$.
        IF ($i \leq K$) THEN
            Include $x_i$ in the set of K-Nearest neighbours
        ELSE IF ( $x_i$ is closer to y than any previous nearest neighbour) THEN
            Delete farthest in the set of K-Nearest neighbours
            Include $x_i$ in the set of K-Nearest neighbours.
        END IF
        Increment i.
        END DO UNTIL
        Determine the majority class represented in the set of K-Nearest neighbours.
        IF (a tie exists) THEN
            Compute sum of distances of neighbours in each class which tied.
        IF (no tie occurs) THEN
            Classify y in the class of minimum sum
        ELSE
            Classify y in the class of last minimum found.
        END IF
    ELSE
        Classify y in the majority class.
    END IF
END
```

3.2.3.6. OPTICS

Ordering Points to Identify the Clustering Structure (OPTICS) is a cluster analysis method [181]. Like DBSCAN, OPTICS considers both *ε* and *MinPoints* parameters. However, like LOF, OPTICS also includes 1) a *core distance* – the minimum epsilon which makes a distinct point a core point and 2) a *reachability distance* – the distance of object *x* to object *y* is the smallest distance from *y,* if *y* is a core object and cannot be smaller than the core distance of *y*. OPTICS-OF is an outlier detection method based on OPTICS and is similar to LOF in application and concept [181].

OPTICS can be presented as Pseudocode as follows [182]:

Algorithm 3.2 – OPTICS Pseudocode

```
OPTICS (SetOfObjects, ε, MinPts, OrderedFile)
    OrderedFile.open();
    FOR i FROM 1 TO SetOfObjects.size DO
        Object := SetOfObjects.get(i);
        IF NOT Object.Processed THEN
            ExpandClusterOrder(SetOfObjects, Object, ε,
                MinPts, OrderedFile)
    OrderedFile.close();
END; // OPTICS
```

3.2.3.7. TF-IDF

TF-IDF is based on the intuition that a query term which appears in many documents is not a 'good' discriminator, and should be given less weighting than a query term which occurs in fewer documents [185]. TF is based on the idea that the frequency of a term within a document itself is an indicator of its importance. Therefore an IDF measure multiplied by a TF measure has generically become known as TF-IDF, and this class weighting has proven robust and difficult to improve upon [185].

TF-IDF is a numerical statistic which attempts to define the relevance of a word within a document in a corpus, and is used as a weighting factor in information retrieval and text mining [186]. Term Frequency f measures how frequently a term t appears in a document. Inverse Document Frequency measures how important a term is.

Therefore

$$Inverse\ Document\ Frequency\ (t) = Log\ \frac{N}{N}$$

(3.10)

Where, *N* is the total number of documents and *Nt* is the number of documents with term *t* [186].

3.2.4. Ensemble Averaging

Committee methods operate on the principle that combining the output of a group of machine learning algorithms can achieve a decision function superior to any individual output [187]. Ensemble averaging is a committee method in artificial neural networks that averages the output of a collection of outputs [188]. Ensemble averaging concerns the following two properties of artificial neural networks [189]. Firstly, in a network, bias can be reduced at the cost of increased variance. Secondly, in a group of networks, the variance can be reduced at no bias cost.

The ensemble average can be calculated through the following, where each expert is y_i, and the overall result $\tilde{y}$ can be defined as [190].:

$$\tilde{y}(x;\alpha) = \sum_{j=1}^{p} \alpha_j y_j(x)$$

(3.11)

For a given input, x, the output of the combined model, $\tilde{y}$, is the weighted sum of the corresponding outputs of the component neural networks, y_j, j = 1,⋯,p, and the α_j's are the associated combination-weights [190].

In the context of outlier detection, this process has been described as Feature Bagging [191]. Through combining the results of multiple outlier detection algorithms using different sets of features.

3.2.5. Human-in-the-Loop

A HILML model must be able to generalise across use-cases and accept a declarative or semi-declarative specification [192]. Due to the declarative specification, a HILML system should capture a model of a Directed Acyclic Graph (DAG) of intermediate data items. Through a declarative specification, HILML can identify the logical operator for each node in the workflow, such as data preparation or model training [193].

The key advantages of an HILML model are as follows [141]:

1. Overcoming limited analyst bandwidth: An analyst can only feasibly examine less than 1% of the overall event volume. Therefore, the use of outlier detection can present the most pertinent events for investigation.

2. Overcoming weaknesses of unsupervised learning: An events rarity, or status as an outlier, does not necessarily constitute maliciousness. Therefore, an events score does not capture intent. Using an HILML model can include an analyst's subjective assessment of malicious intent.

3. Actively adapts and synthesises new models: Analyst feedback provides labelled data regularly, creating a positive feedback loop. The more attacks the machine learning mode detects, the more feedback it receives from an analyst, which then improves the accuracy of future predictions.

3.2.6. Algorithm Comparisons

A comparison table of outlier detection algorithms is presented in Table 7 along their advantages and disadvantages.

Table 7 - Outlier Detection Algorithms

Detection Algorithms	Advantages	Disadvantages
Local Outlier Factor (LOF)	Uses a local approach, which is able to identify outliers in data that would not be an outlier in	Determining a threshold for anomaly score representing an outlier varies

	another area of the data. Therefore, a point with a small distance to a dense cluster can be an outlier, whereas a point in a sparse cluster may be an inlier. Can be generalised to multiple datasets. Often outperforms other algorithms, such as in network intrusion detection [194] and on classification benchmark data [180].	between datasets.
Local Outlier Probabilities (LoOP)	Scales resulting values to a range of (0,1) which can be useful in some applications.	The requirements for detecting anomalous data behaviours require determining an outlier factor threshold for anomalies, rather than being returned in a probability.
Density-based spatial clustering of applications with noise (DBSCAN)	Can determine high- and low-density clusters within a dataset. Can identify outliers in a dataset that do not belong to clusters. No requirement to define a number of clusters.	More applicable to cluster analysis data applications rather than anomaly detection. Clusters with varying densities cannot be easily identified, only high and low densities.
Ordering points to identify the clustering structure (OPTICS)	No requirement to define a radius as in DBSCAN, instead using a priority queue for unprocessed clusters. Can identify clusters of differing densities and no requirement to define the number of clusters.	More applicable to cluster analysis data applications rather than anomaly detection. Slower than DBSCAN.
K-Nearest Neighbour	Very fast to process. Requires no training to make predictions due to instance-based learning. Makes no assumptions regarding the dataset. Often outperforms other algorithms, such as on classification benchmark data [180].	Sensitive to irrelevant features and large datasets.

Based on Table 7, LOF and DBSCAN will be compared in the case study Chapter 6. Both LOF and DBSCAN have the common parameter of neighbourhood size *k* and can identify outliers within a dataset.

3.3. Graphical Applications of Machine Learning

Due to the complex nature of machine learning models, most security approaches involve some visualisation aspect. Visualising complex data facilitates a more comprehensive stage for conveying knowledge. In doing so, machine learning models become more practicable. Within the medical data domain, there is an increasing requirement for valuable and accurate information. Visualisation techniques can, therefore, be used to provide both awareness and modelling capabilities for the benefit of an infrastructure [195]. It is proven that data visualisation enables meaningful inferences to be extracted from raw data, and facilitates cost savings and faster decision support [196]. Detecting anomalous data behaviours in healthcare infrastructures is challenging. Visualisation brings together several related data sets and presents them in such a way as to identify relationships between them. To achieve this, cyber-threat monitoring systems employ IDSs as network sensors [197]. IDSs statistically analyse the time of the attack, the source of the attack, the destination of the attack, and can visualise the result. Visualisations are then used in order to leverage the perceptual abilities of the user, in order to find features in network structures and data.

Visualisation provides a framework for providing situational awareness at both i) a high-level, enabling an overview of data flows within a hospital infrastructure and ii) a low-level, enabling a detailed view for the investigation of a specific interaction within the hospital. There are three levels of situational awareness [198]; i) user must be able to identify all the information relevant to a decision being made, i.e. objects, identities; ii) user must be able to use the situational awareness model to make connections in order to understand what is happening; iii). the user must be presented with the information in such a way so as they can process it in order to accurately make predictions about events that may occur in the near future.

Figure 14 demonstrates cyber-situational awareness as a three-phase process. Situation recognition, situation comprehension and situation projection [199]. These phases can further be broken down into seven aspects. 1) Situation Perception, 2) Impact Assessment, 3) How Situations Evolve, 4) Attacked Behaviour, 5) How the Current Situation is Caused, 6) Quality and Trustworthiness of Information and 7) Assessing Plausible Features. Situation Recognition incorporates aspects 1-6, Situation Comprehension includes aspects 2, 4 and 5, finally Situation Projection includes aspects 3 and 7. Figure 14 is a visual representation of the overlap between the three phases and seven aspects of cyber-situational awareness. For example, all aspects of Situation Comprehension phase are also incorporated into the Situation Recognition phase.

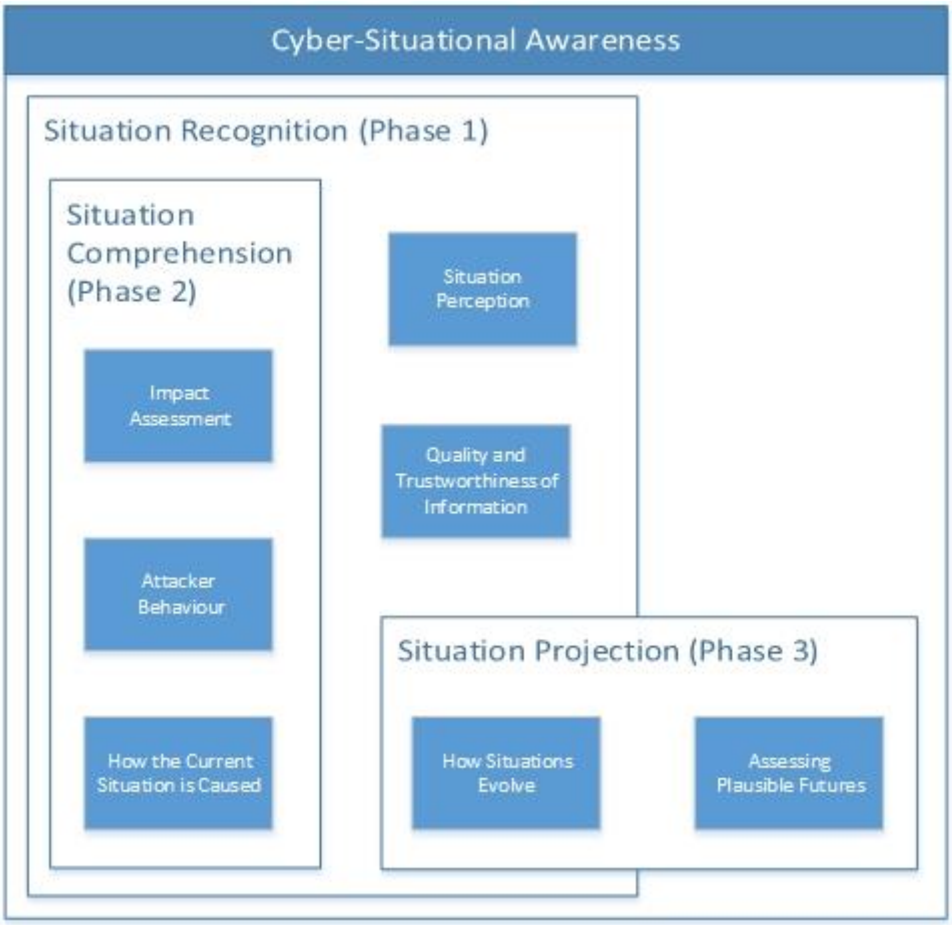

Cyber-Situational Awareness
Situation Recognition (Phase 1)
Situation Comprehension (Phase 2)
Situation Perception
Impact Assessment
Quality and Trustworthiness of Information
Attacker Behaviour
Situation Projection (Phase 3)
How the Current Situation is Caused
How Situations Evolve
Assessing Plausible Futures

Figure 14 - Cyber-Situational Awareness

Firstly, the user must be aware of the current situation, also known as *Situation Perception,* incorporating both situation recognition and identification. Situation perception includes identifying the type of attack, the source of an attack and the target of an attack. Secondly, the user must be aware of the impact of the attack, also known as *Impact Assessment*. Impact assessment involves the assessment of the current impact and damage, and the assessment of the future impact and potential damage. Thirdly, the user must be aware of how situations evolve, often heavily involving situation tracking as a major component. Fourthly, the user must be aware of attacker behaviour. This is achieved through attacker trend and intent analysis, which analyses attacker behaviours within situations. Fifthly, the user must be aware of how and why the situation is caused, including causality analysis and forensics. Sixthly, the user must be aware of both the quality and trustworthiness of the situational awareness information items. Quality metrics used to assess this include truthfulness, completeness and freshness. Finally, the user must be able to assess the plausible futures of the situation.

3.3.1. The Theory of Gamified Learning

Research has shown that gamification can be utilised to enable situational awareness [147]. The Theory of Gamified Learning infers that gamification can positively affect learning and decision making through a more direct mediating process and a less direct moderating process [200]. Gamification affects learning via mediation when a user's behaviour is encouraged in such a way that it improves learning outcomes itself, such as a fitness app [201]. The theory therefore mediates the relationship between game elements and learning. For the next stage, gamification affects learning via moderation when pre-existing information is improved through strengthening the relationship between instructional design quality and outcomes [202]. For the moderation theory, the moderator does not influence the outcome construct independently of the causal construct, therefore the pre-existing information must be of high quality, or the addition of gamification techniques would be of no benefit. The Theory of Gamified Learning is outlined in Figure 15 [200].

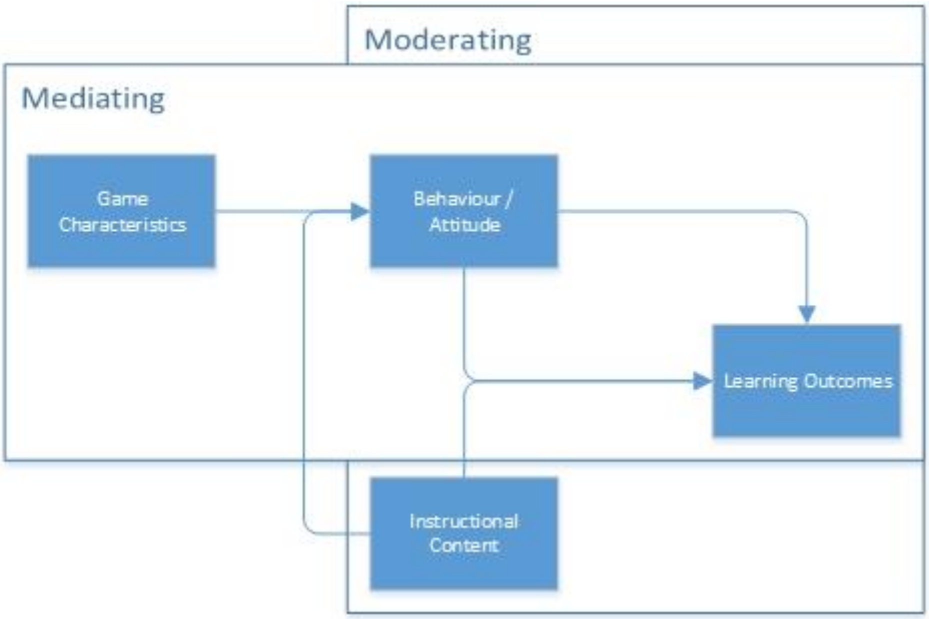


Figure 15 - The Theory of Gamified Learning

The Mediating process follows the flow of 1) Game Characteristics, 2) Behaviour/Attitude and finally 3) Learning Outcomes. The Moderating process is described by the influence of 1) Behaviour/Attitude on 2) Instructional Content to 3) Learning Outcomes.

3.3.2. Feature Testing

Given the mean expressed previously in (15), the scatter matrix is the *m*-by-*m* positive semi-definite matrix. Where *T* denotes matrix transpose, and multiplication is with regards to the outer product [203], as expressed in (19).

$$S = \sum_{i=1}^{m} (x_i - \mu)(x_i - \mu)^T = \sum_{i=1}^{m} (x_i - \mu) \otimes (x_i - \mu)^T = \left(\sum_{i=1}^{m} x_i x_i^T \right) - m\mu\mu^T$$

(3.12)

A scatter matrix visualises the relationship between the features to predict the most appropriate for the machine learning classification. The scatter matrix displays the positive and negative correlation between the features. Figure 16 outlines the feature testing process. Through testing the features in this way, the most pertinent features can be selected and applied [204].

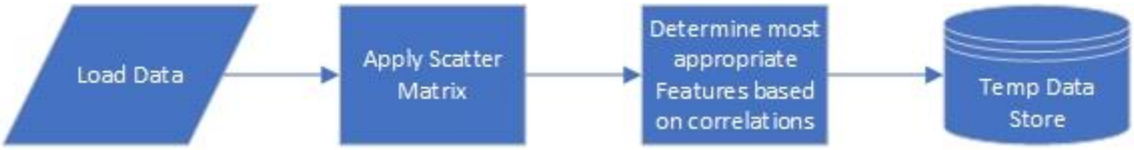


Figure 16 - Feature Testing Process

3.4. Related Applications of Machine Learning

This section presents an overview of related works which incorporate machine learning (or advanced analytics) techniques in order to provide situational awareness of information security issues within audit logs. Four peer-reviewed applications are presented, with a discussion of strengths and weaknesses in relation to PARISS. Finally, a table featuring commercial solutions is presented with a description and a list of limiting factors.

3.4.1. SIEM

Security Information and Event Management (SIEM) systems are distributed systems, which collect and process logs generated by both network hardware and software assets, and perform real-time and centralised event analysis [205]. In doing so, event correlation mechanisms are implemented by the analysis server to identify the occurrence of malicious actions and foresee an attack. However, there are a number of issues present in current SIEM solutions. Specifically, current SIEM solutions have processing constraints which limit the effectiveness of discovering violations within the business logic. Additionally, SIEMs cannot process data at the edge of the deployed architecture. This presents limits in addressing data disclosure and privacy issues, a particularly relevant problem within large scale deployments. Finally, no mechanisms are provided to improve the dependability of data storage systems that contain evidence of security breaches and maintain and store the sensitive data of involved parties [205].

3.4.2. MAP1/MAP2

Monitoring Access Pattern Phase 1 (MAP1) attempts to identify illegitimate access to EPRs and score each access for appropriateness, so the top scoring cases can be prioritised and investigated by privacy officers. The production of scores indicating suspiciousness of access is preferable to simple procedure-based patterns. A training set is created through labelling selected events as either suspicious or appropriate by privacy officers. LR and SVM models is trained on 10-fold cross-validation sets of 1,291 labelled events [95] MAP1 demonstrates that statistical and machine learning methods can assist in identifying potentially illegitimate accesses to EPRs [95]. MAP2 is an extension of the work of MAP1 and relates to fine-tuning the detection algorithm [96]. MAP2 focuses on the construction of classifiers with appropriate filtering techniques to detect rare events. MAP2 uses a combination of Signature detection, Anomaly detection and Classifier detection, extending the capabilities of the previous MAP1 classifier algorithm. Privacy officers identified 78 illegitimate accesses to the EPR during the study period, and MAP2 identified 75 of those accesses independently, demonstrating that the technique has the capability to facilitate the detection of rare, but important events [96].

3.4.3. CADS

Community Based Anomaly Detection Systems (CADS) is an unsupervised learning framework to detect insider threats based on information recorded in audit logs of collaborative environments [126]. 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 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 [126]. CADS does not implement supervised learning techniques to further classify the data with feedback from patient privacy officers.

3.4.4. AI^2

AI^2 is a cyber-security machine learning system, which improves its accuracy over time through feedback from security analysts [141]. AI^2 is composed of the following four components. Firstly, a Big Data Processing System, which quantifies the behaviours and features of raw data [141]. 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 [141]. Thirdly, a Feedback Mechanism and Continuous Learning, which incorporates analyst input through a user interface [141]. 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 refine the model [141]. Raw data is input into AI^2 which 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 [141].

3.4.5. Commercial Solutions

Table 8 compares existing commercial solutions and their limiting factors.

Table 8 – Comparison to Commercial Solutions and their Limiting Factors

Commercial Solution	**Description** **(Taken from websites as most aren't peer reviewed)**	**Limiting Factors**
FairWarning Patient Privacy Intelligence	FairWarning is a Procedure-Based Analytics Patient Privacy Violation detection system [108] deployed in some UK hospitals. Fair Warning's Patient Privacy Intelligence is an open platform that secures patient data held in mission-critical applications. We give healthcare providers the tools and support they need to manage the full lifecycle of security incidents as required by regulations like HIPAA, HITRUST, and NIST.	Procedure-Based Analytics approach, rather than a machine learning approach. Lack of visualisation tools to provide situational awareness. Although deployed in NHS Trusts the system is tailored to HIPAA (US Health Insurance Portability and Accountability Act of 1996) compliance, rather than an ICO (UK Information Commissioner's Office) focus.
Protenus	Protenus uses artificial intelligence to help healthcare organizations protect patient privacy and secure health data [206].	Lack of visualisation tools to provide situational awareness. A focus on HIPAA compliance,

		rather than an NHS focus.
Iatric Security Audit Manager	Security Audit Manager iQ [206] empowers privacy auditors to: • See ranked suspicious activities automatically in personalized worklists, dynamically increasing productivity • Increase accuracy and ensure fewer false positives by combining our proven expert-based deterministic algorithms with machine-learning • Uncover patterns once difficult to find through role-based behavioural analysis • Detect, prevent, and respond to privacy incidents and breaches for a complete end-to-end solution	Lack of visualisation tools to provide situational awareness. A focus on HIPAA compliance, rather than an NHS focus.
Maize Analytics	Maize Analytics' Explanation-Based Auditing System (EBAS) reporting and filtering capabilities allow healthcare privacy officers to more quickly identify suspicious activities .	Lack of visualisation tools to provide situational awareness. A focus on HIPAA compliance, rather than an NHS focus.
Spher	SPHER is the front line defence against the day-to-day threat of patient privacy violations (PHI data breaches) resulting from inappropriate access to Protected Health Information [206]. As required by HIPAA, HITECH and MACRA, every comprehensive compliance strategy must include User Activity Monitoring, a requirement that SPHER is specifically designed to achieve.	Lack of visualisation tools to provide situational awareness. A focus on HIPAA compliance, rather than an NHS focus.
Intruno	Intruno uses advanced behavioural analysis to provide the ultimate intelligent notification against data breaches and privacy violations originating from both stolen credentials by external hackers and malicious insiders [206].	Lack of visualisation tools to provide situational awareness. A focus on HIPAA compliance, rather than an NHS focus.
Arcus Data	As electronic health records (EHRs) steadily replace paper records, healthcare institutions struggle to prevent security breaches without resorting to laborious manual audits of EHR accesses. Arcus Data dashboards and reporting capabilities allow healthcare privacy officers to more quickly identify suspicious activities [206].	Lack of visualisation tools to provide situational awareness. A focus on HIPAA compliance, rather than an NHS focus.
PatternEx	PatternEx is the commercial solution for AI^2, as	Lack of focus on healthcare

	described in 3.4.4. PatternEx delivers Artificial Intelligence, combining Analyst Intuition with machine learning to defend the enterprise against cyber security threats [206].	infrastructures. Lack of focus on EPRs.
SolarWinds	Improve security and compliance with an easy-to-use, affordable SIEM tool [206]. Detect suspicious activity - Identify threats faster with event-time detection of suspicious activity. Mitigate security threats - Conduct security event investigations and forensics for mitigation and compliance with SolarWinds SIEM software. Regulatory compliance readiness - Demonstrate compliance with audit-proven reporting for HIPAA, PCI DSS, SOX, DISA STIG, and more. Maintain continuous security - Improve security measures with SolarWinds® Log & Event Manager (LEM) SIEM tool, a hardened virtual appliance with encryption capabilities for data in transit and at rest, SSO/smart card integration, and more.	Lack of focus on healthcare infrastructures. Lack of focus on EPRs.
DarkTrace	DarkTrace [207], based in the UK, is among the world's most advanced machine learning technologies for cyber defence and an advocate for using AI for safeguarding critical systems.	Lack of focus on healthcare infrastructures. Lack of focus on EPRs.

3.5. Summary

A proactive approach to data confidentiality is required to safeguard EPR systems. Visualisation and machine learning techniques have the potential to enhance situational awareness of patient privacy violations within EPRs. Visualisations allow the user to explore the data and understand the patterns and trends within the comprehensive EPR audit data sets. Unsupervised machine learning techniques are able to classify this data as there is limited abnormal data and a lack of labelled training data. Feedback from the analysts can 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 Procedure-Based Analytics. Through detecting abnormal behaviours using machine learning and visualising the results, patient privacy officers can avoid the ‘needle in a haystack’ challenge of detecting this activity manually.

In this chapter, artificial intelligence and machine learning techniques are described in order to determine a suitable density-based outlier detection algorithm for the novel system framework. Additionally, related applications and work in this area are described, in both an academic and commercial context, detailing their use and limitations. In chapter 4, the proposed system framework is presented and detailed.

4. Proposed System Framework

4.1. Introduction

There is a clear need for an anomaly detection system to ensure patient confidentiality within EPR systems. PARISS is an information security system, which improves its accuracy over time through feedback from security analysts. The system assists information security officers within healthcare organisations to improve the situational awareness of patient data confidentiality risks. The issues of scalability require the system to be deployed on a cloud domain due to the requirements of storage of the EPR audit data and the processing of the algorithms. 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.

4.2. System Development Life Cycle

The novel contribution in this framework involves the use of advanced data analytics techniques, a Human-In-The-Loop (HILML) and the use of visualised attack events. Low-risk data is analysed, processed and pre-filtered using advanced data analytics techniques. The output 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 with an overall understanding of system behaviour in order to detect security breaches and general anomalies. The system provides situational awareness to detect anomalous behaviour within EPR audit activity.

4.2.1. Requirements Specification

EPR audit data is input to PARISS and features are extracted and scaled. The data is processed using the technique, Local Outlier Factor (LOF). The results are visualised to highlight potentially malicious activity in a broad overview of the activity within the EPR. From here the analyst explores the visualisation and the highlighted data points.

Interaction with the visualisation allows for in-depth exploration of the data, providing detailed technical information regarding the data points. This provides insight as to why the data is ranked as potentially malicious by LOF. HILML techniques are implemented to compliment the unsupervised LOF scores. Upon investigation an analyst can label the data as legitimate or illegitimate. Through this process, a semi-supervised approach is applied to the challenge of detecting EPR misuse. In doing so a feedback loop is embedded in the system to continuously improve PARISS.

4.2.2. Location

Figure 17 demonstrates where PARISS is situated within the hospital infrastructure layout. Data from the EPR (and any other relevant hospital systems if the EPR is not a comprehensive and holistic view of patient record activity) is extracted and input into the Data Warehouse. This is hardcoded and occurs daily at 4AM as this is statistically the period of quietest activity within a hospital. Data from the data warehouse is then extracted and input to PARISS.

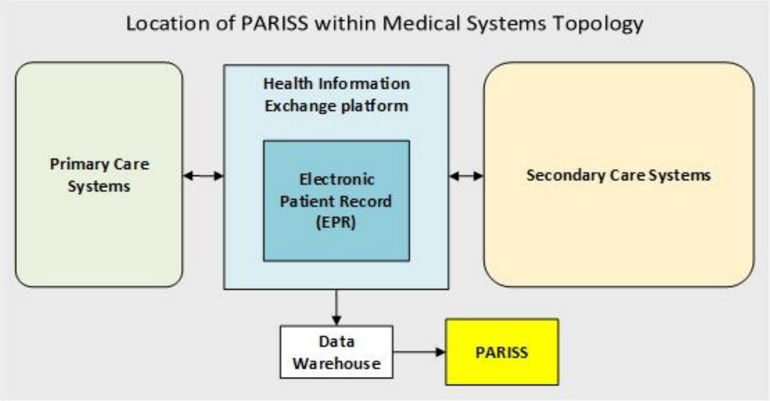

Location of PARISS within Medical Systems Topology
Primary Care Systems
Health Information Exchange platform
Electronic Patient Record (EPR)
Secondary Care Systems
Data Warehouse
PARISS

Figure 17 - PARISS Location

4.3. Algorithm Design

The algorithm process steps are described. The novelty of this work is implementing previous techniques in a novel way. Broadly, the use of the Local Outlier Factor algorithm, in conjunction with a Human-in-the-Loop and a visualisation element, applied to the context of patient privacy in EPRs. Detailed diagrams of the system components are presented and discussed throughout the remainder of this Chapter.

1. *EPR Audit Data*: This audit data is captured by the EPR and records at every interaction with the EPR. This data is extracted into the data warehouse where it is stored and input into the PARISS pipe-delimited format on a daily basis.
2. *Feature Extraction*: Features of the EPR audit data are extracted for the analysis purposes. A statistical features-based approach is taken.
3. *User Profiling*: This phase is only executed on initial set up. Through profiling users, the anomaly threshold benchmark is set. This enables the selection of a local outlier factor anomaly threshold.
4. *Feature Scaling:* Features are scaled in order to conform to a common scale for the classification algorithms.
5. *Feature Testing:* Features are tested and selected to determine the most suitable features for the Local Outlier Factor algorithm.
6. *Local Outlier Factor*: LOF is applied to extract anomalies from the data. After the data has been processed by LOF, the data is ranked and selected based on previous user interaction. This is to ensure the most notable data points are presented to the user first.
7. *Quantifying LOF*: LOF results are quantified for Infinite (Inf) and Not a Number (NaN) values. Inf indicates a point that is next to identical points, but is not a member, and is therefore anomalous and assigned an anomalous numerical value of 2. NaN indicates many neighbours at the same location, therefore the ratio of densities is undefined and is not anomalous, so assigned a non-anomalous numerical value of 1.
8. *Visualisation*: This component generates the visualisation for the user. The component uses the system operator’s input and calls upon the data stored in the database component, which is then processed and visualised and passed onto the UI Output.

9. *Data Analyst*: The operator interacts with and manipulates the visualisation in order to set their own data parameters. This increases their situational awareness of the data flow within the healthcare infrastructure.
10. *Human-in-the-Loop Machine Learning (HILML)*: HILML techniques are applied in order to extract knowledge from the analysts. An analyst can rank a detected outlier on an accuracy scale, which sets a benchmark multiplier score for the IDs, affecting their anomaly score in future iterations.
11. *Data Storage*: This component stores the data in a database when not in use by the other components.

4.3.1. UML Diagram

The PARISS UML Diagram is detailed in Figure 18.

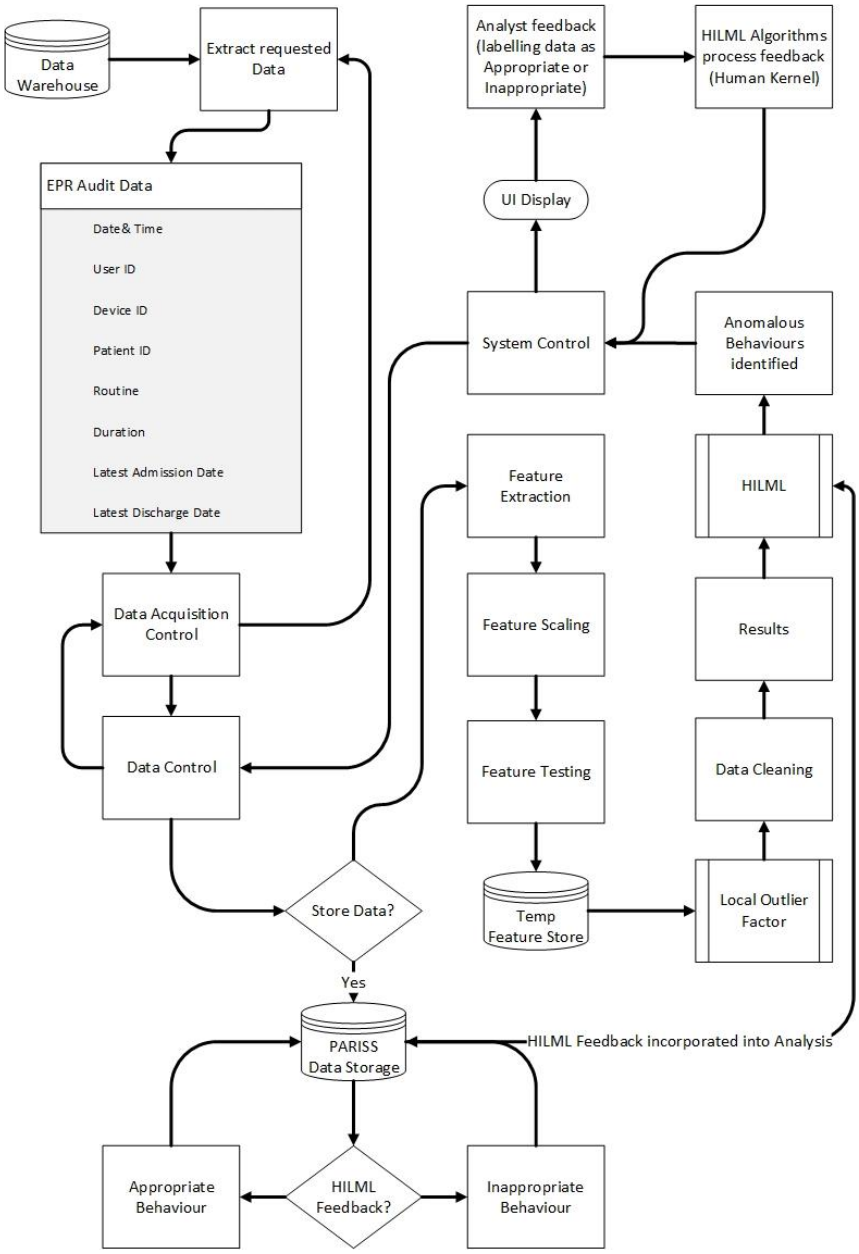


Figure 18. PARISS UML Diagram

Figure 18 presents the flow for how PARISS functions and processes the EPR audit data. Each component is discussed throughout the remainder of this section. Figure 18 highlights how the Data Acquisition Control function requests data from the data warehouse and functions with the Data Control. Figure 18 also displays the PARISS Data Storage incorporating both appropriate and inappropriate behaviour accesses as defined by the HILML feedback. This feedback is then retrieved during the analysis stage after Local Outlier Factor has processed a new batch of data.

4.3.2. Input Data

Input Data to PARISS is extracted from the Data Warehouse. Due to the disparate nature of systems in hospitals, a data warehouse is managed by the Business Intelligence team in order to extract insight and prevent data siloing. EPRs integrate many aspects of care into a single system, which are audited. Each encounter with patient data results in an audit footprint, which is stored in the data warehouse. Data is extracted into pipe delimited (|) values format and extracted into PARISS. A pipe is employed as it is rarely used in normal text within EPRs or numeric processing. This process is detailed in Figure 19.

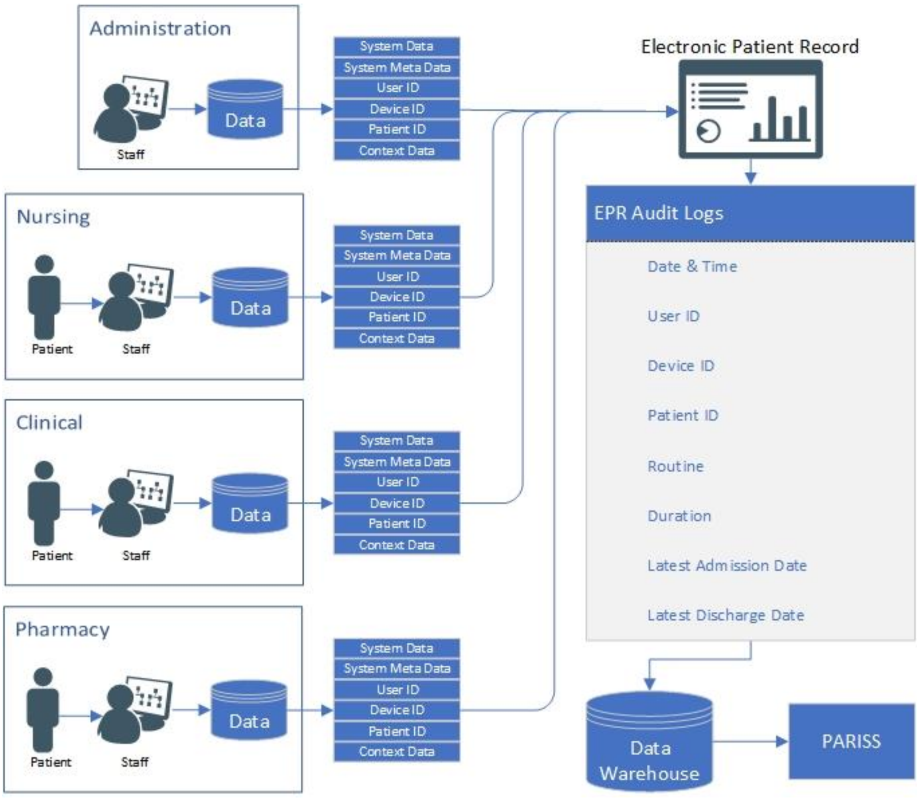


Figure 19 - PARISS Input Data

EPR Audit Logs consist of the following fields (with slight variance between EPRs):

- *Date & Time*: The date/time the patient record was accessed;
- *User ID:* The User ID who accessed the patient record;
- *Device ID*: The Device ID the patient record was accessed on;
- *Patient ID*: The patient ID record that was accessed;
- *Routine*: The routine performed whilst accessing the patient record (was the record updated, was a letter printed etc.);

- *Duration*: The number of seconds the patient record is accessed for (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 is last admitted to the hospital
- *Latest Dis Date*: The date the patient is last discharged from the hospital.

4.3.3. Feature Extraction

The PARISS Feature Extraction process is detailed in Figure 20.

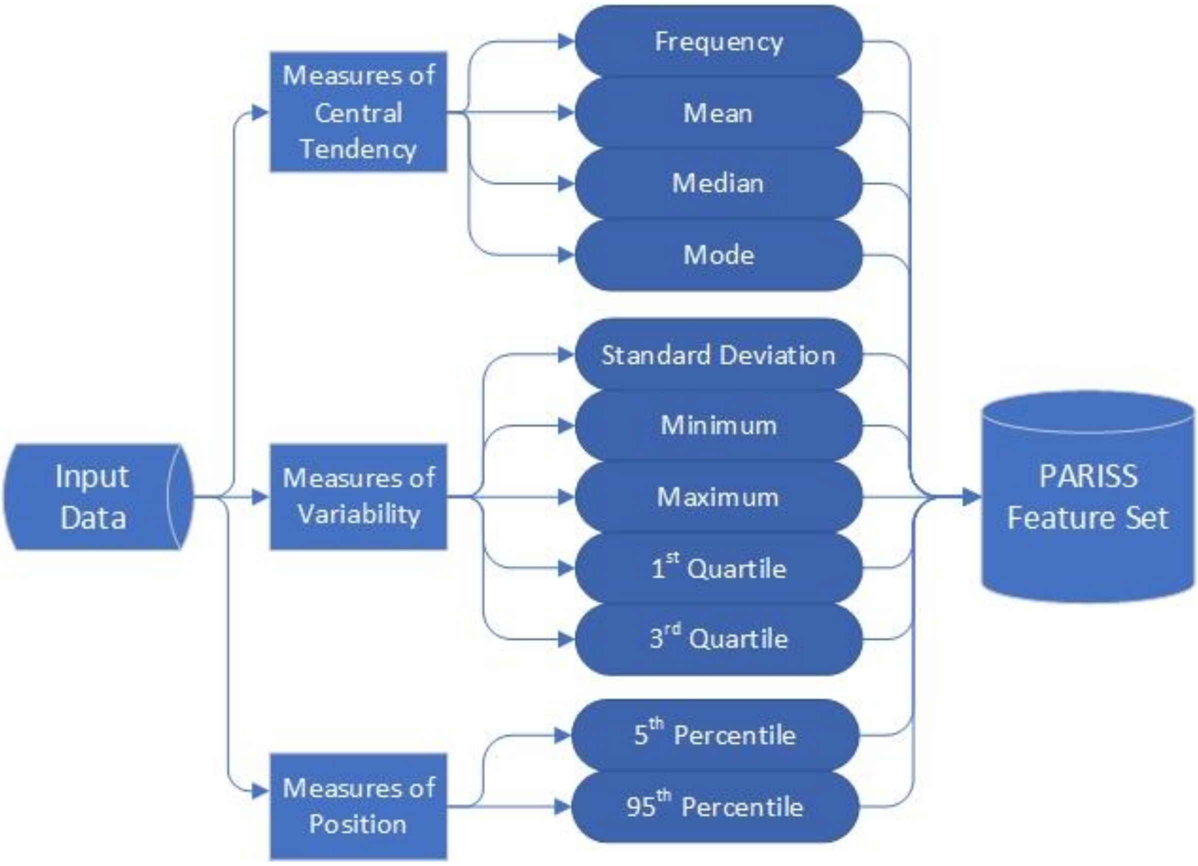


Figure 20 - PARISS Feature Extraction

Features of the EPR audit data are extracted for the LOF classification process. During the pre-processing stage, a statistical features based approach is implemented [208]. Four 'measures of central tendency' are calculated through the Frequency, Mean, Median and Mode feature extraction process. Five measures of variability are calculated through the Standard Deviation, Minimum, Maximum, 1st Quartile and 3rd Quartile features. Finally, two measures of position are calculated through the 5th Percentile and 95th Percentile features.

The resulting eleven features are extracted from the dataset for each ID (User, Patient, Device and Routine). Table 9 displays the features selected, with an accompanying description.

Table 9 - Dataset Feature Names and Descriptions

Feature Name	Feature Description
Frequency	The number of times the ID featured in the dataset (using a Pivot Table)
Mean	The ‘average’ ID value in the dataset. The sum of the durations for all values for a particular ID, divided by the frequency of that ID.
Mode	The value that appears most in the ID range
Standard Deviation	The measure of the dispersion of the ID range from its mean value
Minimum	The data value that is less than or equal to all other values in the ID range
5th Percentile	The value below which the lowest 5% of the data falls
1st Quartile	The median of the lower half of the data set
Median	The value that separates the higher and lower half of the ID range
3rd Quartile	The median of the upper half of the data set
95th Percentile	The value above which the upper 5% of the data falls
Maximum	The data value that is greater than or equal to all other values in the ID range

The mean (μ) is calculated using the equation outlined in (20).

$$\mu = \frac{1}{m}\sum_{i=1}^{m} x_i$$

(4.1)

From this, the standard deviation (σ) is calculated using the equation outlined in (21):

$$\sigma = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(x_i - \mu)^2}$$

(4.2)

The remaining frequency, mode, median, minimum, maximum, 5th percentile, 95th percentile, 1st quartile and 3rd quartile is calculated using sort functions. For example, the mode employs the computation outlined in the pseudo code (22).

$$X = sort(x);$$

$$indices = find(diff([X;\ realmax]) \ > \ 0);$$

$$[modeL, i] = max\ (diff([0;\ indices]));$$

$$mode \ = X(indices(i));$$

(4.3)

4.3.4. Profiling IDs

The approach for profiling IDs for benchmarking typical and atypical behaviours within the EPR for PARISS is detailed in Figure 21. This process is undertaken upon initialisation of PARISS in a new hospital infrastructure. Initially there is a need to determine a cross section of typical behaviour for each staff role. This may be provided by the hospital. Otherwise a cross section of staff roles is selected to profile based on the data. If staff information is unavailable (due to anomalous data) then the most active ID types are selected. These ID types are then filtered and visualised against duration. Through this, a benchmark can be determined for typical behaviours and agreed with the hospital. This provides a starting point for PARISS. Once PARISS is in use, the HILML feedback tailors the benchmark for each hospital. A case study of this process is presented in Chapter 5.

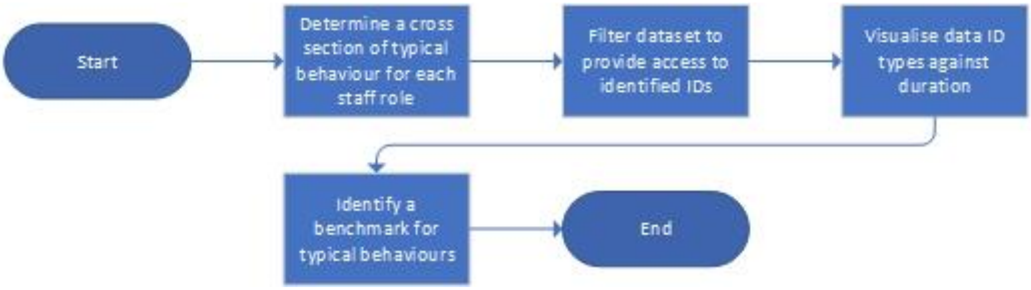

Start
Determine a cross section of typical behaviour for each staff role
Filter dataset to provide access to identified IDs
Visualise data ID types against duration
Identify a benchmark for typical behaviours
End

Figure 21 - PARISS Profiling IDs

4.3.5. Feature Scaling

Three traditional approaches considered for this process are, 1) *Min-Max scaling*, 2) *Decimal scaling* and 3) *Z-score normalisation*.

The Min-Max approach scales the data to a fixed range, between 0-1. The normalised value is obtained using the method outlined in (23).

$$MM(x_{ij}) = \frac{x_{ij} - x_{min}}{x_{max} - x_{min}}$$

(4.4)

Having a bounded range results in lower standard deviations and suppresses the effect of outliers. Decimal scaling normalises by moving the decimal point of values of feature x. Therefore, a $DS(x)$ value is obtained using the method outlined in (24).

$$DS(x_{ij}) = \frac{x_{ij}}{10^c}$$

(4.5)

Where $\max[\backslash(\mathrm{DS}(\mathrm{x_{ij}}))\backslash] < 1$ and c is the smallest integer. The Z-score normalisation approach rescales features so that they have the properties of a standard normalisation. The Z-score approach scales the data to a standard normal distribution. The scaled value is obtained using the method outlined in (25).

$$x_{ij} = Z(x_{ij}) = \frac{x_{ij} - \overline{x}_j}{\sigma}$$

(4.6)

Where $\overline{x}_j$ and σ_j are the sample mean and standard deviation of the *jth* attribute, respectively [143].

The bespoke Min-Max feature scaling process for PARISS is outlined in Figure 22. The data is loaded and a Min-Max boundary value of 0,1 is assigned. A two-phase process is applied. Firstly, PARISS must determine the Min-Max values of each feature. It does this by processing each line individually and assessing its feature class. It then assesses if this is a new maximum or minimum value for this class and if so, it updates the value stored for that feature. This process continues until every line is processed. The second phase normalises the data. PARISS processes each line and determines its feature class. It then normalises this data against the Min-Max value it has determined for that feature class based on the first phase. This process continues until every line is processed.

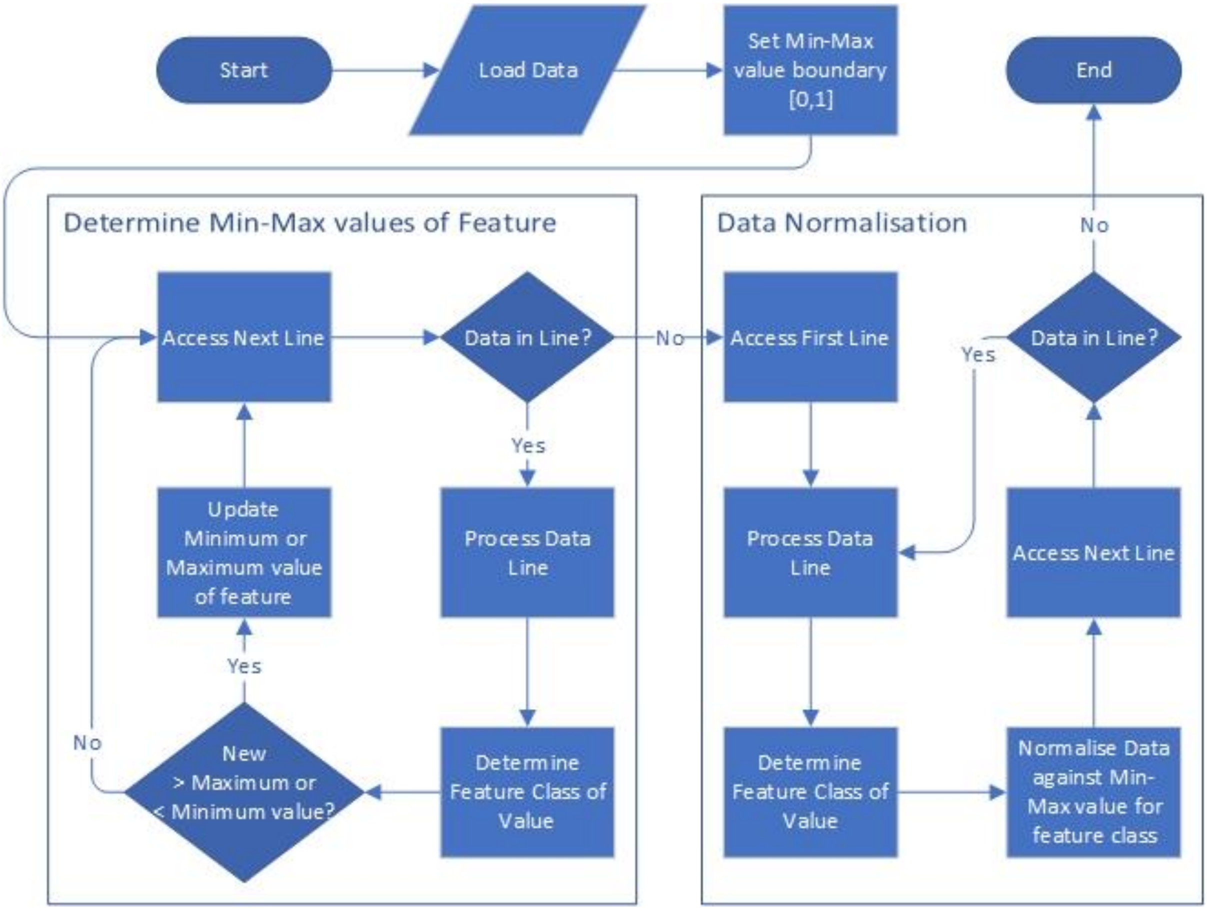

Start
Load Data
Set Min-Max value boundary [0,1]
End
Determine Min-Max values of Feature
Access Next Line
Data in Line?
No
Yes
Update Minimum or Maximum value of feature
Process Data Line
Yes
No
New > Maximum or < Minimum value?
Determine Feature Class of Value
Data Normalisation
No
Access First Line
Yes
Data in Line?
Process Data Line
Access Next Line
Determine Feature Class of Value
Normalise Data against Min-Max value for feature class

Figure 22 - PARISS Min-Max Feature Scaling

4.3.6. Local Outlier Factor

A literature review of outlier detection methods was undertaken in chapter 3 and LOF was chosen as the algorithm for PARISS. This was primarily due to LOF providing an anomaly score, which allowed nuanced scoring for anomalous behaviour. This enables to behaviours to be ranked for prioritisation for the HILML process. For every value of each of the four IDs, a LOF anomaly score is calculated. The LOF anomaly score measures the local deviation of density through determining how isolated the value given by *k*-nearest neighbours (*k* is initially set to 5 as this is the recommended default [26]). A value of 1 indicates that an object is comparable to its neighbours and represents an inlier. A value below 1 indicates a dense region, and would therefore also be an inlier. A value significantly above 1 therefore indicates an outlier (anomaly). Any value below a 1 is an inlier, so all values within the range 0-1 are classified as inliers. Due to a range of 0-1 being classified as inliers, values within the range 1-2 were also determined to be classified as inliers. Therefore initially, any value above 2 is considered to indicate an outlier.

A LOF anomaly score is calculated by taking the number of variants according to the mathematical combination and is calculated using the equation in (26). As there are ten features, 45 LOF scores are calculated to account for all the feature combinations for every ID in the dataset.

$$\binom{n!}{k!} = \frac{n(n-1)\dots(n-k+1)}{k(k-1)\dots 1}$$

(4.7)

A flowchart of the Local Outlier Factor Detection process is presented in Figure 23. This algorithm is presented in further detail in 3.2.3.1.

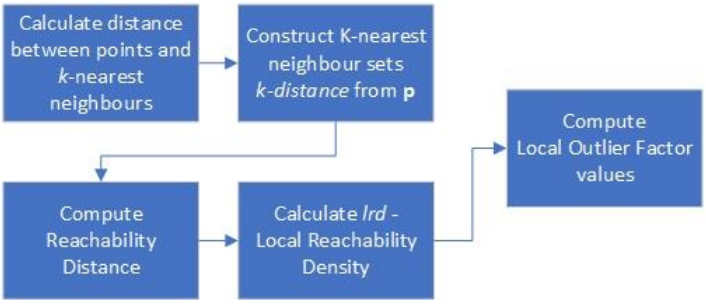

Calculate distance between points and *k*-nearest neighbours
Construct K-nearest neighbour sets *k-distance* from **p**
Compute Local Outlier Factor values
Compute Reachability Distance
Calculate *lrd* - Local Reachability Density

Figure 23 - Local Outlier Factor Detection process

4.3.6.1. Ensemble Averaging

Highlighting an ID of interest, such as a user, is useful in some cases, where repeated inappropriate behaviour is evident. However, if the inappropriate behaviour occurred only once then an analyst would need to investigate the user's entire behaviour for patterns, which is not feasible. A model of combining LOF values into ensemble averaged LOF values is used in order to identify individual audit logs for review. Through highlighting a specific audit log for review, the analyst can review the context around the EPR access and determine whether the access was appropriate or inappropriate. The Ensemble Averaging process is outlined in Figure 24.

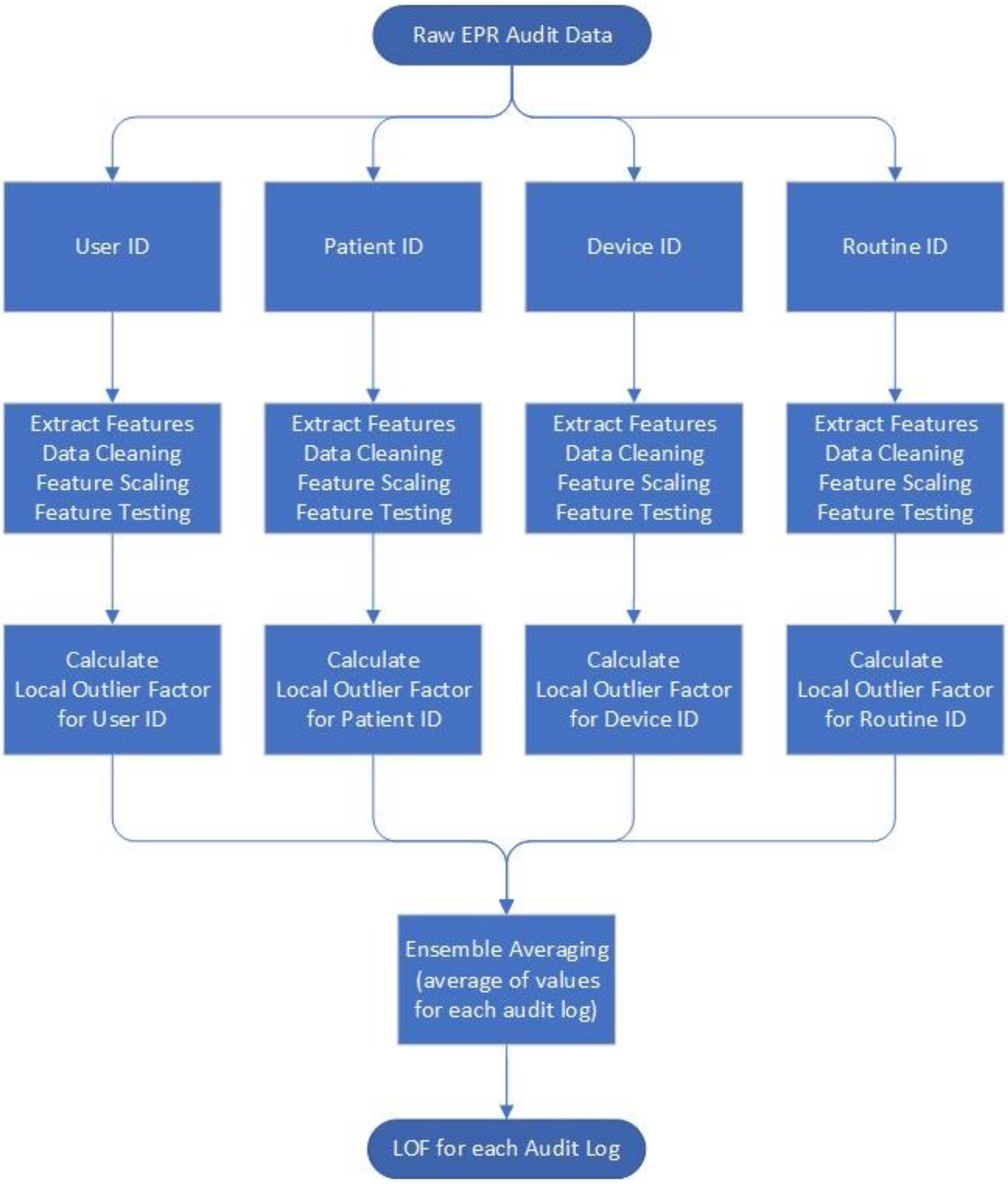


Figure 24 - Ensemble Averaging for PARISS

An ID Anomaly Score does not indicate exactly when a potential inappropriate access has occurred. In order to assign an anomaly score to a specific audit log, rather than a specific ID, the LOF anomaly scores need to calculate the ensemble average. In order to achieve this, a weighted average is applied to each audit log. An additional column is added next to each of the IDs with that ID's associated anomaly score. For every audit log, a weighted average of the four anomaly ID scores is calculated. The calculated ensemble average anomaly score can then be plotted against the Date & Time stamp and visualised to the analyst.

4.3.7. Quantifying LOF

Quantifying LOF is then performed order to convert the Not a Number (NaN) and infinite (Inf) values. A NaN value indicates that a point has many neighbours in the same location, therefore the ratio of densities is undefined, and the points are not outliers. An Inf value occurs when a point is next to several identical points, but is not itself a member of that cluster; they are therefore 'infinite' and can be classified as anomalous. The NaN values are therefore assigned a value of 1, to indicate they are not anomalous, and the inf values are assigned a value of 2, to indicate they are anomalous. Any missing or null values, such as a missing mode value due to lack of data, are assigned the median value for their feature class. The quantifying LOF process for PARISS is detailed in Figure 25.

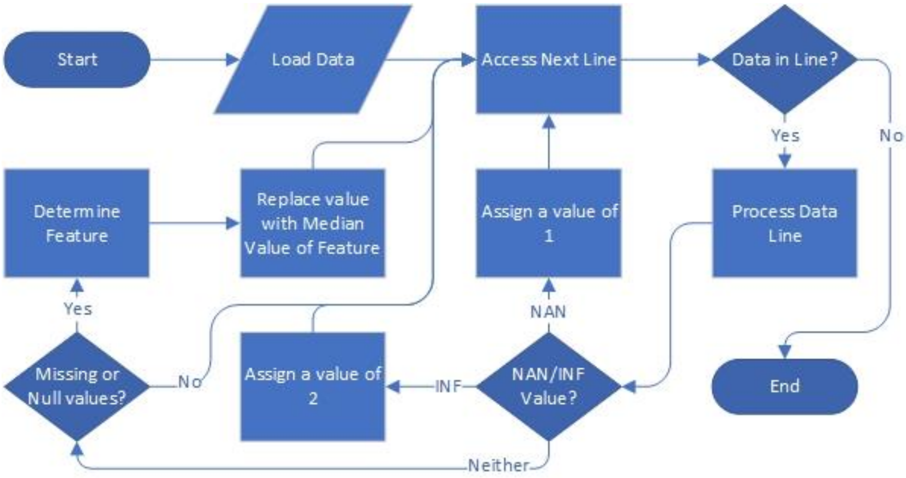


Figure 25 - PARISS Quantifying LOF Process

4.3.8. User Interface Design

The User Interface concept for PARISS is detailed in Figure 26. The User Interface was designed in consultation with a partner hospital. The key requirements were for the primary audit log graph to be the focus of the user interface, accompanied by the table of data to enable further investigation. There was also a requirement for the graphs of the individual ID types to be available to be investigated. The user interface needed to be uncluttered and easy to use, so results could be interpreted and actioned quickly.

Highlighting a single audit log as an outlier enables an analyst to review it within the context of the other audit logs and determine intent. Additionally, an event being an outlier does not constitute maliciousness. Focusing attention on a single event of interest allows analyst intuition to be leveraged in determining context and intent. Therefore, employing an HILML method overcomes the limitations of an unsupervised learning model and incorporates analyst feedback to adapt and use new models. In doing so, the analyst's attention can be focused to the most pertinent events within the dataset.

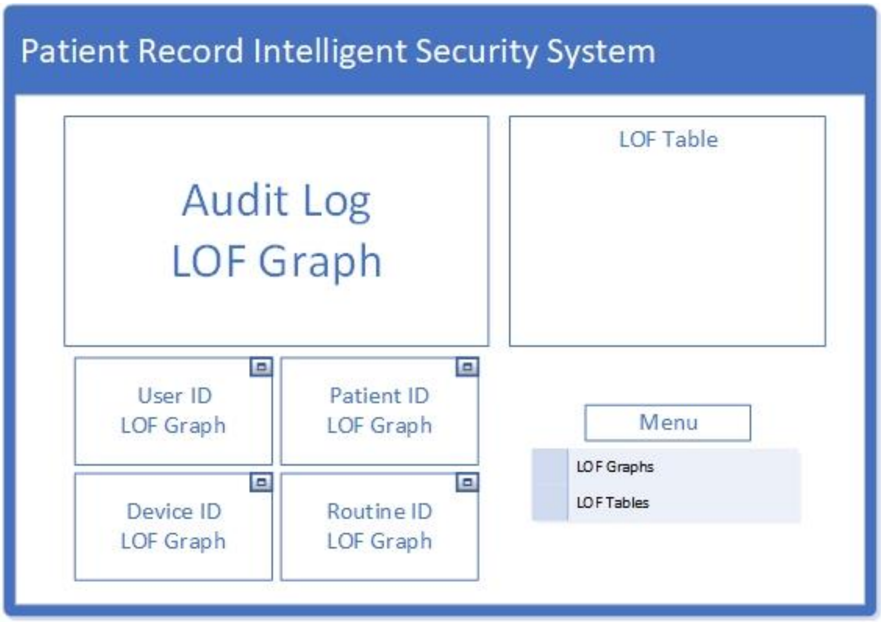


Figure 26 - PARISS User Interface – Dashboard

The menu section allows the user to change what PARISS presents to the user. The LOF Graphs option is the default layout, displayed in Figure 26. In this layout, there are five graphs presented to the user, and one table. The LOF Tables option is similar to LOF Graphs but swaps the graphs for tables. In this layout, there are five tables presented to the user, and one graph (whichever table the user has selected as the primary table).

The visualisation clearly displays the key logs of interest to be investigated by an Analyst. Hovering over a data point displays the date and time of the EPR access (in yy/mm/dd hh:mm format), in addition to its anomaly score. The ensemble averaged score initially is displayed in the primary LOF graph position, with the ID LOF scores displayed in secondary graphs beneath. If a User selects the maximise button on any of the secondary graphs, the graph swaps position with the primary graph. The LOF table lists the LOF results for the primary graph. This initially lists in order of highest LOF anomaly score though this can be changed, for example listing them in date and time order.

Through visualising the anomalies in this way, outliers can be highlighted to an analyst for scrutiny. In our visualisation, outliers in the top quarter of each ID range are highlighted as red, to be investigated as a priority. Outliers in the 3rd quarter appear orange, and outliers in the 2nd quarter appear light orange. This creates an interactive live task list for the analyst, with an anomaly priority ordering. Clicking on a point displays the ID number, which allows the analyst to investigate the activity associated with the ID. The display updates when new data is input and new LOF scores are calculated, providing a current view of anomalous EPR activity within a hospital. Activity such as insider threats (a staff member misusing their access privileges), or external threats (such as credentials accessed through social engineering and utilised for data exfiltration), can be investigated. In this way, the system provides situational awareness to aid patient privacy officers to monitor for malicious or unusual activity proactively.

4.3.9. HILML Algorithm Design

The Human-in-the-Loop Machine Learning process for PARISS is detailed in **Error! Reference source not found.**.

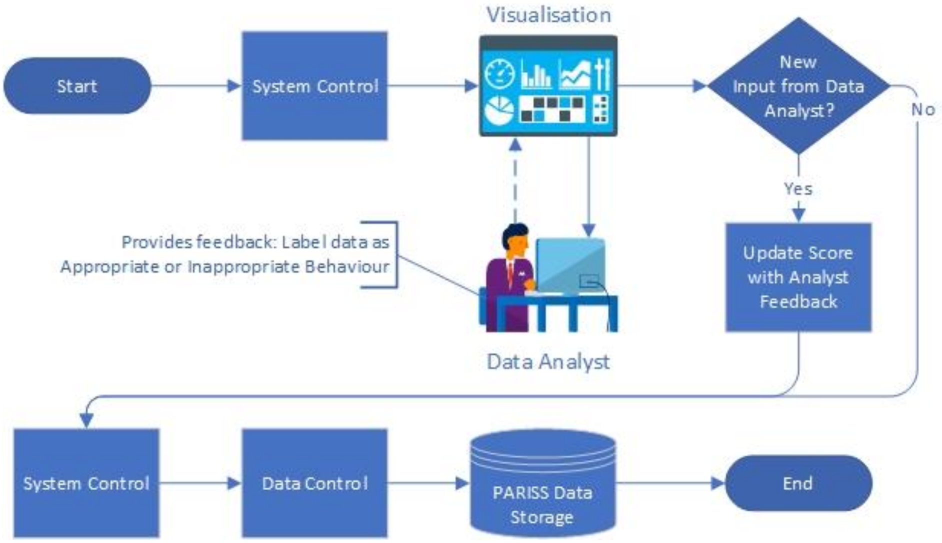

Visualisation
Start
System Control
New Input from Data Analyst?
No
Yes
Update Score with Analyst Feedback
Provides feedback: Label data as Appropriate or Inappropriate Behaviour
Data Analyst
System Control
Data Control
PARISS Data Storage
End

Figure 27 - Human-in-the-Loop Machine Learning process

By including the HILML model, an active learning approach employs analyst feedback to train the machine learning model. An analyst is asked to review the audit log anomalies and assign a feedback score. By default, the feedback score for every anomaly is 1. The analyst can provide a feedback score in the range of 0.1 to 2. This LOF score is multiplied by the feedback score to provide the final score. Therefore, if a feedback score of 2 is given, indicating anomalous behaviour, it multiplies the anomaly score by 2 for the relevant IDs and therefore makes them more likely to be rated highly in future. If an analyst gives a score as low as 0.1, this multiplies the anomaly scores by 0.1, which effectively whitelists the IDs, making them unlikely to appear as an anomaly in future.

The feedback scores are being updated throughout the use of PARISS, incorporating analyst feedback into the anomaly identification process. The range of 0.1 to 2 is chosen to reflect that anomalies within the 0-2 range are classified as appropriate behaviour within PARISS. This process is required as some users, such as consultants, will be required to access a variety of patients and would otherwise regularly be flagged as acting anomalously.

4.3.10. Database Design

Figure 28 outlines the database solution for PARISS.

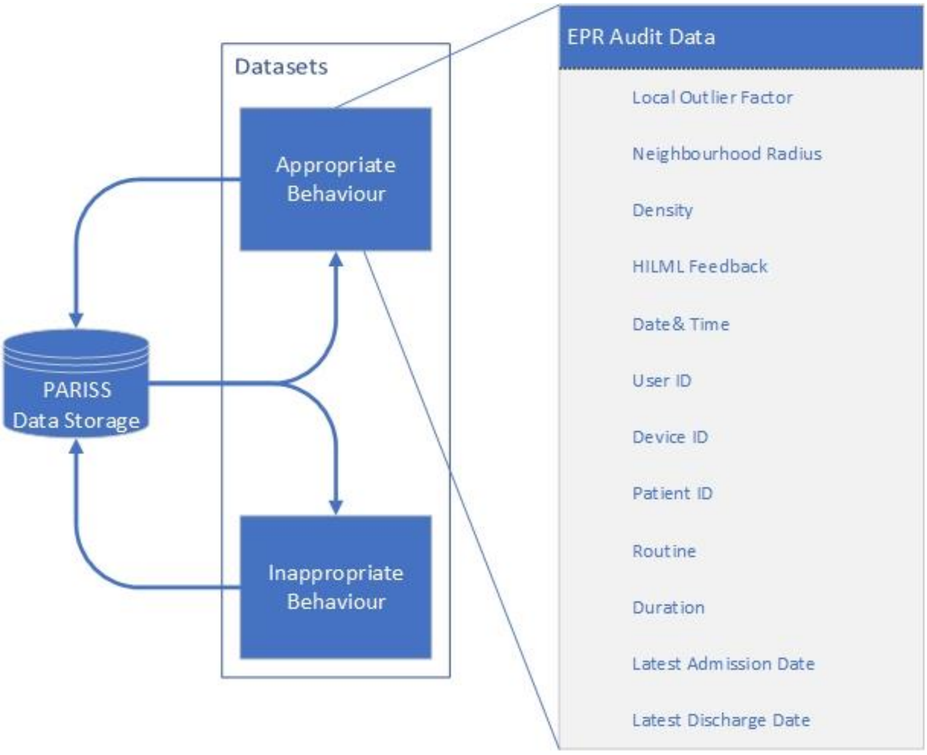


Figure 28. Database Diagram

In addition to the extracted data, the stored data includes the assigned LOF score (including corresponding neighbourhood radius and density values) and HILML feedback for the audit log.

4.3.11. Overall System Design

A high-level view of the PARISS Architecture is presented in **Error! Reference source not found.**.

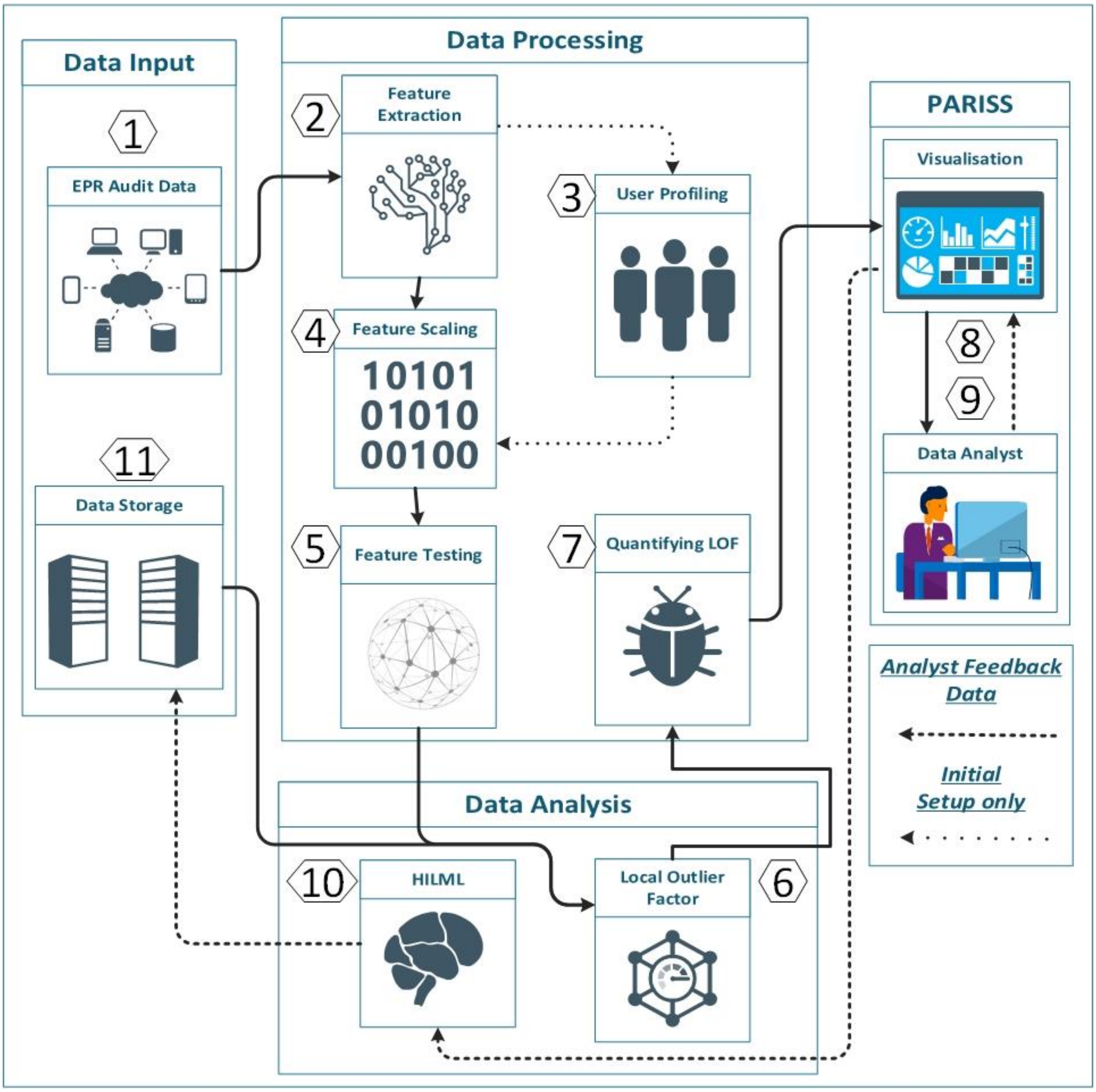


Figure 29. Overall System Design

4.4. Summary

In this chapter, the Patient Record Intelligent Security System (PARISS) is presented. The system is deployed within the hospital infrastructure and extracts data from the data warehouse. PARISS then pre-processes the data by extracting features, scaling the data and testing for appropriate features, before calculating the Local Outlier Factor for each ID type. Ensemble averaging is then applied to provide a LOF anomaly score to individual audit logs. This is then visualised and presented to the user. The user can provide feedback to PARISS through a Human-in-the-Loop machine learning approach. This data is then stored and called upon when more audit data is extracted from the data warehouse. In chapter 5, a case study will be used to test the proposed system using a real-world EPR dataset containing 1,007,727 rows of audit logs is presented and explored.

5. Case Study: EPR Data Validation

5.1. Introduction

Data behaviour within healthcare infrastructures needs proactive monitoring for malicious, erratic or unusual activity. Patients need to be assured of three crucial security principles *1)* the data stored is trustworthy and accurate. *2)* Data can be reliably accessed by healthcare professionals when needed. *3)* Only authorised healthcare professionals have access to the data, and only access it when it is appropriate to do so. Issues also surround data being exchanged across multiple countries that have different laws and regulations concerning data traversal, protection requirements, and privacy laws [209].

A case study of actual EPR audit data is presented as an exploration of user behaviours within an EPR, in order to understand anomalous activity. The data used in this research was supplied in three stages by the hospital partner (who currently use the FairWarning commercial solution). 1) They provided one month (May 2017) of data to allow for the development of a proof of concept. 2) Following on from the successful proof-of-concept stage, a further 6 months of data was provided (July-December 2016) in order to explore the effects of a larger dataset on the PARISS process. Finally, 3) the hospital provided 18 months of data (February 2016 – August 2017), the maximum that could be provided, as data was deleted after that point due to lack of data storage at the hospital.

According to the hospital, the number of anomalous alerts generated are affected by three factors.

1. The number of staff proportionally increases the number of alerts (a hospital with 4 times the staff would have 4 times the alerts).
2. The number of patients also affects proportionally the number of alerts, creating a compound effect (similarly if there were 4 times the number of patients, this would have another 4 times the number of alerts).
3. A complexity factor, which is a considerable challenge to define. This refers to the complexity of specialties within the hospital (such as in an acute hospital) and the innate curiosity of the staff.

The task of navigating this data for anomalous activity is therefore considerable.

5.2. EPR Data

This rich dataset contains 1,007,727 rows of audit logs of every user and their EPR activity in a single UK specialist hospital over a period of 18 months (28-02-16 – 21-08-17). The data used in this research is from a specialist hospital. A large teaching hospital would have approximately four times the number of staff, and would, therefore, have a proportional increase in data quantity.

5.2.1. EPR Data Fields

The EPR in this case study uses a unique hierarchal relationship data structure. For this reason, the data cannot be queried directly and extracted. Instead, it is hard-coded into the EPR to push this data on a daily basis at 04:30AM. This data is pushed to a shared data file in pipe-delimited format (comma-delimited may cause issues with certain fields such as name, or routine). This data remains in .txt format until it is synthesised into a single .csv data file using the command prompt. A diagram of this process is presented in Figure 30.

The process outlined in Figure 30 however is unusual. Many EPRs (and other medical systems) use a relationship data structure, therefore this data can instead be extracted using an SQL query and the .csv file will be created.

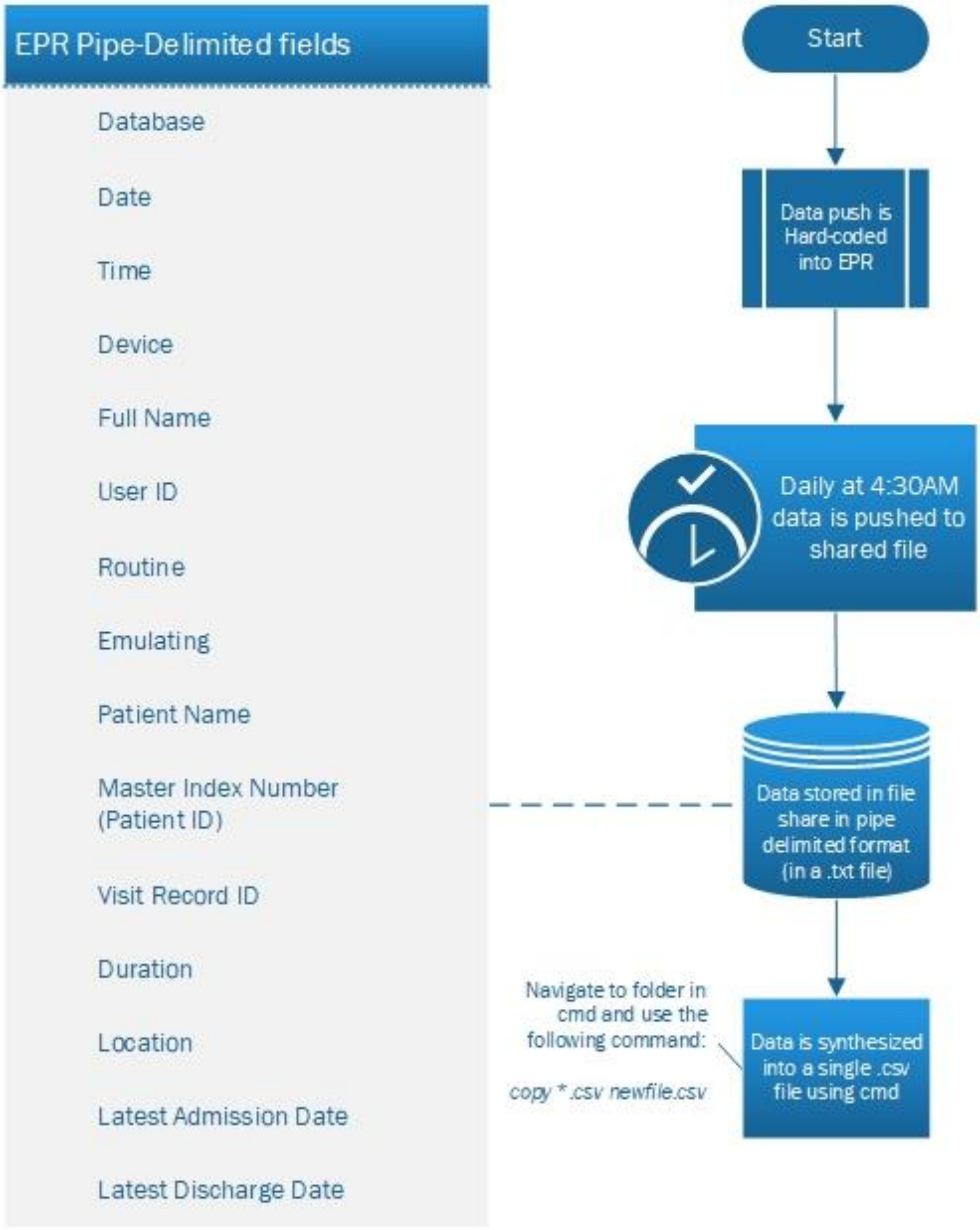

EPR Pipe-Delimited fields
Database
Date
Time
Device
Full Name
User ID
Routine
Emulating
Patient Name
Master Index Number (Patient ID)
Visit Record ID
Duration
Location
Latest Admission Date
Latest Discharge Date
Start
Data push is Hard-coded into EPR
Daily at 4:30AM data is pushed to shared file
Data stored in file share in pipe delimited format (in a .txt file)
Navigate to folder in cmd and use the following command:
copy *.csv newfile.csv
Data is synthesized into a single .csv file using cmd

Figure 30 – Extraction of EPR Data

The data provided in this research was previously extracted for input into FairWarning. FairWarning is a procedure-based detection system for patient privacy. This work aims to expand on the capability of FairWarning through applying machine learning and visualisation techniques. Other fields extracted for the benefit of FairWarning (but unfortunately not made available to this project due to information governance and staff privacy concerns) are outlined in Table 10. This data can be used to give a further picture of potentially anomalous behaviour. For example, using a member of staff’s address in the Electronic Staff Record (ESR) system along with the patient’s address in the Patient

Administration System (PAS) may indicate inappropriate behaviour (such as a member of staff reading details of a neighbour without clinical reason). The key alerts generated by FairWarning are for i) Self-Exam, ii) Family Snooping, iii) Employee Snooping and iv) Neighbour Snooping.

Table 10 - Other FairWarning Input Data

EDMS Fields	ESR Fields	PAS Fields	ICE Fields
Timestamp	Employee Number	Hospital Number	Time Stamp
User ID	First Name	GenderDesc	User ID
User First Name	Last Name	Address Line 1	User First Name
User Last Name	Birth Date	Address Line 2	User Last Name
Patient ID	Age	Address Line 3	Patient ID
Patient First Name	Position Title	Address Line 4	Patient First Name
Patient Last Name	Department	Postcode	Patient Last Name
Application	Department ID	Date of Birth	Application
Event Description	Manager Surname	Confidential Flag	Event Description
Event Type	Manager First Name	Marital Status	Event Type
	User Status	Family Name	
	User Hire Date	Given Name	
	Ctr Hrs		
	Location		
	User Gender		
	Address Line 1		
	Address Line 2		
	Address Line 3		
	Town or City		
	County		
	Post Code		
	Country		
	Telephone Home		
	Marital Status		

A sample of the EPR data used in this research is presented in Table 11. In the first row of Table 11, User 865 accesses the ‘Pharmacy Orders’ function of the EPR on Patient 58991 whilst using Device 362. Each User UID, Patient UID and Device ID is tokenised through isolating the unique entries and assigning each value an incrementing number. There are 1,515 unique User IDs, 72,878 unique Patient IDs, 2,270 unique Devices IDs and 13,722

Routine ID combinations within the dataset. Therefore, there are 90,385 unique IDs in the dataset in total (for user, patient, device and routine combined).

Table 11 - EPR Audit Sample Data

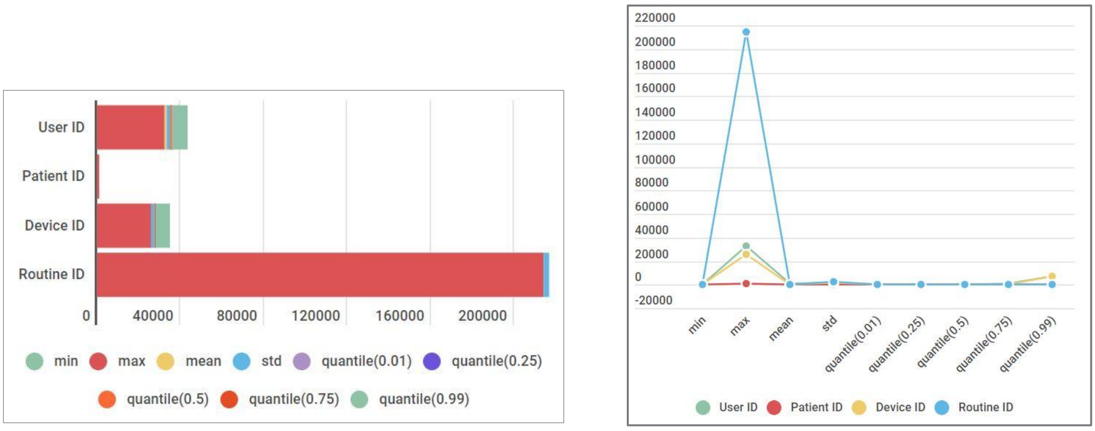
Date&Time	Device	User UID	Routine	Routine Description	Patient UID	Duration	Location Latest Adm Date	Latest Dis Date
28/02/16 00:00	362	865	PHA.ORDS	Pharmacy Orders	58991	54	28/02/2016	29/02/2016
28/02/16 00:02	923	199	REC REC:(DRP) UK.OE	Recent Clinical Results Recent Clinical Results:(Departmental Reports) UK.View Orders	17278	77	15/02/2016	15/02/2016
28/02/16 00:02	103	677	ASF	Assessment Forms	4786	13	22/07/2008	22/07/2008
28/02/16 00:02	103	677	ASF	Assessment Forms	4786	54	22/07/2008	22/07/2008
28/02/16 00:04	923	199	REC UK.OE	Recent Clinical Results UK.View Orders	62121	147	08/02/2016	08/02/2016
28/02/16 00:04	103	677	ASF VH	Assessment Forms Visit History	14067	39	28/09/2004	28/09/2004
28/02/16 00:04	845	1489	PHA.ORDS	Pharmacy Orders	49304	22	23/01/2002	23/01/2002
28/02/16 00:06	923	199	REC REC:(DRP) UK.OE	Recent Clinical Results Recent Clinical Results:(Departmental Reports) UK.View Orders	60948	165	08/01/2016	08/01/2016
28/02/16 00:08	775	568	NOTE	Patient Care Notes	32826	75	25/01/2012	25/01/2012
28/02/16 00:10	748	797	REC REC:(DRP)	Recent Clinical Results Recent Clinical Results:(Departmental Reports)	2166	20	28/01/2016	28/01/2016

A distribution of durations for each of the ID types in the dataset is presented in Table 12.

Table 12 - Distributions of Durations

Value	User ID	Patient ID	Device ID	Routine ID
min	1	1	1	1
max	32,557	1,011	25,739	214,345
mean	665.166	13.828	443.933	59.476
std	1,842.86	25.556	1,373.30	2,153.38
quantile(0.01)	1	1	1	1
quantile(0.25)	34	1	8	1
quantile(0.5)	220	4	63	1
quantile(0.75)	651	15	295	2
quantile(0.99)	7,411	116	7,138	253

A visualisation of this data as a stacked bar chart and line graph is presented in Figure 31. Visualising the distribution highlights clear anomalies in the data. The max value is significantly higher than the 0.99 quantile for each ID. This can have significant effects on the mean. For example, with Routine ID, the median is 1, but the mean is 59.476. The visualisation demonstrates how anomalous the max is compared to the rest of the dataset.

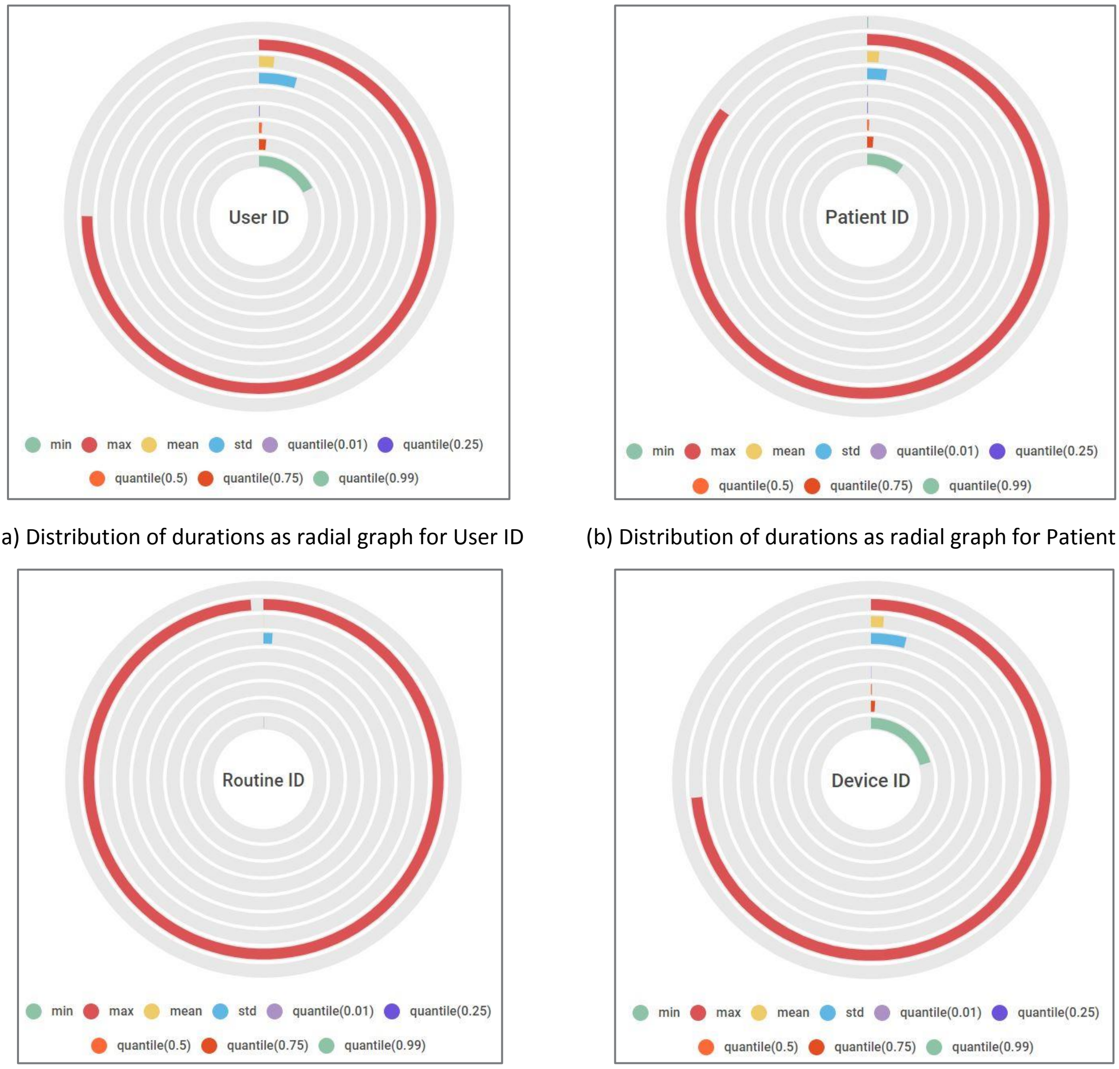


(a) Distribution of durations as a stacked bar chart

(b) Distribution of durations as a line graph

Figure 31 – Distribution of durations

A visualisation of each ID type is presented as radial graphs in Figure 32.



(a) Distribution of durations as radial graph for User ID

(b) Distribution of durations as radial graph for Patient ID

(c) Distribution of durations as radial graphs for Device ID

(d) Distribution of durations as radial graphs for Routine ID

Figure 32 – Distribution of durations as radial graphs

The visualisation demonstrates for each ID type the extent to which the max can skew the feature sets and demonstrates clear anomalies. For example, when compared to the green line showing the 0.99 quantile, the extent to which the max value is an anomaly is clear.

5.2.2. EPR Data Tokenisation

The dataset contains four distinct ID types, User, Patient, Device and Routine. Each User ID, Patient ID and Device ID is tokenised by isolating the unique entries and assigning each value an incrementing number, using this bespoke algorithm:

Algorithm 5.1 – Tokenisation algorithm for PARISS

```
Sort values into alphabetical order
    Set 1st a tokenised value of 1
        IF next value is same as previous value
            Assign same tokenised value
        ELSE
            Assign a value +1 to the previous tokenisation value
        END IF
        All values have been tokenised
END
```

This is done to anonymise the dataset as mandated by the hospital Information Governance. The Routine ID was not tokenised as it denotes the tasks performed by the User on the EPR for the interaction. This tokenisation process is displayed in Figure 33.

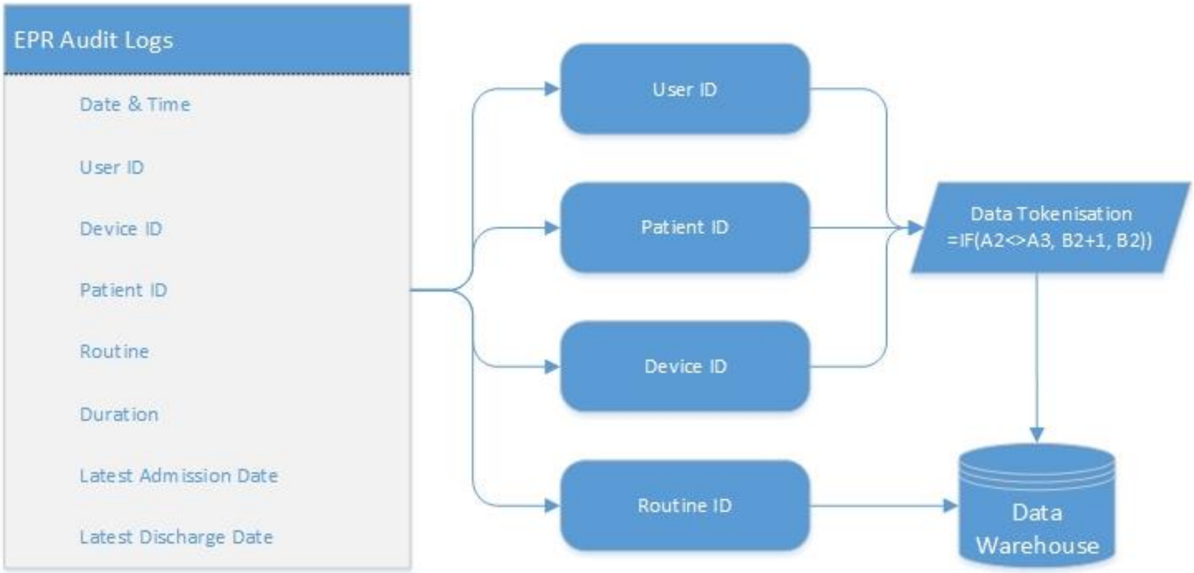
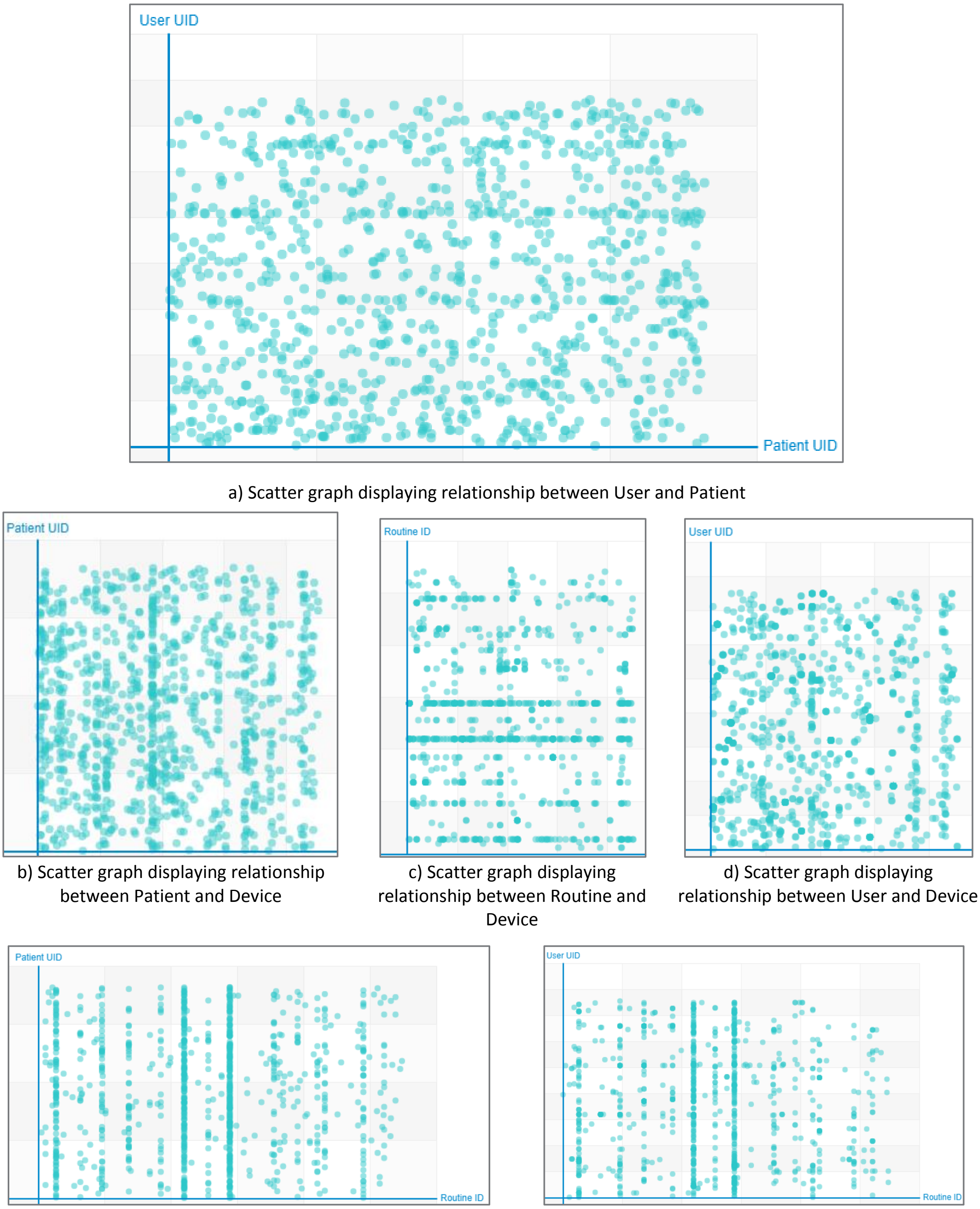


Figure 33 – ID extraction and tokenisation

5.2.3. EPR Data Discussion

In Figure 34, scatter graphs display the relationship between different IDs in the dataset.



a) Scatter graph displaying relationship between User and Patient

b) Scatter graph displaying relationship between Patient and Device

c) Scatter graph displaying relationship between Routine and Device

d) Scatter graph displaying relationship between User and Device

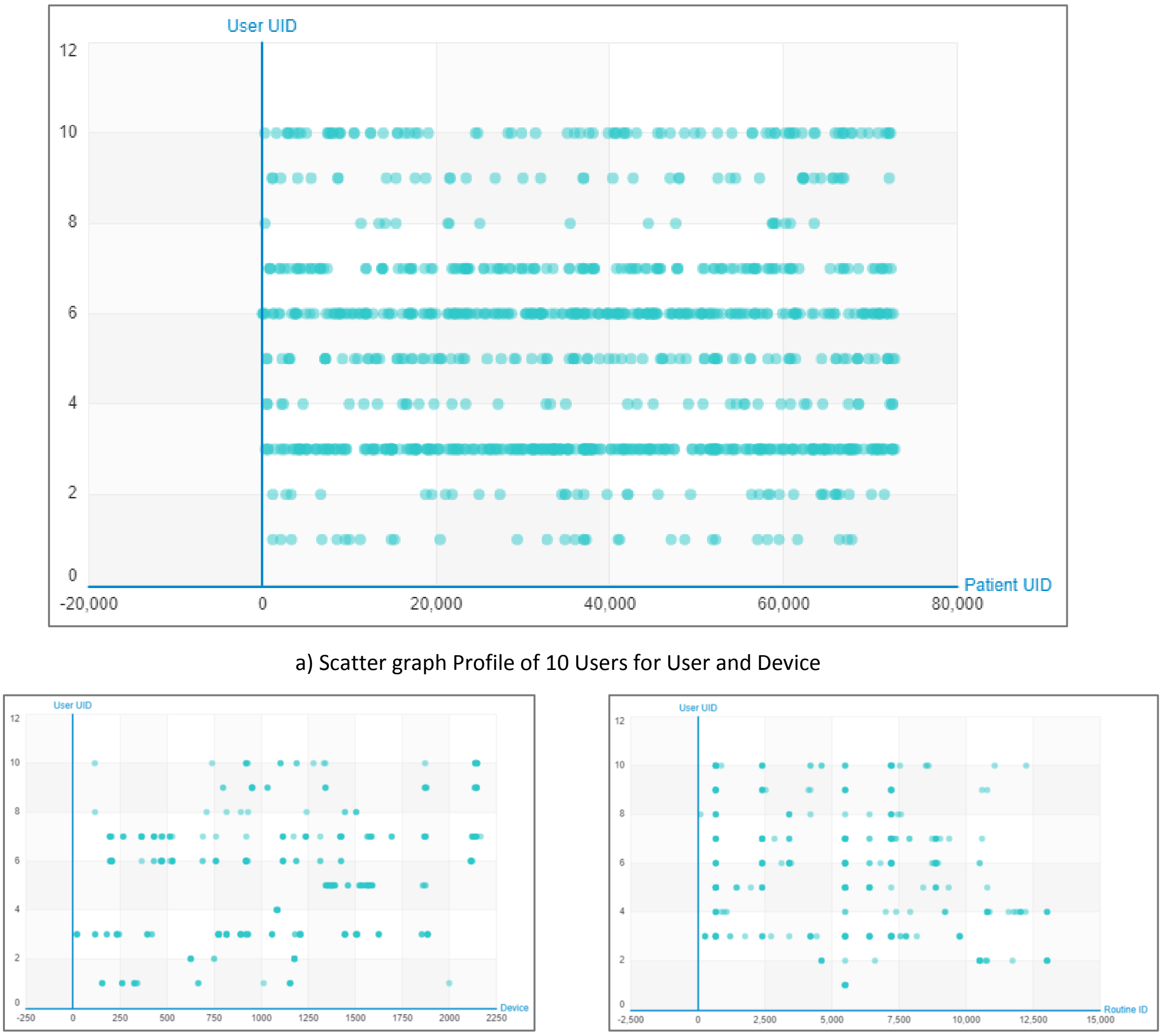
e) Scatter graph displaying relationship between Patient and Routine

f) Scatter graph displaying relationship between User and Routine

Figure 34 - Scatter graph displaying relationship between ID types

Specifically, this includes the relationships between (a) User and Patient, (b) Patient and Device, (c) Routine and Device, (d) User and Device, (e) Patient and Routine, (f) User and Routine. The graphs are high level representations of which IDs interact with which other IDs over 18 months in order to identify trends. The graphs demonstrate the complexity of the data, as there is no clear structure at face value. The graphs do not show clustering or any clear patterns of behaviour and do not create useful insights within datasets. The data therefore needs to have meaningful features selected and users of interest identified. In doing so, legitimate accesses can be removed from the visualisation and illegitimate accesses highlighted to a privacy officer. The exceptions are in Figure 34(c), (e) and (f), which have distinct lines through the centres of the graphs. These graphs have the Routine ID and due to high frequency (20%) of some routine IDs, within the dataset, such as Pharmacy Orders (which is used 214,345 times), this creates a distinct pattern in the dataset.

In Figure 35, a profile of 10 Users is presented. Scatter graphs display the relationships between User ID and (a) Patient, (b) Device and (c) Routine is displayed. The 10 users are a random selection of users, as visualising 10 users represents a reflection of the dataset as a whole. Figure 35 is a representation of the potential of the system to filter the larger dataset of Figure 34 to users of interest. The selected users are the first 10 in the dataset (having been assigned a random number through the tokenisation process). In Figure 34(a), Users 3, 6 and 7 access a larger number of patients than the rest of the users. This likely indicates they are a similar staff role as they are users that access a larger number of patient records than other staff, such as admin staff who check patients in. When comparing these same users in Figure 34(b), the same users access a large number of devices, indicating that these members of staff work in many areas of the hospital, which is in contrast to User 4, who only uses 1 device. There's is less of a correlation when comparing the Users Routines in Figure 34(c), indicating most users access a variety of routine functions in the system. In this way, roles can be clustered within the data and features extracted. Unusual or erratic spikes in activity would indicate illegitimate activity that would warrant further investigation.



a) Scatter graph Profile of 10 Users for User and Device

b) Scatter graph Profile of 10 Users for User and Patient

c) Scatter graph Profile of 10 Users for User and Routine

Figure 35 – Scatter graph Profile of 10 Users for ID Types

5.3. Profiling IDs

In order to profile behaviour within the hospital, the most active IDs for each ID type is described to establish typical behaviour patterns. In PARISS, this process can either be provided by the hospital, or can be calculated as detailed in this section. This dataset was provided on the condition of anonymity of data; therefore, the latter process is taken in order to identify an anomaly within the dataset. Once initial benchmarking values are set, the HILML process improves accuracy and reliability and tailors to the hospital as detailed in section 4.3.9. The initial benchmark values for each of the ID types are defined by reviewing the density of scatter graphs and agreeing the values with the hospital.

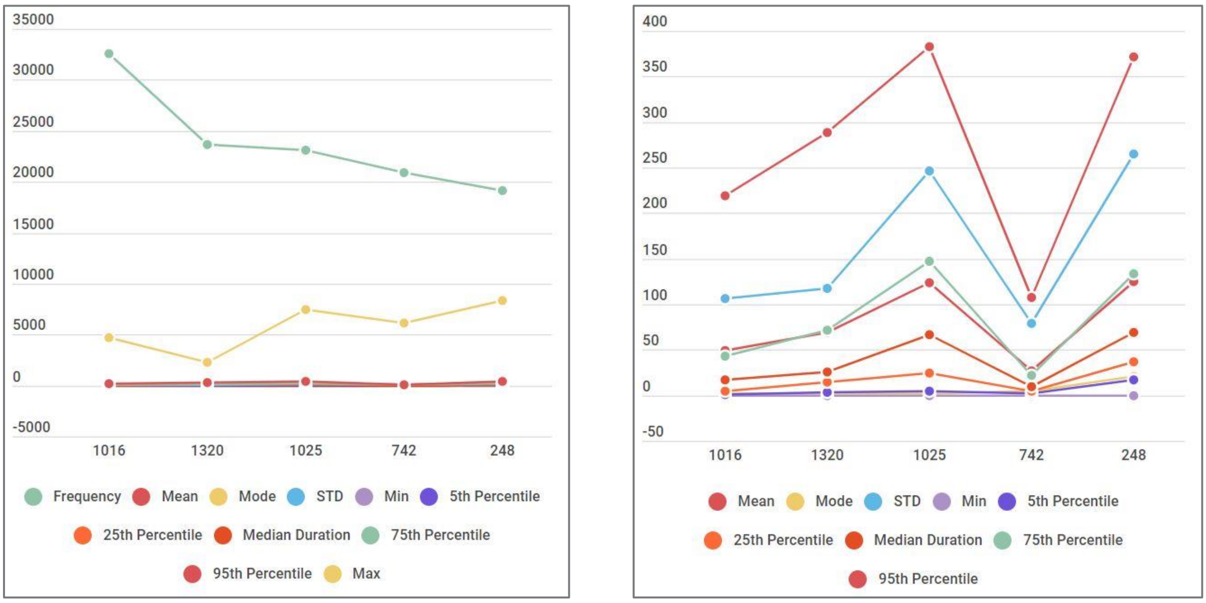
5.3.1. Profiling User ID

The most active profiles (as defined by highest frequency value) for User ID are displayed in Table 13. The table therefore shows the 5 users' IDs with the highest frequency values along with other features associated with the user ID.

Table 13 - Profiles for User ID

User ID	Frequency	Mean	Mode	STD	Min	5th Percentile	25th Percentile	Median Duration	75th Percentile	95th Percentile	Max
1016	32557	49.50	2	106.66	0	1	5	17	43	219	4751
1320	23674	69.57	2	117.32	0	3	14	26	72	289	2268
1025	23104	124.06	2	246.66	0	4	24	67	147	383	7469
742	20907	27.39	5	79.29	0	2	5	9	22	108	6081
248	19160	125.23	21	264.26	0	17	37	69	134	371	8360

A line graph displaying the data profiles is presented in Figure 36(a). The same data with Max and Frequency removed is presented in Figure 36(b) to give a more detailed overview of the lower data ranges.



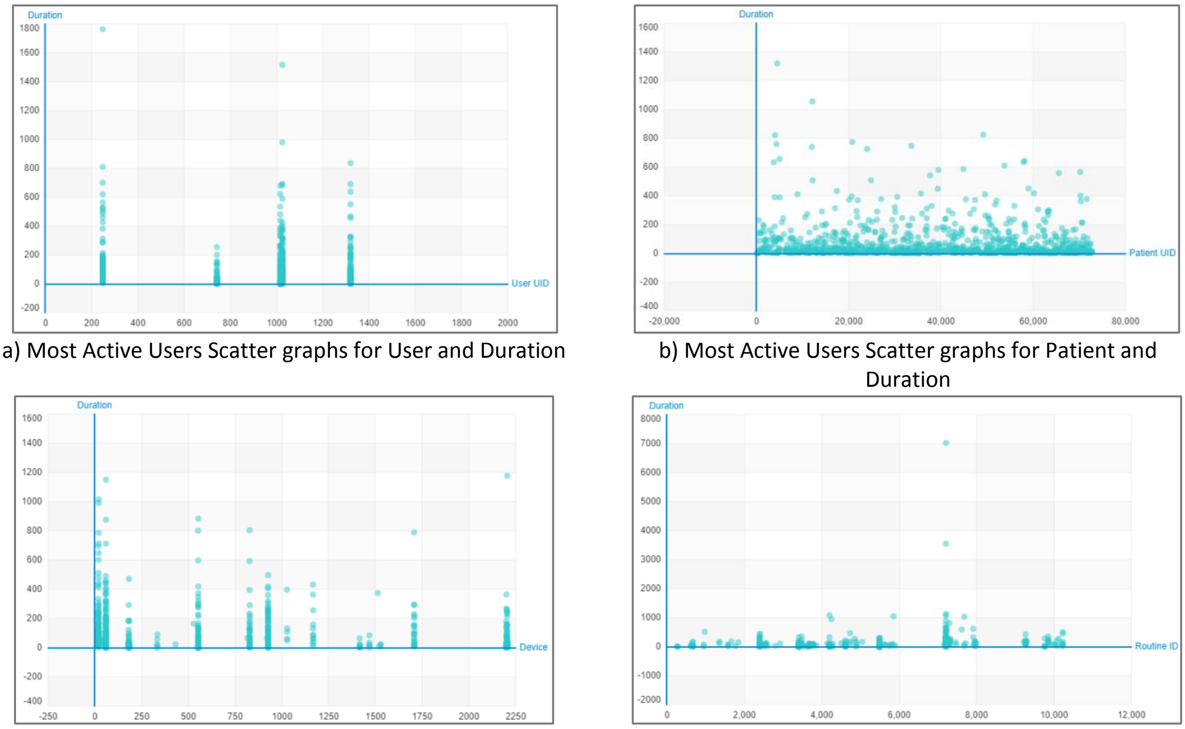
a) User ID Profiles

b) User ID Profiles (exc. Frequency and Max)

Figure 36 – User ID Profiles Line Graphs

Figure 37 displays the 5 most active User IDs and compares each ID with duration for (a) User, (b) Patient, (c) Device and (d) Routine.

In Figure 37(a), the density of the scatter graph is around 400 seconds, with one of the most active users in the dataset never exceeding 400 seconds. Additionally, in Figure 37(b), a user typically spends 300 seconds (5 minutes) or less performing an action on a patient. Figure 37(c) demonstrates that users spend longer on certain devices than others, for example some devices are never accessed for routines that are longer than 200 seconds, whereas others frequently exceed that. Finally, Figure 37(d) demonstrates that most routines have consistent usage times within the dataset, with some exceptions, such as the routine that is accessed for 7,000 seconds (almost 2 hours). Therefore, 400 seconds is the initial benchmark for typical user behaviour in the dataset.



a) Most Active Users Scatter graphs for User and Duration

b) Most Active Users Scatter graphs for Patient and Duration

c) Most Active Users Scatter graphs for Device and Duration

d) Most Active Users Scatter graphs for Routine and Duration

Figure 37 - Most Active Users Scatter graphs

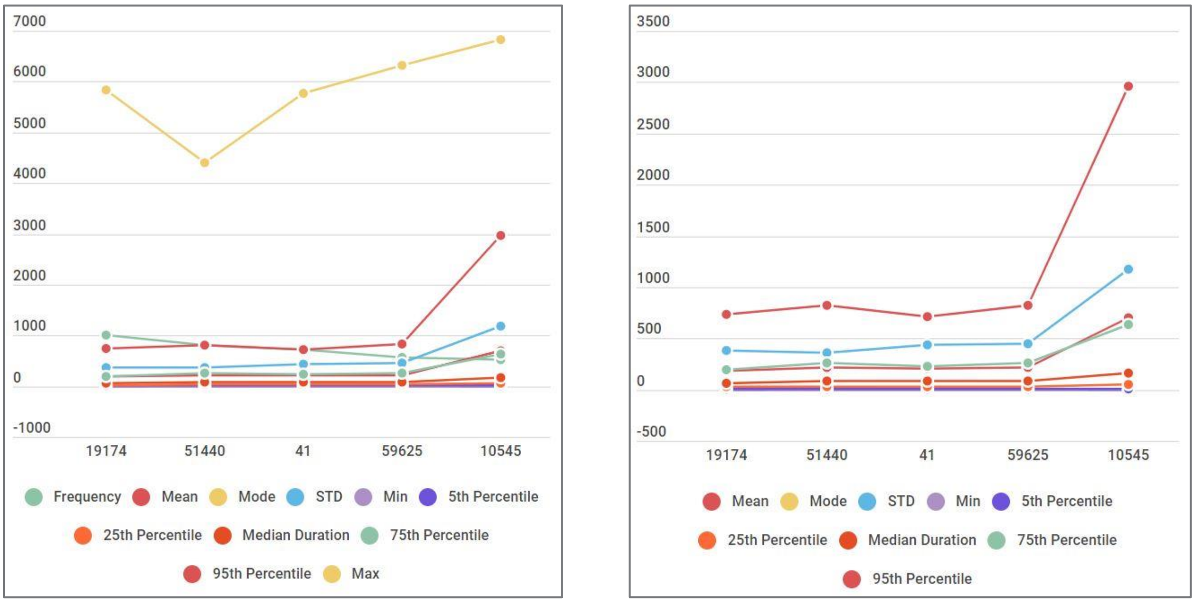
5.3.2. Profiling Patient ID

The most active profiles (as defined by highest frequency value) for Patient ID are displayed in Table 14.

Table 14 - Profiles for Patient ID

Patient ID	Frequency	Mean	Mode	STD	Min	5th Percentile	25th Percentile	Median Duration	75th Percentile	95th Percentile	Max
19174	1011	187.59	2	377.12	1	4	23.5	66	196.5	730.5	5822
51440	800	218.70	2	363.80	1	5	29	88	263.25	817.45	4395
41	715	206.75	20	439.57	1	6	28.5	82	229.5	715.1	5772
59625	574	219.28	7	445.48	0	3	28.25	79	258.5	821.3	6312
10545	527	698.13	2	1174.52	1	3	54	163	633.5	2960.8	6826

Figure 38(a) is a line graph with the profiles for these patient IDs. Max and Frequency are removed in Figure 38(b).

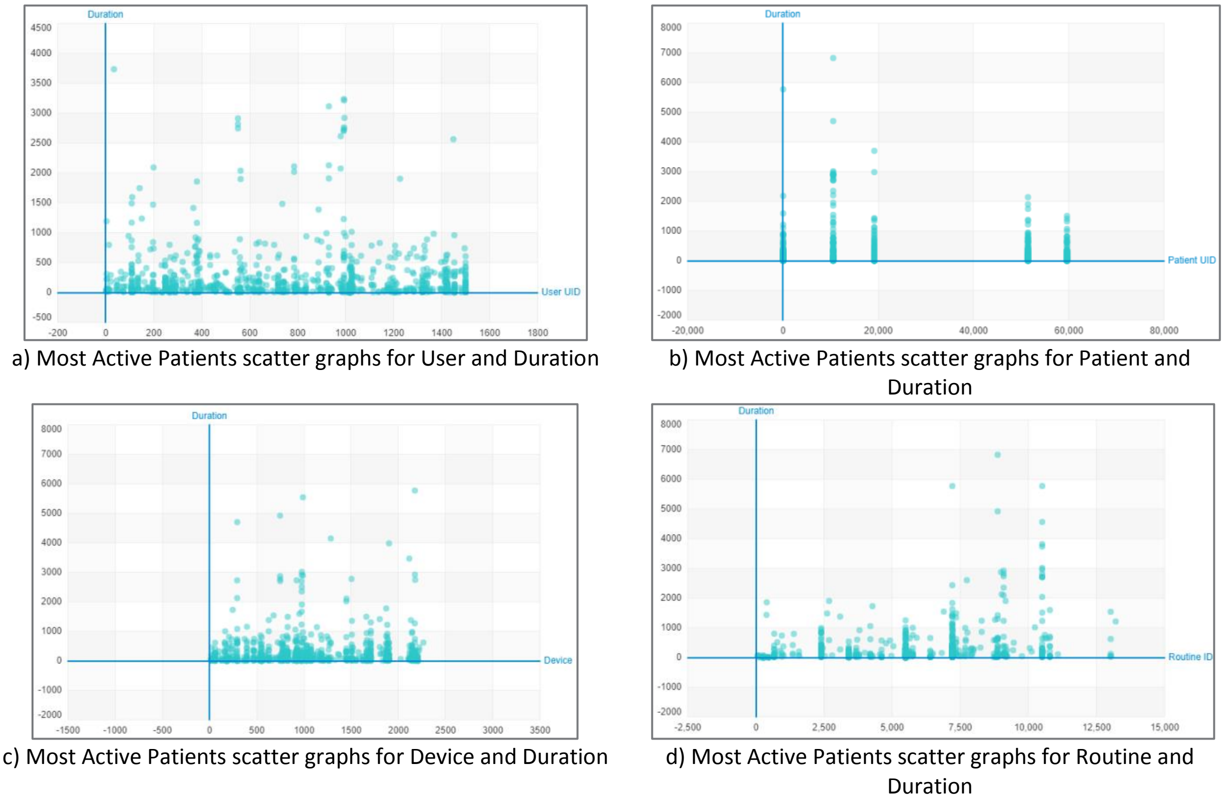


a) Patient ID Profiles

b) Patient ID Profiles (exc. Frequency and Max)

Figure 38 – Patient ID Profiles Line Graphs

Figure 39 displays the 5 most active Patient IDs and compares each ID with duration for (a) User, (b) Patient, (c) Device and (d) Routine.



a) Most Active Patients scatter graphs for User and Duration

b) Most Active Patients scatter graphs for Patient and Duration

c) Most Active Patients scatter graphs for Device and Duration

d) Most Active Patients scatter graphs for Routine and Duration

Figure 39 - Most Active Patients scatter graphs

In Figure 39(a), the density of the scatter graph is around 1,000 seconds (17 minutes), with some notable anomalies, such as the user that accesses a patient ID for 3,750 seconds. This observation is strengthened by Figure 39(b), which indicates that 1,000 seconds is a typical time for a patient record to be accessed. Most clinic sessions last 15 minutes, which would confirm this observation. Figure 39(c) also has 1,000 seconds as a typical access time, with a few exceptions of over 5,000 seconds. However, unlike with User IDs there are no clear observations that patient IDs are accessed for longer on different devices than others. Finally, Figure 39(d) demonstrates that most routines have consistent usage times of under 1,000 seconds within the dataset, with some exceptions, such as the routine that is accessed for 7,000 seconds (almost 2 hours). Therefore, 1,000 seconds is the initial benchmark for typical patient behaviour in the dataset.

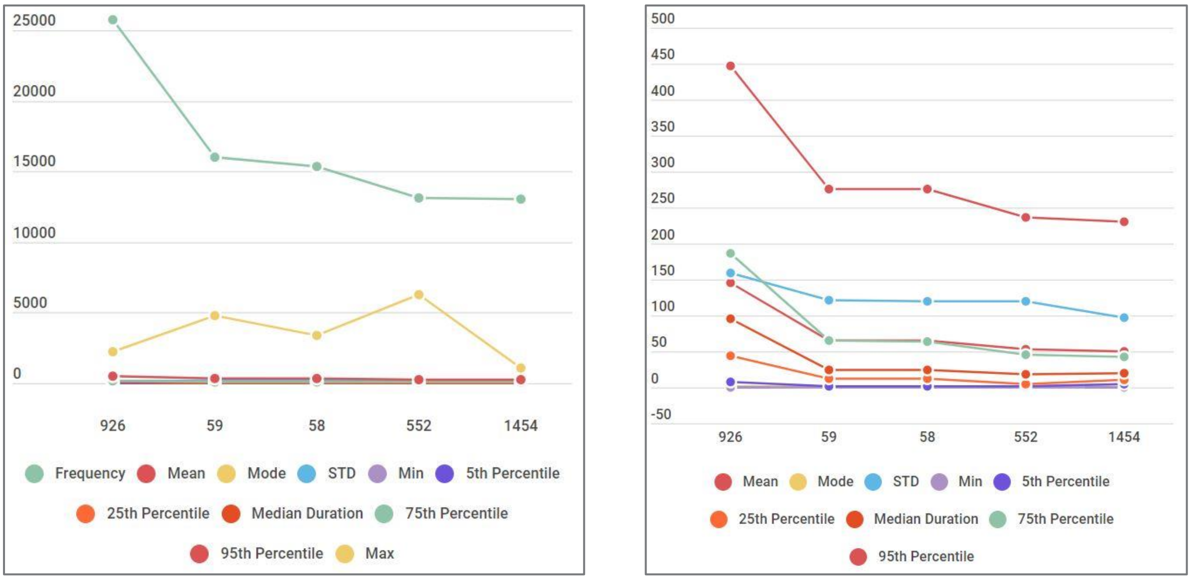
5.3.3. Profiling Device ID

The most active profiles for Device ID are displayed in Table 15.

Table 15 - Profiles for Device ID

Device ID	Frequency	Mean	Mode	STD	Min	5th Percentile	25th Percentile	Median Duration	75th Percentile	95th Percentile	Max
926	25739	144.74	2	159.19	0	8	44	96	187	447	2196
59	16011	65.18	2	121.28	0	2	12	24	65	275.5	4751
58	15386	65.56	2	120.07	0	2	12	24	64	276	3372
552	13140	52.95	2	119.31	0	1	5	18	46	237.05	6268
1454	13079	49.65	5	96.22	0	4	10	20	43	230	1064

Figure 40(a) displays the data as a line graph and Figure 40(b) removes the Max and Frequency values.

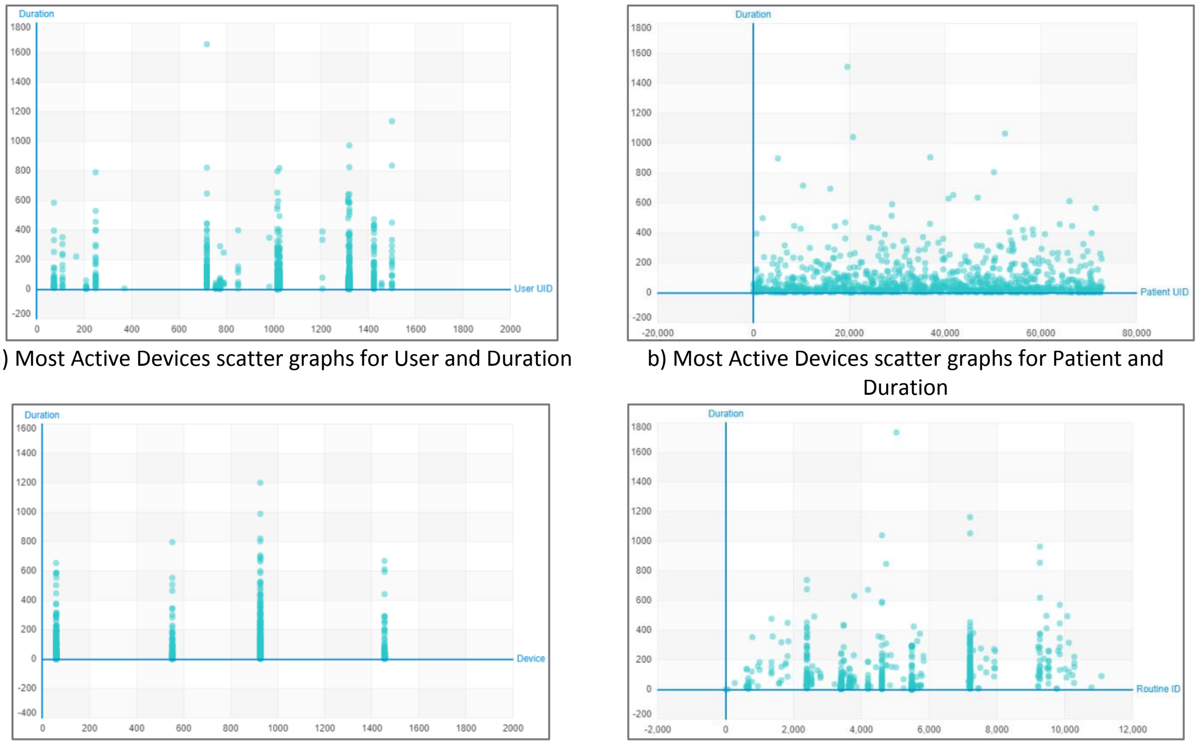


a) Device ID Profiles

b) Device ID Profiles (exc. Frequency and Max)

Figure 40 – Device ID Profile Line Graphs

Figure 41 displays the 5 most active Device IDs and compares each ID with duration for (a) User, (b) Patient, (c) Device and (d) Routine.



a) Most Active Devices scatter graphs for User and Duration

b) Most Active Devices scatter graphs for Patient and Duration

c) Most Active Devices scatter graphs for Device and Duration

d) Most Active Devices scatter graphs for Routine and Duration

Figure 41 - Most Active Devices scatter graphs

In Figure 41(a), the density of the scatter graph is under 400 seconds, with some users accessing a device for no longer than a few seconds. However, one user does perform a routine on a device for over 1,600 seconds. Additionally, in Figure 41(b), 300 seconds (5 minutes) or less is typically spend on a device performing an action on a patient. Figure 41(c) demonstrates similar observations, with varying datapoints but it is atypical for a device to be used for longer than approximately 600 seconds. Finally, Figure 41(d) again demonstrates that 400 seconds is the typical time spent on a device in the dataset, with some exceptions, such as the routine that is accessed for 1,700 seconds. Therefore, 400 seconds is the initial benchmark for accessing devices in the dataset.

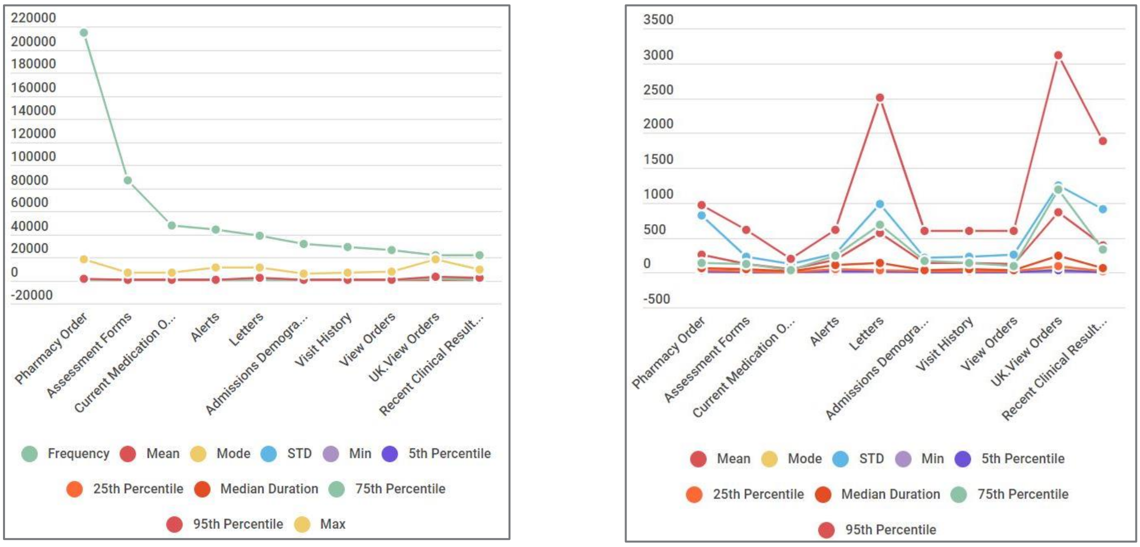
5.3.4. Profiling Routine ID

The most active profiles for Routine ID are displayed in Table 16. For User, Patient and Device only the 5 most active IDs are profiled. However, for Routines a more comprehensive view of profiling is required as each routine requires a distinct profile to understand typical behaviour for each routine performed within the EPR.

Table 16 - Profiles for Routine ID

Routine ID	Frequency	Mean	Mode	STD	Min	5th Percentile	25th Percentile	Median Duration	75th Percentile	95th Percentile	Max
Pharmacy Order	214345	252.22	24	818.53	1	12	30	56	131	968	18471
Assessment Forms	86872	128.66	7	228.04	2	6	16	41	117	616	6767
Current Medication Orders	47987	45.23	5	119.33	1	4	8	15	31	189	6290
Alerts	43696	188.02	23	276.78	1	16	44	101	244	617	10769
Letters	39091	567.71	15	986.00	3	13	38	134	689	2502.5	10700
Admissions Demographics Data	31402	140.52	9	211.89	1	6	16	37	160	593	5577
Visit History	28622	137.82	10	224.30	2	8	18	44	142	594	7014
View Orders	26227	123.29	10	250.20	2	6	15	31	93	603	7294
UK.View Orders	21971	863.34	70	1242.84	2	27	88	239	1188.5	3107.5	18380
Recent Clinical Results: Department Reports	21626	382.52	16	912.81	2	9	24	61	324	1878	9083

Figure 42(a) shows the 10 most frequently accessed routines in a line graph. Max and Frequency are removed in Figure 42(b).

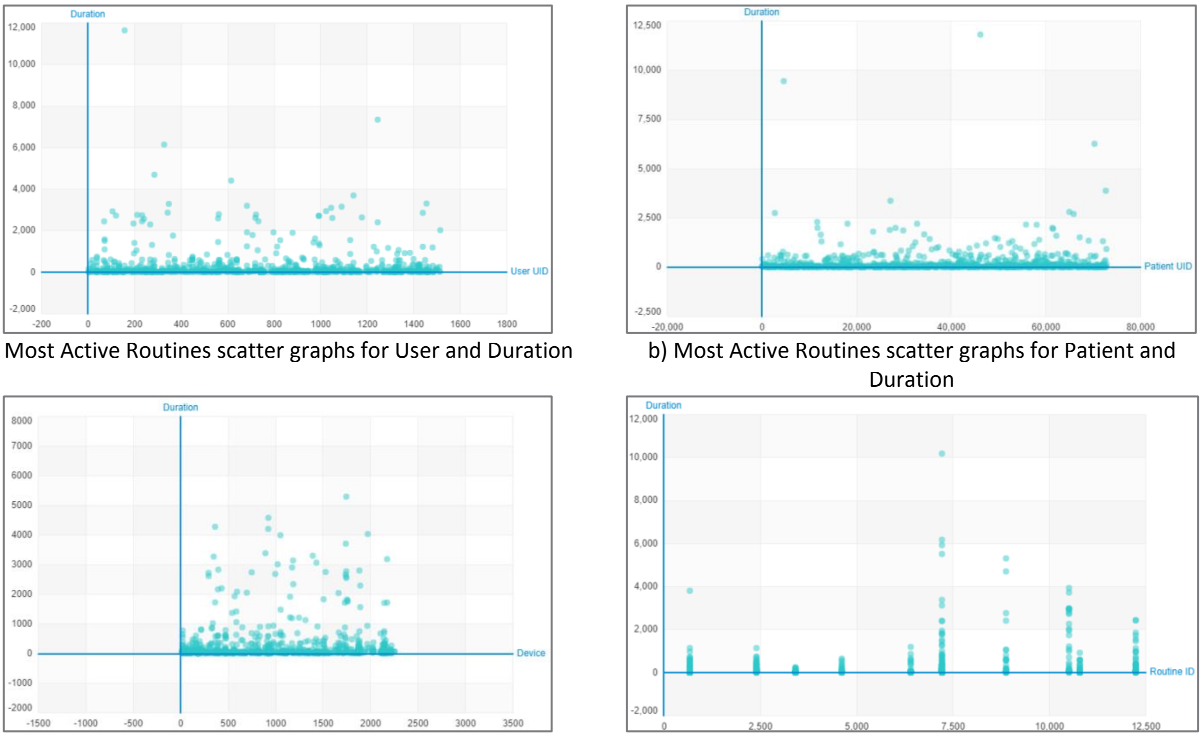


a) Routine ID Profiles

b) Routine ID Profiles (exc. Frequency and Max)

Figure 42 – Routine ID Profile Line Graphs

Figure 43 displays the 5 most active Routine IDs and compares each ID with duration for (a) User, (b) Patient, (c) Device and (d) Routine.



a) Most Active Routines scatter graphs for User and Duration

b) Most Active Routines scatter graphs for Patient and Duration

c) Most Active Routines scatter graphs for Device and Duration

d) Most Active Routines scatter graphs for Routine and Duration

Figure 43 - Most Active Routines scatter graphs

In Figure 43(a), the density of the scatter graph is around 1,000 seconds. Due to the routines having extreme anomalies within the dataset (such as 12,000 seconds), the scale makes observations difficult to determine for routine ID. In Figure 43(b), a routine is typically performed in under 1,000 seconds on a patient. Figure 43(c) also has 1,000 seconds as a typical behaviour benchmark, although the number of devices that exceed that threshold is much greater than for Figure 43(a) and Figure 43(b). Finally, Figure 43(d) demonstrates that most routines have consistent usage times within the dataset of 1,000 seconds, however others have much more variety, with datapoint exceeding 10,000 seconds, whereas others never exceed 1,000 seconds. Therefore, 1,000 seconds is the initial benchmark for typical routine behaviour in the dataset.

5.3.5. Discussion

A behaviour profile of each ID type is created. Typical behaviour is determined in **Error! Reference source not found.**. If activity significantly deviates from this it should be flagged to an analyst for review. This activity represents the baselines that activity will be measured against. Through the application of HILML each ID benchmark will be refined as results are presented to an analyst.

Table 17 – Typical Behaviour

	User ID	Patient ID	Device ID	Routine ID
Duration (Secs)	400	1,000	400	1,000

Limitations are presented here due to the tokenised nature of the data. If the data had not been tokenised more insights would be extracted at this stage. For example, all Users do not interact with the EPR comparably. A doctor would access the EPR for typically different routines from those that a nurse would. Similarly, a pharmacist would interact in a different way to an administrative member of staff. If this data was available, role based behavioural benchmarks would be produced, giving more nuanced results. Similarly, for Patient ID, complex patients, or VIPs may need to be monitored differently to typical patients. A device used on a busy ward will be used differently to a device used in a doctor’s office, or on a reception desk.

5.4. Summary

In this chapter, a case study is presented for the real-world dataset provided by the specialist Liverpool based hospital. The extraction of the data is discussed and the tokenisation process is described. An exploration of the data including a distribution of values is presented, with emphasis on the anomalous nature of max values, which significantly deviate from all other data points, including in the 0.99 quantile. In chapter 6, the results are detailed and discussed for three datasets, performing the system framework on each dataset. The results of each dataset are then compared and discussed.

6. Results and Rationale Discussion

6.1. Introduction

The PARISS process is applied to the dataset and the results are detailed in this chapter. Initial Results of one month (May 2017) of data are detailed in section 6.2 as a benchmark experiment comparing the LOF approach with DBSCAN. Follow up results of six months of data (July-December 2016) are detailed in section 6.3. Finally, the process is applied to 18 months of data and these results are discussed in section 6.4. A summary and discussion are provided in section 6.5.

PARISS uses as much information as possible to determine an anomaly score for each ID and therefore each audit log. However, with this dataset, due to the way the routine ID is captured, it may give misleading results. In order for the LOF scores for routine to be of value for routine ID, each routine (rather than the routine combination) would need to be calculated. Unfortunately, this cannot be differentiated within the dataset. For each Routine ID, all routines involved in that interaction with the User, Device and Patient are stored as the Routine ID. Rather than separate Routine IDs for each routine performed on the patient record within that interaction. For example, if the LOF scores for each routine are calculated individually (rather than as a routine set), such as ‘Assessment Forms’ and ‘Maternity Data’, then these values can be compared with other instances of that routine, to determine whether certain log accesses are anomalous. However, as these cannot be separated within the combinations of routines, then an informative LOF score cannot be determined for

Routine ID. Therefore, an anomaly score is calculated both including and excluding the routine ID in order to determine which is more effective for the dataset in the case study.

6.2. Initial Results

The initial dataset consists of one month of data in May 2017. Based on the background research, LOF is selected for use in PARISS; however, DBSCAN experimentation is included for comparison. For the initial results, the data contains 60,454 rows. There are 828 User IDs, 11,068 Patient IDs, 1,123 Device IDs and 1,891 Routine IDs.

6.2.1. Feature Extraction

Table 18 displays the eleven features extracted for the ten most frequent user IDs. Each of the 14,910 unique IDs in the dataset has these eleven features extracted, resulting in 164,010 features in total.

Table 18 - Feature Extraction for User ID (One Month)

User ID	Frequency	Mean	Mode	STD	Min	5th Percentile	25th Percentile	Median Duration	75th Percentile	95th Percentile	Max
720	2191	51.74	2	87.55	0	2	11	22	50	209.5	1012
556	1870	118.25	2	245.48	0	4	22	59	138	385.2	7007
202	1796	93.11	2	213.37	0	1	2	4	44	557	1969
551	1607	67.85	2	126.21	0	2	5	20	69	312.8	1436
354	1166	204.78	2	235.12	1	2	31	130	298.25	636.5	3421
405	1124	24.75	5	107.7	0	2	5	7	19	83.85	3240
295	1122	202.37	1	377.63	0	6	49	113	230.75	600.9	6142
355	845	65.83	11	89.56	1	6	19	42	74	200.4	909
119	807	201.31	6	511.81	1	6	24.5	77	223	633.5	7331
138	758	132.98	45	253.69	0	17	41	73	139	370.15	4150

6.2.2. Feature Scaling

Figure 44 displays the Min-Max feature scaling of the dataset.

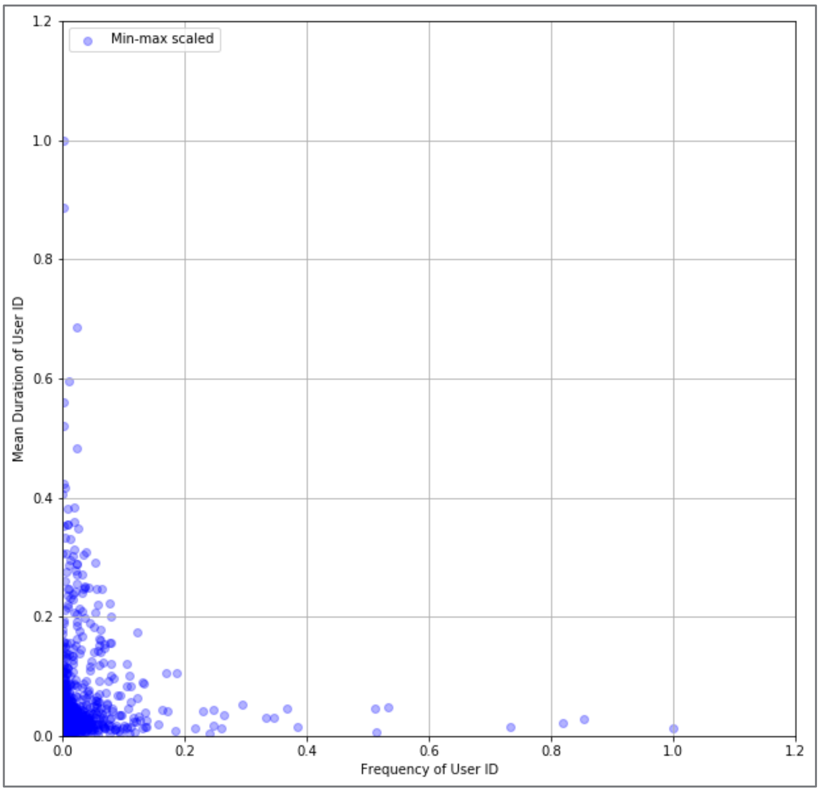


Figure 44 - Min-Max Feature Scaling (One Month)

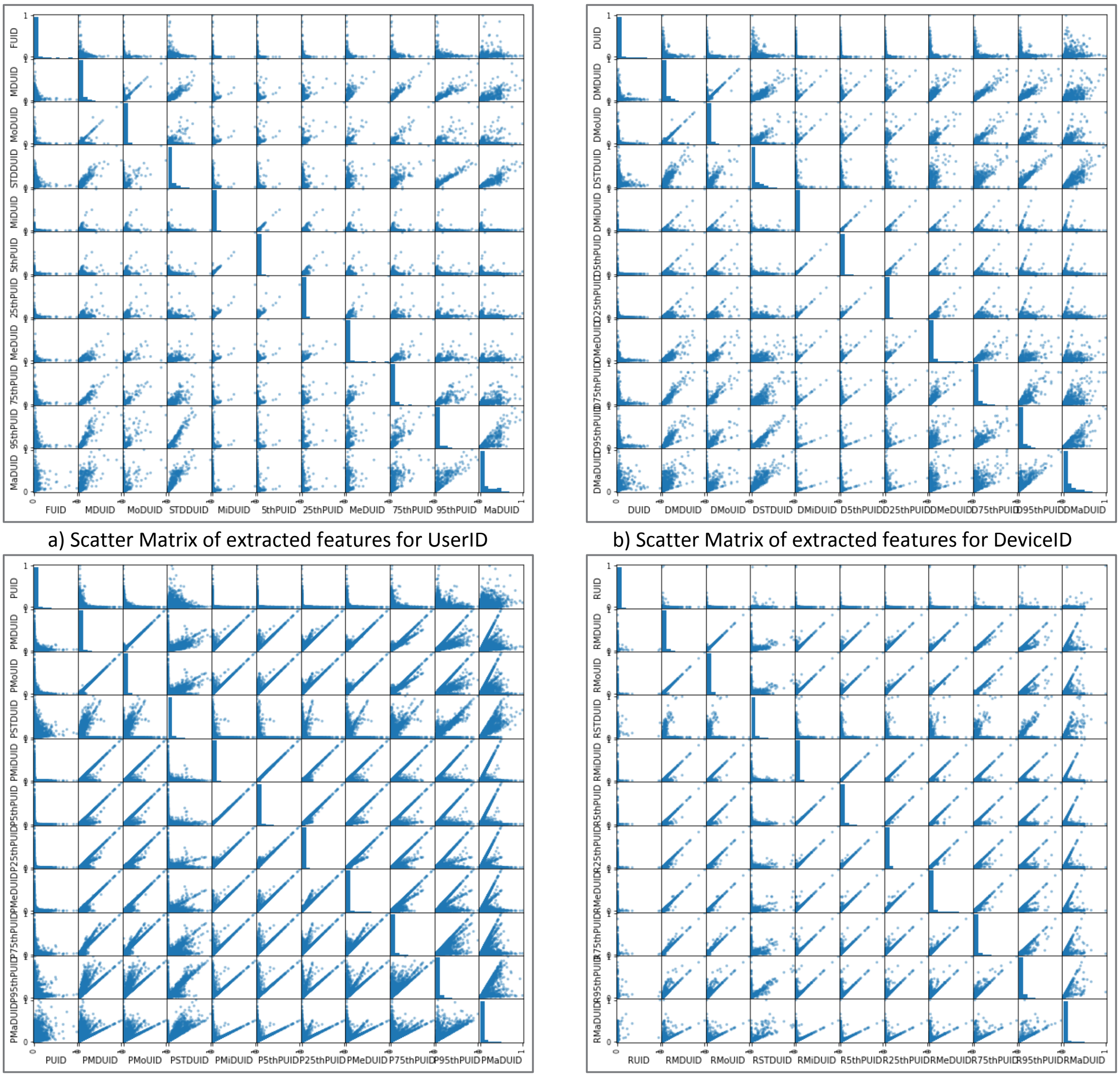
Table 19 displays the top ten most frequent User IDs, with min-max scaling applied to their features.

Table 19 - MinMax Scaled Features for User ID (One Month)

User ID	Frequency	Mean	Mode	STD	Min	5th Percentile	25th Percentile	Median Duration	75th Percentile	95th Percentile	Max
720	1.000	0.012	0.000	0.016	0.000	0.001	0.005	0.005	0.008	0.020	0.067
556	0.853	0.027	0.000	0.046	0.000	0.002	0.009	0.013	0.021	0.037	0.461
202	0.820	0.021	0.000	0.040	0.000	0.000	0.000	0.001	0.007	0.054	0.130
551	0.733	0.016	0.000	0.023	0.000	0.001	0.002	0.004	0.011	0.030	0.094
354	0.532	0.047	0.000	0.044	0.001	0.001	0.014	0.029	0.046	0.062	0.225
405	0.513	0.006	0.001	0.020	0.000	0.001	0.002	0.001	0.003	0.008	0.213
295	0.512	0.047	0.000	0.070	0.000	0.003	0.022	0.025	0.036	0.058	0.404
355	0.385	0.015	0.002	0.017	0.001	0.003	0.008	0.009	0.011	0.019	0.060
119	0.368	0.046	0.001	0.095	0.001	0.003	0.011	0.017	0.034	0.061	0.483
138	0.346	0.031	0.010	0.047	0.000	0.010	0.018	0.016	0.021	0.036	0.273

6.2.3. Feature Testing

Figure 45 displays the feature testing scatter matrix of the extracted features.



a) Scatter Matrix of extracted features for UserID

b) Scatter Matrix of extracted features for DeviceID

c) Scatter Matrix of extracted features for PatientID

d) Scatter Matrix of extracted features for Routine

Figure 45 - Scatter Matrix of extracted features

6.2.4. DBSCAN: User, Patient, Device and Routine ID

As a comparison to the LOF results, the DBSCAN algorithm is applied to the initial dataset. DBSCAN is selected for the comparison with LOF as they both use a core and a reachability distance in order to determine outliers (as outlined in Chapter 3). Table 20 presents 20 rows of DBSCAN results for User ID, Patient ID, Device ID and Routine ID. DBSCAN does not apply a weighted score to the results, therefore the results are classified as one of three point types. A core point is classified as a point that belongs to a cluster. A boundary point is within the epsilon of a core point, but does not meet the criteria of min_points to be considered a core point. Finally, noise points are not assigned to any cluster.

Table 20 - DBSCAN Benchmark Results Example

User ID	User ID cluster_id	User ID type	Patient ID	Patient ID cluster_id	Patient ID type	Device ID row_id	Device ID cluster_id	Device ID type	Routine ID row_id	Routine ID cluster_id	Routine ID type
119	n/a	noise	**803**	n/a	noise	**1**	n/a	noise	**MPI ZCUS.UK.SCH ZCUS.UK.LETTER**	n/a	noise
126	n/a	noise	**804**	n/a	noise	**2**	n/a	noise	**ZCUS.UK.LETTER VH SPC OE**	n/a	noise
144	n/a	noise	**805**	n/a	noise	**3**	n/a	noise	***** ASF**	6	core
203	n/a	noise	**806**	n/a	noise	**4**	n/a	noise	***** ASF MPI**	6	core
226	n/a	noise	**807**	n/a	noise	**5**	n/a	noise	***** ASF NOTE ZCUS.UK.LETTER**	6	core
297	n/a	noise	**4764**	n/a	noise	**6**	n/a	noise	***** ASF NPC**	6	core
359	n/a	noise	**4765**	n/a	noise	**7**	n/a	noise	***** ASF SPC VH ZCUS.UK.SCH SPCUS**	6	core
404	n/a	noise	**4766**	n/a	noise	**8**	n/a	noise	***** ASF SS ZCUS.UK.SCH**	6	core
432	n/a	noise	**4767**	n/a	noise	**9**	n/a	noise	***** ASF ZCUS.UK.LETTER**	6	core
442	n/a	noise	**4768**	n/a	noise	**10**	n/a	noise	***** ASF ZCUS.UK.LETTER ZCUS.UK.SCH**	6	core
526	n/a	noise	**6674**	n/a	noise	**11**	n/a	noise	***** ASF ZCUS.UK.SCH BD**	6	core
770	n/a	noise	**8763**	n/a	noise	**12**	n/a	noise	***** BD**	6	core
775	n/a	noise	**8764**	n/a	noise	**13**	n/a	noise	***** BD CM NOTE**	6	core
793	n/a	noise	**6**	19	core	**299**	n/a	noise	***** BD UK.OE VH OE**	6	core
795	n/a	noise	**7**	19	core	**300**	n/a	noise	***** CM**	6	core
1	3	core	**8**	19	core	**301**	n/a	noise	***** CM PHA.ORDS**	6	core
6	6	core	**10**	19	core	**302**	n/a	noise	***** LAB.DRP**	6	core
14	6	core	**11**	19	core	**303**	n/a	noise	***** LAB.DRP UK.OE REC**	6	core
18	7	core	**12**	19	core	**304**	n/a	noise	***** MED**	6	core
28	0	core	**13**	19	core	**305**	n/a	noise	**MPI ZCUS.UK.SCH ZCUS.UK.LETTER**	6	core

In Table 21 the number of core, boundary and noise types for each of the IDs are presented. Core Points are points within a DBSCAN cluster. Boundary points are within reachability distance of a cluster. Noise points are not within reachability distance and are therefore considered outliers.

Table 21 - DBSCAN point types for User ID, Patient ID, Device ID and Routine ID

	Core	Boundary	Noise
User ID	331	482	15
Patient ID	6,822	4,233	14
Device ID	440	595	88
Routine ID	1,184	705	2

In Figure 46 the number of core, boundary and noise types are presented as a bar chart.

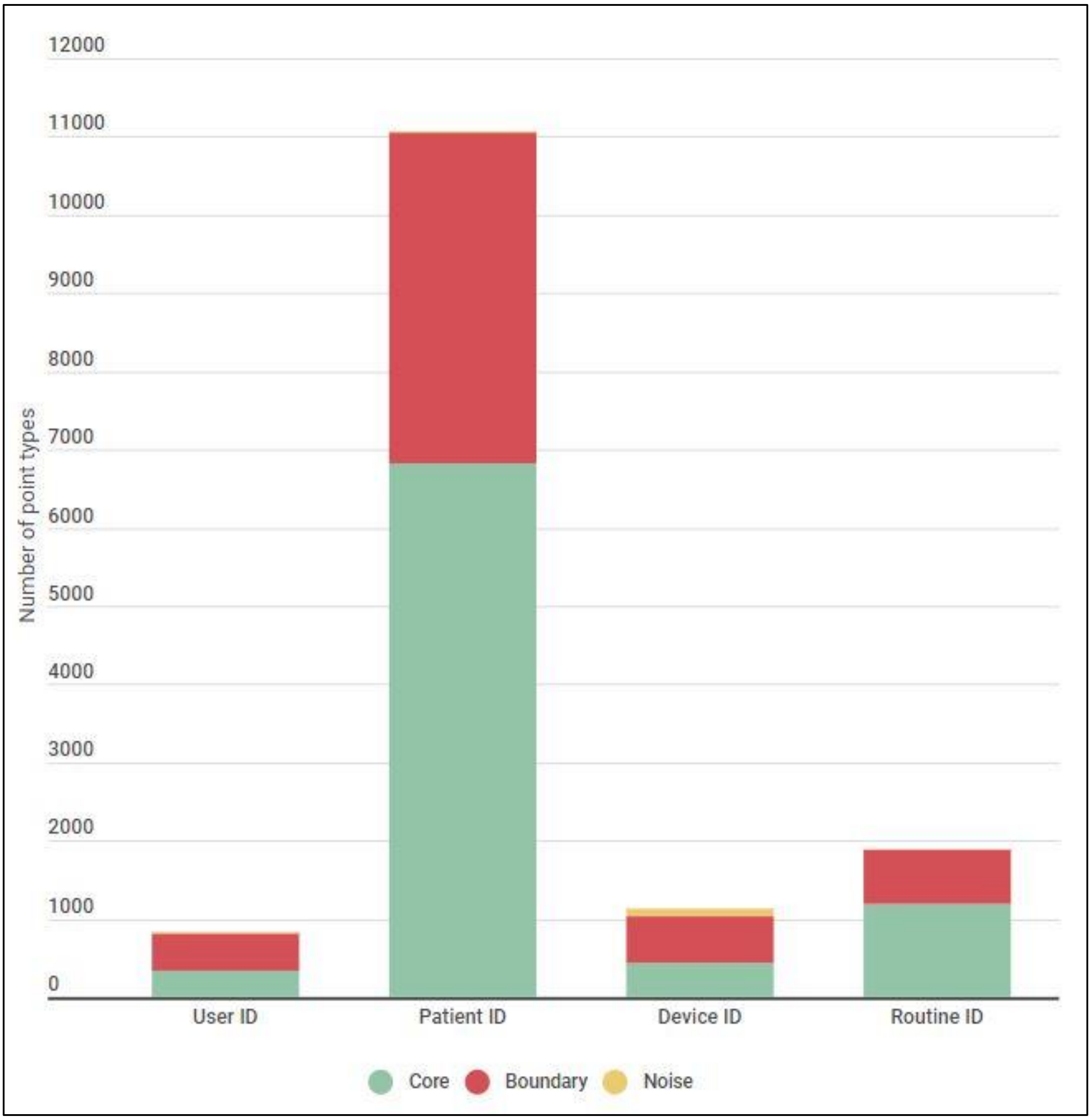


Figure 46 – Bar chart of core, boundary and noise types for DBSCAN results

Due to the lack of a weighted score, DBSCAN does not allow a patient privacy officer to prioritise their investigation into potentially inappropriate behaviour. Once the officer has investigated the noise points, there is the ‘needle-in-a-haystack’ problem of investigating border points. This is insufficient, and a weighted anomaly score is required to enable more nuanced investigation. Therefore, based on the benchmark experimentation, LOF is selected for use in the PARISS system. A full breakdown of the DBSCAN results can be seen in Table 40 in the appendix.

6.2.5. LOF: User, Patient and Device ID

A LOF score is calculated for the 45 combinations (of the 10 features) for each of the 14,910 unique IDs in the dataset. Therefore 670,950 unique LOF scores are calculated in total. An average is then taken to assign an anomaly score. In Table 22, LOF identifies anomalous User IDs, Patient IDs and Device IDs. The neighbourhood radius is defined in stage 3 of the LOF algorithm (Section 3.2.3.1), the density score is defined in stage 4, and the anomaly score is the final LOF value, as defined in stage 5.

Table 22 - LOF (Mean) Anomaly Scores for User ID, Patient ID and Device ID

User ID	Density Score	Anomaly Score	Neighbourhood Radius	Patient ID	Density Score	Anomaly Score	Neighbourhood Radius	Device ID	Density Score	Anomaly Score	Neighbourhood Radius
821	5.65	3.85	0.49	**3760**	118.27	9.32	0.02	**410**	2.66	2.73	0.63
717	149.13	3.35	0.01	**2214**	1062.91	7.33	0.00	**273**	206.22	2.46	0.01
804	18.35	3.15	0.16	**6879**	692.70	6.85	0.00	**331**	286.50	2.39	0.01
813	9.56	2.91	0.24	**2482**	651.25	6.58	0.00	**931**	34.78	2.22	0.04
828	5.79	2.81	0.60	**1293**	718.43	6.45	0.00	**956**	686.16	2.14	0.00
799	30.79	2.77	0.12	**4534**	905.26	6.27	0.00	**307**	212.63	2.11	0.01
822	9.52	2.76	0.13	**3194**	807.05	5.95	0.00	**75**	763.83	2.11	0.00
718	125.30	2.63	0.02	**5124**	547.38	5.89	0.00	**818**	7.36	1.93	0.22
715	209.59	2.52	0.01	**5028**	1695.80	5.70	0.00	**12**	4.28	1.92	0.35
827	152.55	2.28	0.38	**5821**	1695.80	5.70	0.00	**342**	28.50	1.90	0.04

Within the User ID range, the most notable ID is #821, with an anomaly score of 3.852. There are 14 User IDs with an anomaly score above 2. Therefore, LOF has indicated that 2.69% of the User IDs are anomalous. Similarly, the most notable Patient ID is #3760, with an anomaly score of 9.32. There are 82 Patient IDs with an anomaly score above 2; indicating 0.74% of the Patient IDs are anomalous. Finally, the most notable Device ID is #410, with an anomaly score of 2.73. There are 7 Device IDs with an anomaly score above 2, indicating that 0.62% of the Device IDs are irregular. Overall therefore, LOF identifies 1.35%

of IDs as anomalous, which would be highlighted to a patient privacy officer for investigation.

6.2.6. LOF: Routine ID

However, the LOF technique cannot be applied as effectively to the Routine ID. This is due to the concatenation of every routine performed during the interaction with the patient being recorded as a single routine within the dataset. Table 23 presents a sample of the highest LOF anomaly scores for the Routine ID dataset.

Table 23 - LOF (Mean) Anomaly Scores for Routine ID

Routine Set Description	Density Score	Anomaly Score	Neighbourhood Radius
RAD.DRP ZCUS.UK.SCH ASF	1105.03	8.40	0.00
REC ASF MED	1050.68	8.17	0.00
SPC PHA.ORDS	587.39	7.34	0.00
PHA.ORDS ASF ZCUS.UK.SCH	1153.80	6.92	0.00
SS OE	559.91	6.64	0.00
VH ASF NPC UK.OE MPI ZCUS.UK.SCH SS PHA.ORDS	568.13	6.47	0.00
SS ZCUS.UK.LETTER WL	368.16	5.85	0.01
SS MPI ASF ZCUS.UK.LETTER	707.26	5.82	0.00
SS ASF	648.75	5.30	0.00
PHA.ORDS REC REC:(DRP)	658.48	4.94	0.00

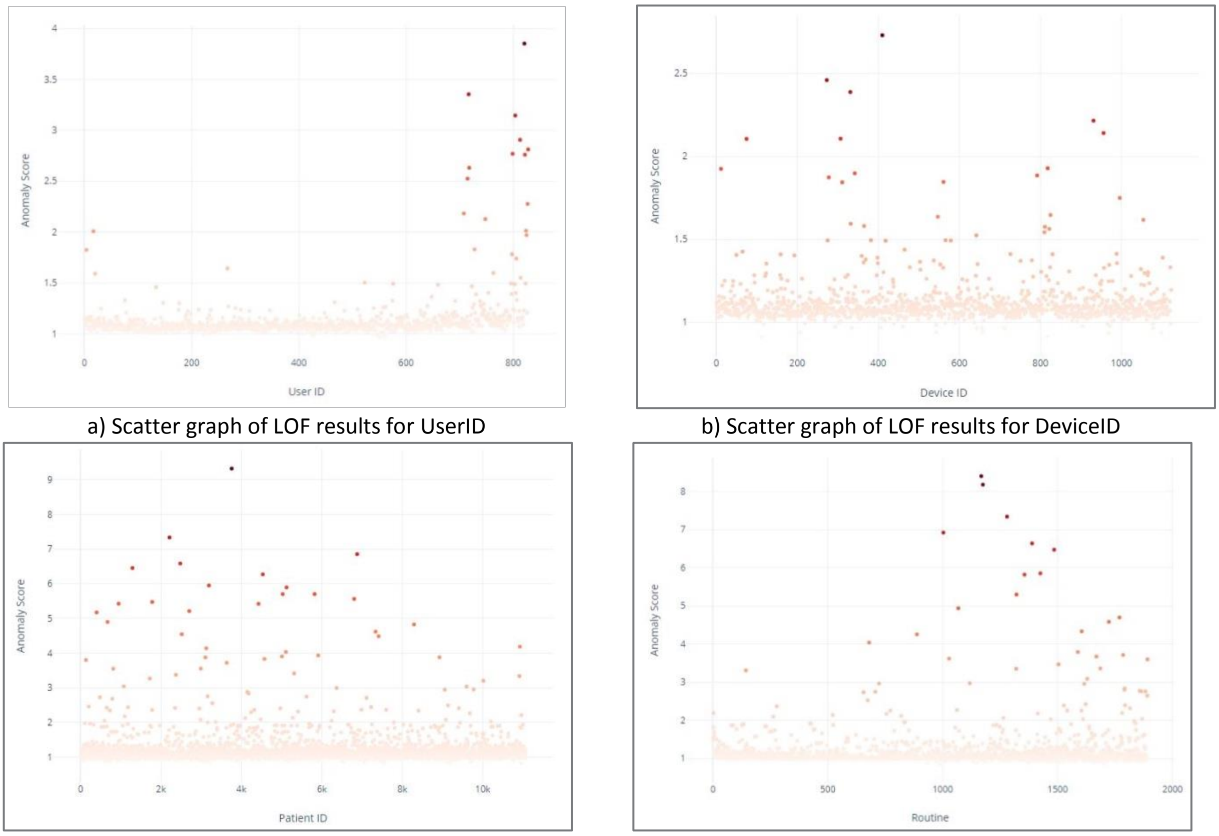
There are 57 routine sets with an anomaly score above 2. Therefore, LOF has indicated that 3.01% of the routine sets are anomalous. The most notable routine set is the combination 'Radiology Reports | Cancelled Account.UK.Scheduling | Assessment Forms', with an anomaly score of 8.40.

6.2.7. Quantifying LOF

The data is quantified as per the process outlined in chapter 4. NAN and Inf values are replaced with 1 and 2 respectively, whilst missing or null values are assigned the median value for their feature class.

6.2.8. Visualisation of LOF Results

A visualisation of the LOF results for each ID is presented in Figure 47.



a) Scatter graph of LOF results for UserID

b) Scatter graph of LOF results for DeviceID

c) Scatter graph of LOF results for PatientID

d) Scatter graph of LOF results for Routine

Figure 47 – Scatter graph of LOF results

6.2.9. LOF: Anomaly Score Ensemble Averaging

A sample of EPR data with a calculated ensemble average LOF anomaly score is presented in Table 24. The table is ordered by the highest LOF anomaly scores. Within the date range, the most notable audit log occurred on 16th May 2017 at 02:56. User #344 accessed Patient #3760 on Device #951 performing the following Routine combination ‘UK.View Orders View Orders’, with an anomaly score of 3.18. There are 241 audit logs with an anomaly score above 2. Therefore, PARISS has indicated that 0.399% of the EPR Audit Logs are anomalous.

Table 24 - EPR Audit Data with Ensemble Averaging Applied to LOF Anomaly Score (including Routine)

Date & Time	Device	Device Anomaly Score	User ID	User Anomaly Score	Routine ID	Routine Anomaly Score	Patient ID	Patient Anomaly Score	Duration (sec)	Adm Date	Dis Date	Ensemble Averaging Anomaly Score (inc Routine)
16/05/17 02:46	951	1.251	344	1.108	UK.OE OE	1.041	3760	9.319	1741	22/07/2011	22/07/2011	**3.180**
16/05/17 01:40	951	1.251	344	1.108	UK.OE	0.967	3760	9.319	1369	22/07/2011	22/07/2011	**3.161**
26/05/17 03:20	141	1.047	701	1.080	RAD.DRP ZCUS.UK.SCH ASF	8.397	5574	1.051	57	17/09/2012	17/09/2012	**2.894**
26/05/17 15:48	141	1.047	701	1.080	RAD.DRP ZCUS.UK.SCH ASF	8.397	5574	1.051	193	17/09/2012	17/09/2012	**2.894**
01/05/17 13:33	1046	1.040	800	1.241	REC ASF MED	8.175	3657	1.045	114	05/09/2016	05/09/2016	**2.875**
08/05/17 20:48	498	1.015	437	1.080	ZCUS.UK.SCH	2.806	2482	6.581	1448	12/12/2016	12/12/2016	**2.870**
19/05/17 23:50	278	1.873	479	1.092	UK.OE	0.967	2214	7.334	451	27/02/2017	27/02/2017	**2.817**
30/05/17 17:20	310	1.092	67	1.102	SS	3.353	6805	5.561	597	08/03/2006	08/03/2006	**2.777**
17/05/17 15:59	794	1.140	556	1.093	SPC PHA.ORDS	7.338	1506	1.067	112	07/02/2012	07/02/2012	**2.660**
25/05/27 20:34	1111	1.140	114	1.105	SPC PHA.ORDS	7.338	994	1.053	100	06/07/2008	06/07/2008	**2.659**

A sample of EPR data with a calculated ensemble average LOF anomaly score (excluding Routine ID) is presented in Table 25. The table is ordered by the highest LOF anomaly scores. Within the date range, the most notable audit log occurred on 16th May 2017 at 01:40. User #344 accessed Patient #3760 on Device #951 performing the following Routine combination 'UK.View Orders', with an anomaly score of 3.892. There are 145 audit logs with an anomaly score above 2. Therefore, PARISS has indicated that 0.182% of the EPR Audit Logs are anomalous.

Table 25 - EPR Audit Data with Ensemble Averaging Applied to LOF Anomaly Score (excluding Routine)

Date & Time	Device	Device Anomaly Score	User ID	User Anomaly Score	Routine ID	Routine Anomaly Score	Patient ID	Patient Anomaly Score	Duration (sec)	Adm Date	Dis Date	Ensemble Averaging Anomaly Score (exc Routine)
16/05/17 01:40	951	1.251	344	1.108	UK.OE	0.967	3760	9.319	1369	22/07/2011	22/07/2011	**3.892**
16/05/17 02:46	951	1.251	344	1.108	UK.OE OE	1.041	3760	9.319	1741	22/07/2011	22/07/2011	**3.892**
19/05/17 23:50	278	1.873	479	1.092	UK.OE	0.967	2214	7.334	451	27/02/2017	27/02/2017	**3.433**
24/05/17 15:55	280	1.206	479	1.092	UK.OE	0.967	2214	7.334	401	27/02/2017	27/02/2017	**3.211**
24/05/17 03:04	413	1.302	612	1.067	ZCUS.UK.LETTER	1.323	6879	6.853	550	21/02/2007	21/02/2007	**3.074**
24/05/17 04:22	413	1.302	612	1.067	ZCUS.UK.LETTER	1.323	6879	6.853	487	21/02/2007	21/02/2007	**3.074**
09/05/17 02:53	1025	1.056	718	2.631	MPI	1.056	403	5.171	24	21/01/2014	21/01/2014	**2.953**
05/05/17 06:29	593	1.112	200	1.094	ZCUS.UK.LETTER VH OE	0.976	2482	6.581	368	12/12/2016	12/12/2016	**2.929**
08/05/17 17:28	476	1.099	677	1.103	OE	1.000	2482	6.581	649	12/12/2016	12/12/2016	**2.928**
22/05/17 20:55	496	1.119	701	1.080	OE ZCUS.UK.LETTER ASF RAD.DRP ZCUS.UK.SCH	1.000	2482	6.581	294	12/12/2016	12/12/2016	**2.927**

In comparison, the results in Table 24 and Table 25 are similar. The most anomalous audit log in Table 24 (16th May 2017 at 02:46) is the second most anomalous audit log in Table 25. Similarly the most anomalous audit log in Table 25 (16th May 2017 at 01:40) is the second most anomalous audit log in Table 24. There are many other similar results appearing in both tables. This is likely due to the fact that the patient anomaly scores in the one-month dataset is high, with anomaly scores above 5, compared to the routine anomaly scores ranging around 1, indicating typical behaviour. In the case of the one-month dataset, the routine anomaly score can be included without offsetting the results.

6.2.10. Visualisation of LOF Audit Log Results

In Figure 48, a visualisation of LOF results of calculated ensemble averaged anomaly scores is displayed for all 60,454 audit logs. The x-axis displays the date, and the y-axis displays the calculated ensemble average anomaly score.

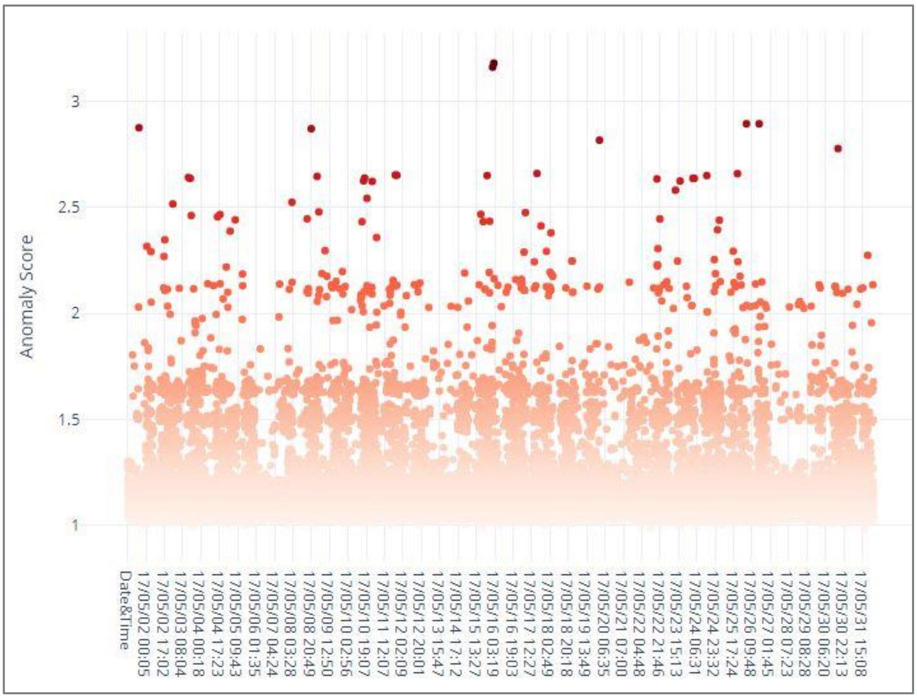


Figure 48 - Visualisation of LOF Results for Ensemble Averaged Anomaly Scores (inc Routine)

In Figure 49, a visualisation of LOF results of calculated ensemble averaged anomaly scores (excluding the Routine ID anomaly score). As discussed in section 4.3.8, outliers in the top quarter of each ID range are highlighted as red, with the intensity of colour reducing as the anomaly score reduces. This creates an interactive live task list for the analyst, with an anomaly priority ordering.

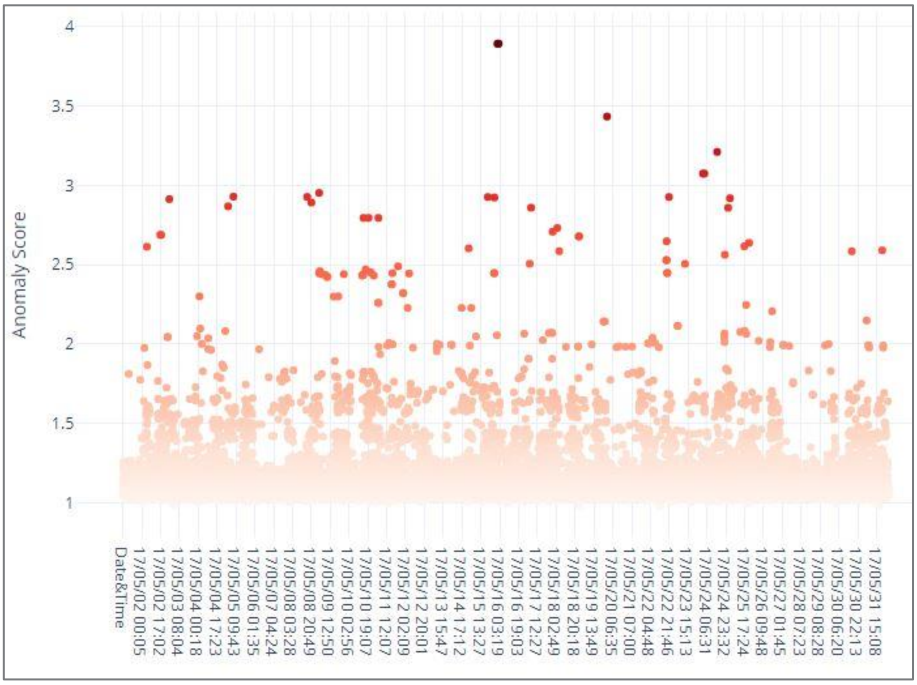


Figure 49 - Visualisation of LOF Results for Ensemble Averaged Anomaly Scores (exc Routine)

6.3. 6 Months EPR Data

The second dataset consists of six months of data from July-Dec 2016. The data contains 340,687 rows of data. There are 1,120 User IDs, 35,159 Patient IDs, 1,671 Device IDs and 6,690 Routine IDs.

6.3.1. Feature Extraction

Table 26 displays the eleven features extracted for the ten most frequent user IDs. Each of the 44,640 unique IDs in the dataset has these eleven features extracted, resulting in 491,040 features in total.

Table 26 - Feature Extraction for User ID (6 Months)

User ID	Frequency	Mean	Mode	STD	Min	5th Percentile	25th Percentile	Median Duration	75th Percentile	95th Percentile	Max
762	13351	44.08	2	97.75	0	1	5	18	37	186.5	4751
79	8867	79.8	2	133.35	0	2	11	31	102	306	3521
548	6969	28.23	6	98.90	0	2	5	8	20	111	6081
768	6604	137.72	2	335.25	0	5	24	68	150.25	411.7	7469
971	6223	88.43	19	143.81	0	3	17	33	94.5	379	1739
188	5941	129.63	21	288.89	0	16	36	70	133	382	7169
480	5674	222.36	291	237.98	0	2	35	148	323	653.35	3458
481	5640	80.90	23	153.86	1	8	24	46	81	274	5855
1054	4767	130.35	2	201.86	0	4	29	77	164	419	6268
210	4006	62.23	15	170.20	1	7	16	30	56	191.5	4839

6.3.2. Feature Scaling

Figure 50 displays the Min-Max feature scaling of the dataset.

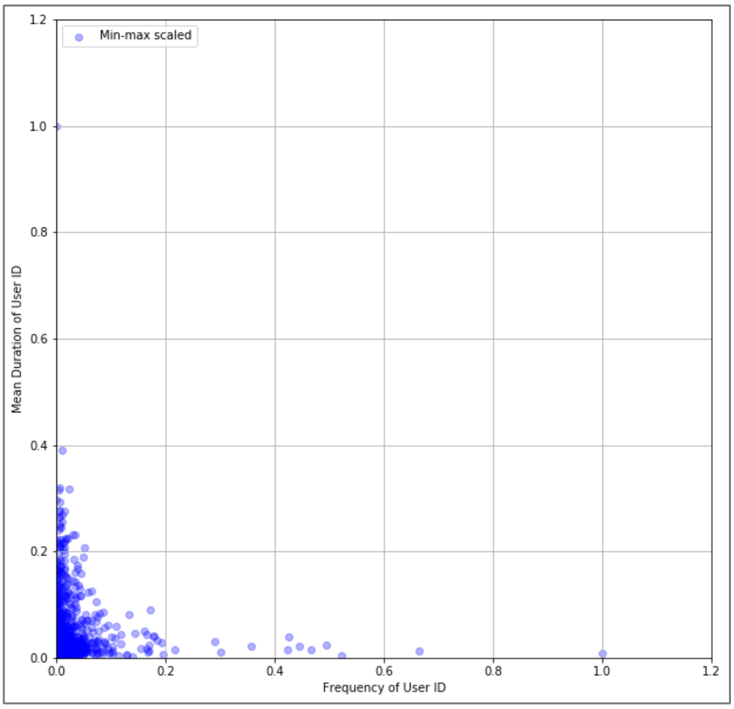


Figure 50 - Min-Max Feature Scaling (Six Months)

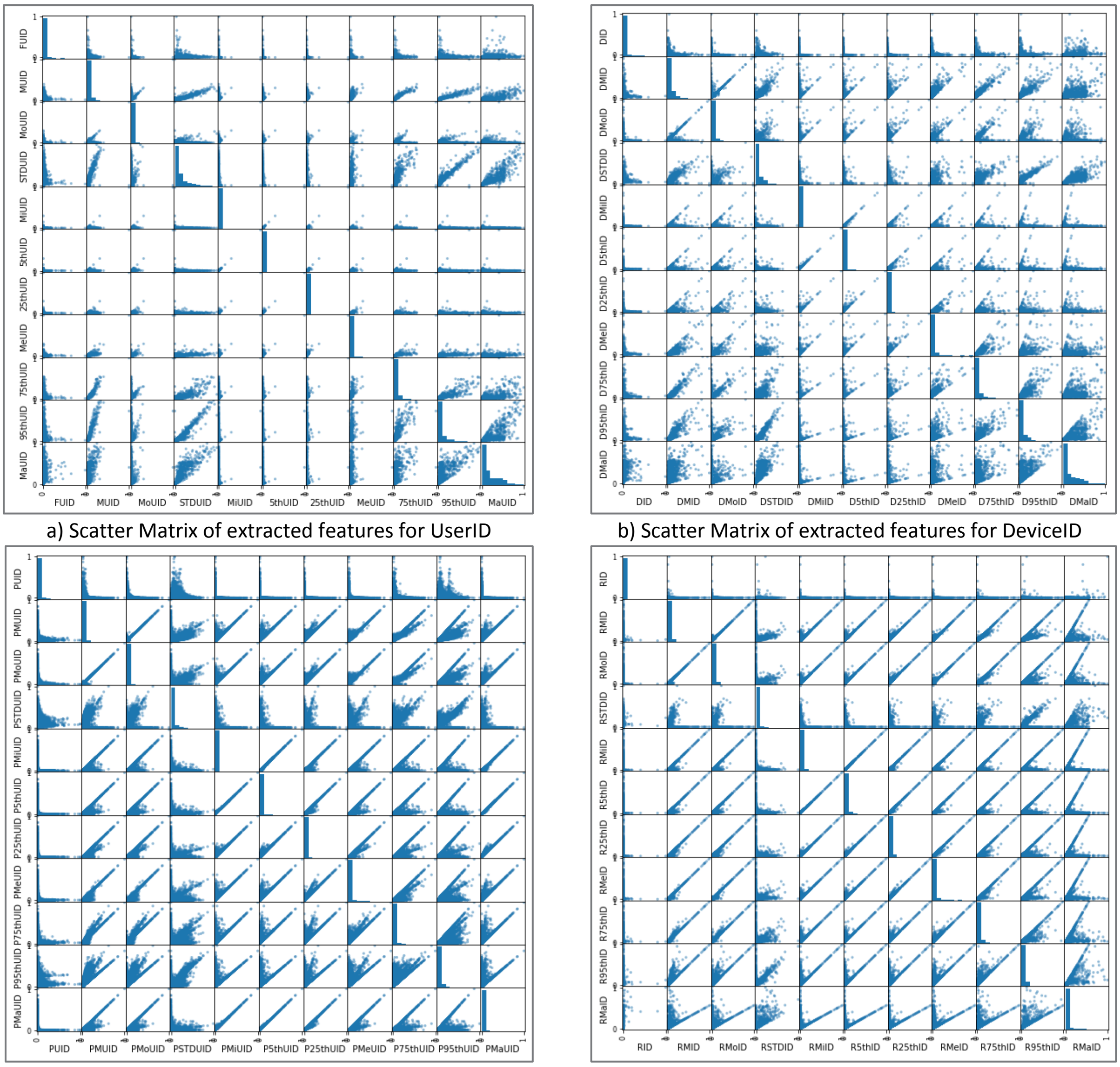
Table 27 displays the top ten most frequent User IDs, with min-max scaling applied to their features.

Table 27 - MinMax Scaled Features for User ID (Six Months)

User ID	Frequency	Mean	Mode	STD	Min	5th Percentile	25th Percentile	Median Duration	75th Percentile	95th Percentile	Max
762	1.000	0.008	0.000	0.034	0.000	0.000	0.001	0.003	0.006	0.024	0.340
79	0.664	0.014	0.000	0.046	0.000	0.000	0.002	0.005	0.018	0.040	0.252
548	0.522	0.005	0.001	0.034	0.000	0.000	0.001	0.001	0.003	0.014	0.436
768	0.495	0.024	0.000	0.116	0.000	0.001	0.004	0.012	0.026	0.054	0.535
971	0.466	0.015	0.003	0.050	0.000	0.000	0.003	0.006	0.016	0.049	0.125
188	0.445	0.023	0.004	0.100	0.000	0.003	0.006	0.012	0.023	0.050	0.514
480	0.425	0.039	0.051	0.082	0.000	0.000	0.006	0.026	0.056	0.085	0.248
481	0.422	0.014	0.004	0.053	0.000	0.001	0.004	0.008	0.014	0.036	0.420
1054	0.357	0.023	0.000	0.070	0.000	0.001	0.005	0.013	0.029	0.055	0.449
210	0.300	0.011	0.002	0.059	0.000	0.001	0.003	0.005	0.010	0.025	0.347

6.3.3. Feature Testing

Figure 51 displays the feature testing scatter matrix of the extracted features.



a) Scatter Matrix of extracted features for UserID

b) Scatter Matrix of extracted features for DeviceID

c) Scatter Matrix of extracted features for PatientID

d) Scatter Matrix of extracted features for Routine

Figure 51 - Scatter Matrix of extracted features for UserID

6.3.4. LOF: User, Patient and Device ID

A LOF score is calculated for the 45 combinations (of the 10 features) for each of the 44,640 unique IDs in the dataset. Therefore 2,008,800 unique LOF scores are calculated in total. An average is then taken to assign an anomaly score. In Table 28, LOF identifies anomalous User IDs, Patient IDs and Device IDs. The neighbourhood radius is defined in stage 3 of the LOF algorithm (Section 3.2.3.1), the density score is defined in stage 4, and the anomaly score is the final LOF value, as defined in stage 5.

Table 28 - LOF (Mean) Anomaly Scores for User ID, Patient ID and Device ID

User ID	Density Score	Anomaly Score	Neigbour hood Radius	Patient ID	Density Score	Anomaly Score	Neigbour hood Radius	Device ID	Density Score	Anomaly Score	Neigbour hood Radius
778	1.51	16.95	0.97	**17999**	627.47	9.19	0.00	**1661**	179.68	2.78	0.01
196	6.40	6.18	0.23	**17103**	278.22	9.12	0.01	**869**	18.83	2.66	0.27
325	18.39	4.33	0.08	**2756**	1282.14	9.11	0.00	**1643**	5.23	2.24	0.31
469	24.30	2.42	0.05	**9349**	107.54	8.88	0.02	**1348**	337.71	2.13	0.00
1108	36.96	2.26	0.03	**22110**	524.28	8.16	0.01	**704**	398.25	2.00	0.00
374	44.83	1.99	0.03	**25233**	1402.74	8.00	0.00	**1139**	417.52	1.95	0.00
674	232.47	1.97	0.01	**32437**	1980.12	7.62	0.00	**205**	3.70	1.89	0.50
1037	29.86	1.96	0.07	**13656**	811.91	6.99	0.00	**361**	28.75	1.87	0.04
37	33.17	1.95	0.05	**19685**	2086.13	6.83	0.00	**881**	38.93	1.76	0.05
553	43.59	1.87	0.03	**15853**	626.51	6.70	0.00	**391**	21.22	1.72	0.08

Within the User ID range, the most notable ID is #778, with an anomaly score of 16.95. There are 5 User IDs with an anomaly score above 2. Therefore, LOF has indicated that 0.45% of the User IDs are anomalous. Similarly, the most notable Patient ID is #17999, with an anomaly score of 9.19. There are 108 Patient IDs with an anomaly score above 2; indicating 0.31% of the Patient IDs are anomalous. Finally, the most notable Device ID is #1661, with an anomaly score of 2.78. There are 5 Device IDs with an anomaly score above 2, indicating that 0.3% of the Device IDs are irregular. Overall therefore, LOF identifies 0.35% of IDs as anomalous, which would be highlighted to a patient privacy officer for investigation.

6.3.5. LOF: Routine ID

However, the LOF technique cannot be applied as effectively to the Routine ID. Table 23 presents a sample of the highest LOF anomaly scores for the Routine ID dataset.

Table 29 - LOF (Mean) Anomaly Scores for Routine ID

Routine Set Description	Density Score	Anomaly Score	Neighbourhood Radius
BD VH ZCUS.UK.SCH OE	1030.12	12.28	0.00
SPC SS ASF PHA.MEDS PHA.ORDS	920.94	10.88	0.00
SS MED MPI	881.46	10.41	0.00
RAD.DRP LAB.DRP	629.85	9.65	0.01
ASF CM MED PHA.ORDS	268.97	9.53	0.01
MPI ASF RAD.DRP	377.45	9.42	0.01
NPC NOTE MED	1407.53	8.95	0.00
MPI ZCUS.UK.SCH ZCUS.UK.LETTER SS	416.31	8.87	0.01
PHA.ORDS OE ZCUS.UK.SCH	865.28	8.66	0.00
ZCUS.UK.LETTER SS ZCUS.UK.SCH MPI	74.72	8.23	0.03

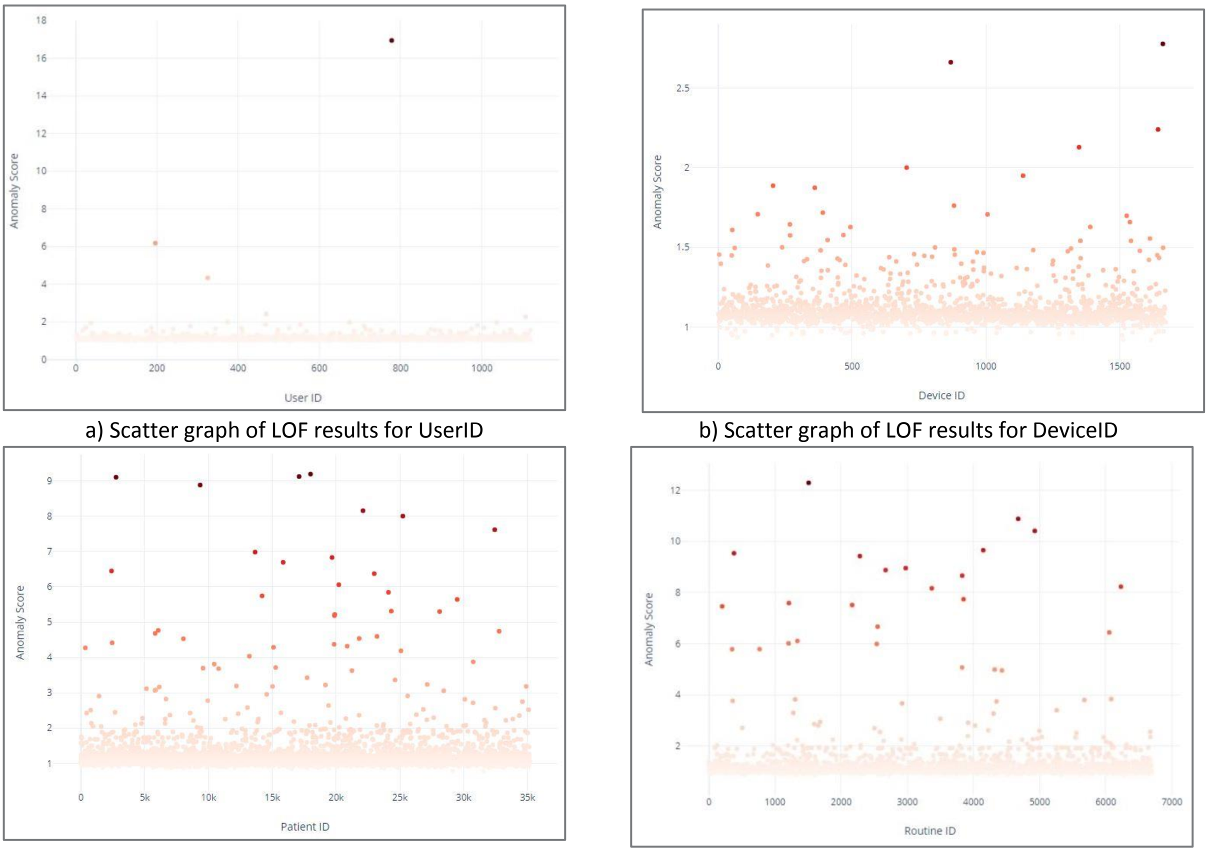
There are 72 routine sets with an anomaly score above 2. Therefore, LOF has indicated that 1.08% of the routine sets are anomalous. The most notable routine set is the combination ‘Bulletin Board (Alerts) | Visit History | Cancelled Account.UK.Scheduling | View Orders’, with an anomaly score of 12.28.

6.3.6. Quantifying LOF

The data is cleaned as per the process outlined in chapter 4. NAN and Inf values are replaced with 1 and 2 respectively, whilst missing or null values are assigned the median value for their feature class.

6.3.7. Visualisation of LOF Results

A visualisation of the LOF results for each ID is presented in Figure 57.



a) Scatter graph of LOF results for UserID

b) Scatter graph of LOF results for DeviceID

c) Scatter graph of LOF results for PatientID

d) Scatter graph of LOF results for Routine

Figure 52 – Scatter graph of LOF results

6.3.8. Anomaly Score Ensemble Averaging

A sample of EPR data with a calculated ensemble average LOF anomaly score (including Routine ID) is presented in Table 30. The table is ordered by the highest LOF anomaly scores. Within the date range, the most notable audit log occurred on 16th July 2015 at 22:10. User #778 accessed Patient #34072 on Device #967 performing the following Routine combination ‘Assessment Forms’, with an anomaly score of 5.186. There are 115 audit logs with an anomaly score above 2. Therefore, PARISS has indicated that 0.034% of the EPR Audit Logs are anomalous.

Table 30 - EPR Audit Data with Ensemble Averaging Applied to LOF Anomaly Score (including Routine)

Date & Time	Device	Device Anomaly Score	User ID	User Anomaly Score	Routine ID	Routine Anomaly Score	Patient ID	Patient Anomaly Score	Duration (sec)	Adm Date	Dis Date	Ensemble Averaging Anomaly Score (inc Routine)
15/07/16 22:10	967	1.470	778	16.946	ASF	1.091	34072	1.037	5703	27/06/2007	27/06/2007	**5.136**
27/07/16 21:40	113	1.079	480	1.137	BD VH ZCUS.UK.SCHOE	12.284	8275	1.012	436	06/04/2016	06/04/2016	**3.878**
14/11/16 01:09	850	1.093	26	1.097	BD VH ZCUS.UK.SCHOE	12.284	352	1.018	481	11/01/2010	11/01/2010	**3.873**
07/12/16 20:17	727	1.107	532	1.130	SPC SS ASF PHA.MEDS PHA.ORDS	10.876	18453	1.027	706	02/04/2012	02/04/2012	**3.535**
21/11/16 13:55	727	1.107	532	1.130	SPC SS ASF PHA.MEDS PHA.ORDS	10.876	24489	1.025	757	07/07/2008	07/07/2008	**3.535**
07/08/16 07:07	1471	1.078	346	1.087	SS MED MPI	10.405	14112	1.067	692	20/07/2016	20/07/2016	**3.409**
26/09/16 03:12	108	1.069	346	1.087	SS MED MPI	10.405	19699	1.035	634	19/01/2016	19/01/2016	**3.399**
09/09/16 17:35	378	1.217	944	1.389	ASF CM MED PHA.ORDS	9.533	33536	1.059	1030	14/07/2016	14/07/2016	**3.300**
13/10/16 01:48	591	1.046	142	1.031	RAD.DRP LAB.DRP	9.651	2396	1.128	672	23/06/2015	01/07/2015	**3.214**
15/11/16 03:35	1213	1.042	552	1.055	RAD.DRP LAB.DRP	9.651	4287	1.082	539	31/08/2016	17/09/2016	**3.208**

A sample of EPR data with a calculated ensemble average LOF anomaly score (excluding Routine ID) is presented in Table 31. The table is ordered by the highest LOF anomaly scores. Within the date range, the most notable audit log occurred on 16th July 2015 at 22:10. User #778 accessed Patient #34072 on Device #967 performing the following Routine combination ‘Assessment Forms’, with an anomaly score of 9.726. There are 1,250 audit logs with an anomaly score above 2. Therefore, PARISS has indicated that 0.367% of the EPR Audit Logs are anomalous.

Table 31 - EPR Audit Data with Ensemble Averaging Applied to LOF Anomaly Score (excluding Routine)

Date & Time	Device	Device Anomaly Score	User ID	User Anomaly Score	Routine ID	Routine Anomaly Score	Patient ID	Patient Anomaly Score	Duration (sec)	Adm Date	Dis Date	Ensemble Averaging Anomaly Score (exc Routine)
15/07/16 22:10	967	1.470	778	16.946	ASF	1.091	34072	1.037	5703	27/06/2007	27/06/2007	**9.726**
29/12/16 17:27	1496	1.177	581	1.108	ZCUS.UK.SCH	1.147	17999	9.195	847	11/09/2015	11/09/2015	**5.740**
11/10/16 20:00	188	1.101	947	1.069	ZCUS.UK.LETTER	1.148	17999	9.195	944	11/09/2015	11/09/2015	**5.683**
11/10/16 01:17	1532	1.152	139	1.050	ZCUS.UK.SCH	1.147	17103	9.125	1823	N/A	N/A	**5.663**
17/11/16 19:05	1495	1.083	1010	1.065	ZCUS.UK.LETTER	1.148	2756	9.106	828	15/05/2006	15/05/2006	**5.626**
08/11/16 22:37	448	1.267	198	1.099	ZCUS.UK.LETTER	1.148	9349	8.884	2352	06/04/1994	06/04/1994	**5.625**
06/10/16 09:49	47	1.083	339	1.057	NOTE	1.150	2756	9.106	883	15/05/2006	15/05/2006	**5.623**
28/09/16 01:24	264	1.038	1066	1.050	ZCUS.UK.LETTER	1.148	17103	9.125	1472	N/A	N/A	**5.606**
29/09/16 18:44	836	1.102	198	1.099	ZCUS.UK.LETTER	1.148	9349	8.884	2845	06/04/1994	06/04/1994	**5.542**
04/10/16 18:32	418	1.113	465	1.179	ZCUS.UK.LETTER	1.148	22110	8.158	1332	19/02/2009	19/02/2009	**5.225**

In comparison, the results in Table 30 and Table 31 are dissimilar. The most anomalous audit log in Table 30 (15th July 2016 at 22:10) is also the most anomalous audit log in Table 31, although the ensemble averaged anomaly score ranges significantly from 5.136 to 9.726 respectively. This is because the anomaly score of User 778 is almost 17 which indicates a very anomalous user ID and therefore is the most anomalous behaviour for both results. This user only accesses the EPR once in the six months of data available, but spends 5,703 seconds (over 1.5 hours) accessing the Assessment Forms of a single patient. With the exception of this significantly anomalous result, the remainder of both tables vary significantly. This is due to both the Patient ID and Routine IDs having statistically high anomaly scores within the data (with anomaly scores over 5). Therefore in Table 30 the audit logs that include anomalous routines are prioritised for the attention of the analyst, whereas in Table 31 the audit logs that include anomalous patient IDs are displayed instead. Ultimately, Table 31 is a more useful indicator of anomalous audit logs as discussed with the hospital.

6.3.9. Visualisation of Audit Log Results

In Figure 53, a visualisation of LOF results of calculated ensemble averaged anomaly scores is displayed for all 340,687 audit logs. The x-axis displays the date, and the y-axis displays the calculated ensemble average anomaly score.

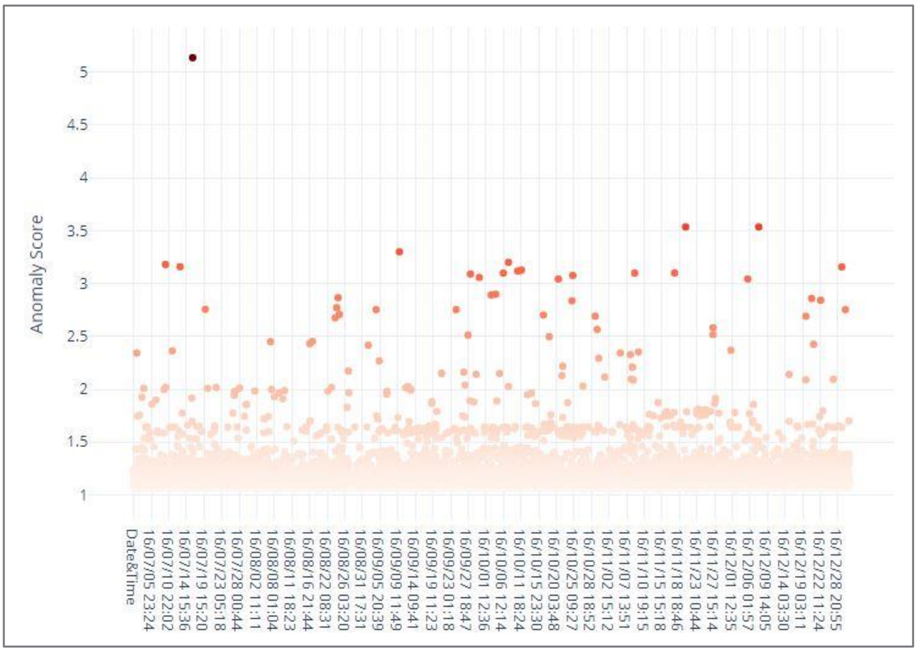


Figure 53 - Visualisation of LOF Results for Ensemble Averaged Anomaly Scores (inc Routine)

In Figure 54, a visualisation of LOF results of calculated ensemble averaged anomaly scores.

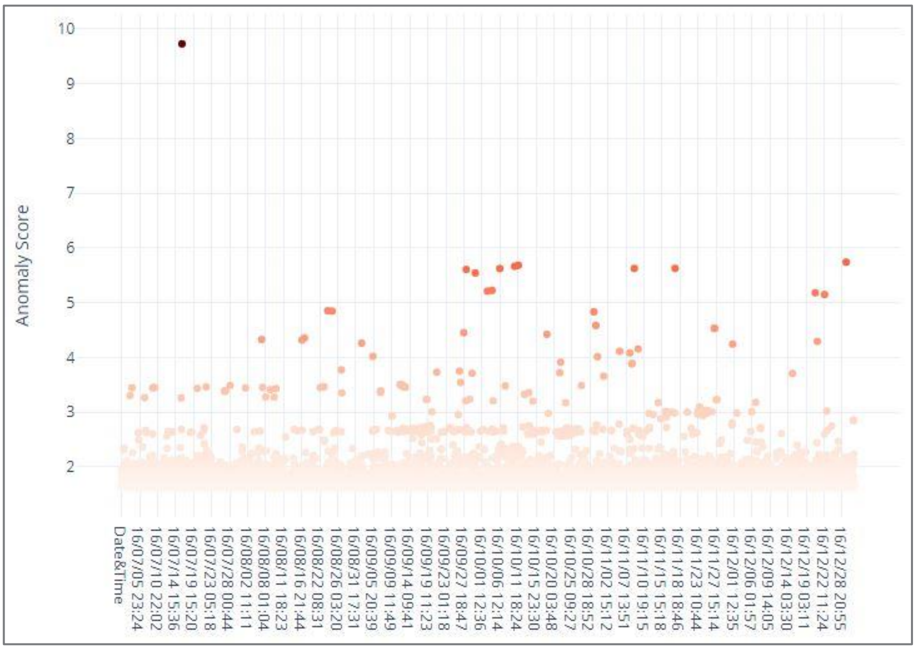


Figure 54 - Visualisation of LOF Results for Ensemble Averaged Anomaly Scores (exc Routine)

6.4. 18 Months EPR Data

The final dataset consists of eighteen months of data (28-02-16 – 21-08-17). The data contains 1,007,727 rows of data. There are 1,515 User IDs, 72,878 Patient IDs, 2,270 Devices IDs and 13,722 Routine IDs.

6.4.1. Feature Extraction

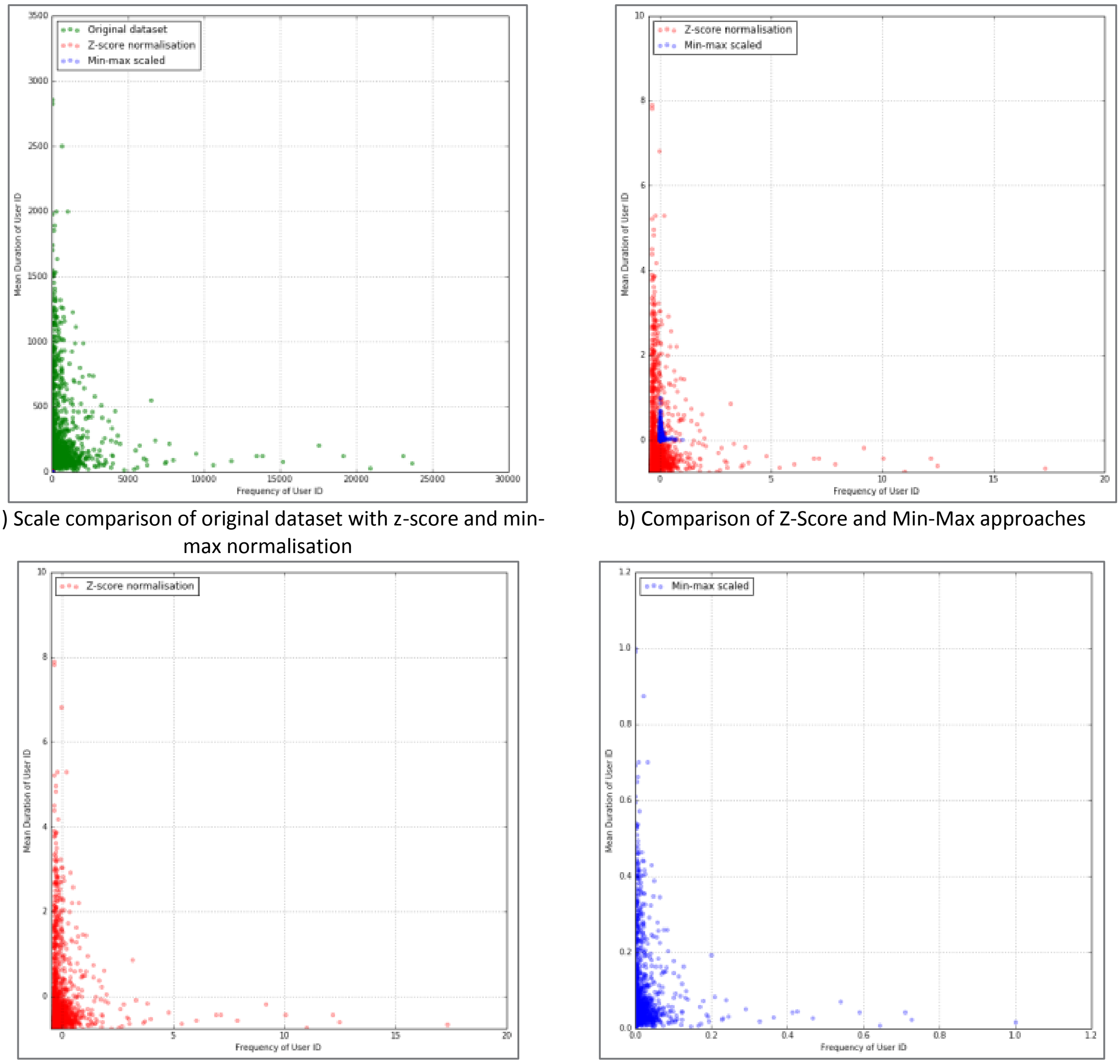
Table 32 displays the eleven features extracted for the ten most frequent user IDs, determined using a pivot table. Each of the 90,385 unique IDs in the dataset has these eleven features extracted, resulting in 994,235 features in total.

Table 32 - Feature Extraction for User ID (Eighteen Months)

User ID	Frequency	Mean	Mode	STD	Min	5th Percentile	25th Percentile	Median Duration	75th Percentile	95th Percentile	Max
1016	32557	49.5	2	106.66	0	1	5	17	43	219	4751
1320	23674	69.57	2	117.32	0	3	14	26	72	289	2268
1025	23104	124.06	2	246.66	0	4	24	67	147	383	7469
742	20907	27.39	5	79.29	0	2	5	9	22	108	6081
248	19160	125.23	21	264.26	0	17	37	69	134	371	8360
639	17543	205.39	2	242.25	0	2	21	120	307	644	5876
640	15159	80.15	16	136.54	0	8	23	46	83	272.1	5855
1424	13824	125.85	1	177.73	0	4	31	77	157	407	6268
372	13450	124.04	2	263.45	0	1	2	5	127.75	635	6572
108	11797	83.7	2	139.66	0	2	11	34	107	314	3521

6.4.2. Feature Scaling

As detailed in Chapter 4, a Min-Max approach is taken for PARISS. Here a Z-Score approach is also detailed and compared to Min-Max. Figure 55(a) displays the data points on three scales, the original dataset (green), the Z-Score standardised features (red) and the min-max normalised features (blue). Figure 55(b) displays a comparison of Z-Score and Min-Max approaches, without the original dataset. Figure 55**Error! Reference source not found.**(c) displays the Z-Score standardised features independently, whereas Figure 55(d) presents the Min-Max approach.



a) Scale comparison of original dataset with z-score and min-max normalisation

b) Comparison of Z-Score and Min-Max approaches

c) Z-score normalisation

d) Min-Max scaling

Figure 55 – (a) Scale comparison of original dataset with z-score and min-max normalisation (b) Comparison of Z-Score and Min-Max approaches

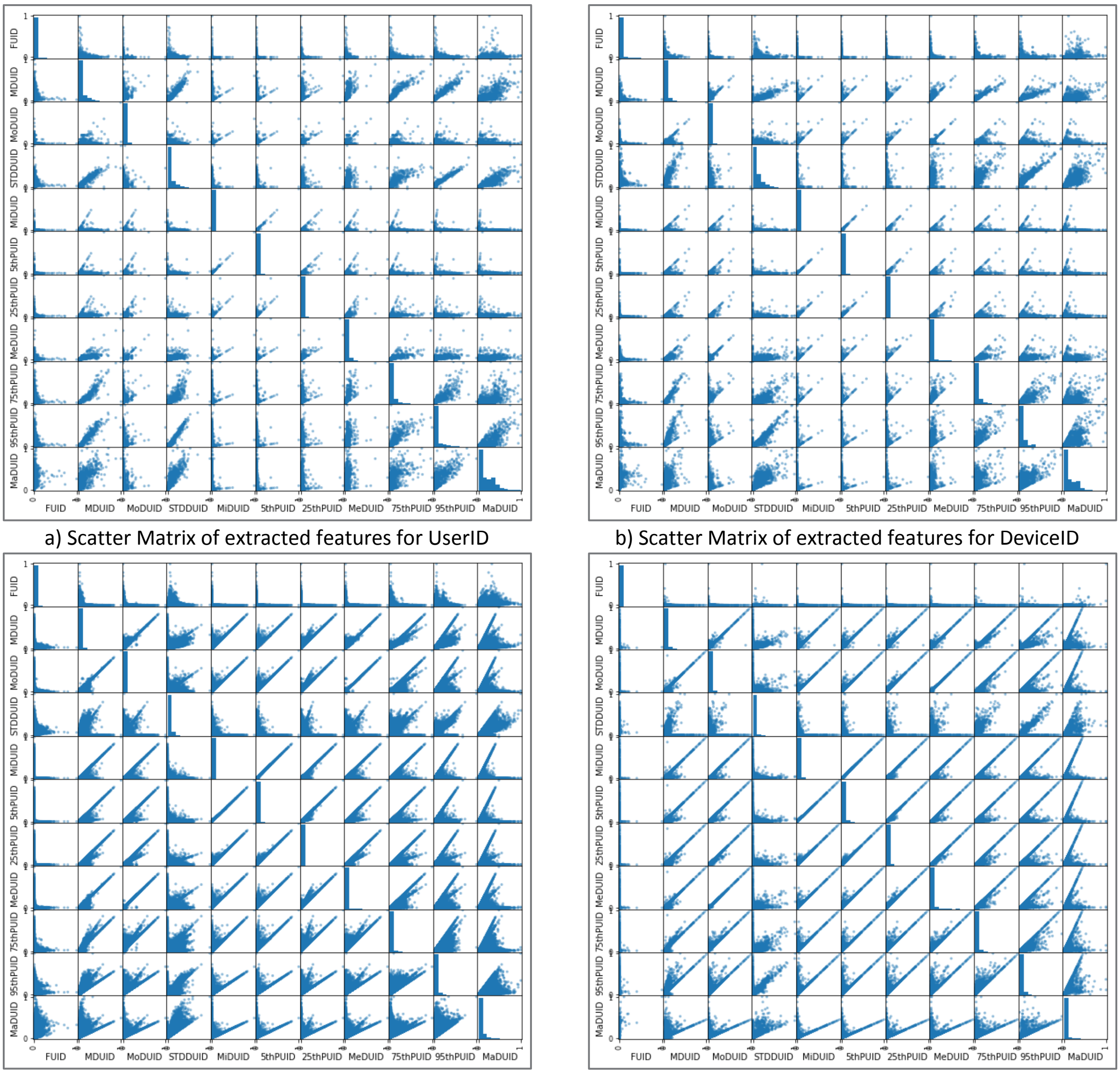
Table 33 displays the top ten most frequent User IDs, with min-max scaling applied to their features.

Table 33 – Feature Scaling for User ID

User ID	Frequency	Mean	Mode	STD	Min	5th Percentile	25th Percentile	Median Duration	75th Percentile	95th Percentile	Max
1016	1	0.017	0.001	0.019	0	0.001	0.002	0.001	0.010	0.020	0.257
1320	0.727	0.024	0.001	0.021	0	0.001	0.008	0.009	0.017	0.026	0.123
1025	0.710	0.043	0.001	0.045	0	0.002	0.014	0.023	0.036	0.034	0.404
742	0.642	0.009	0.001	0.014	0	0.001	0.002	0.003	0.005	0.010	0.329
248	0.588	0.043	0.007	0.048	0	0.010	0.021	0.024	0.033	0.033	0.453
639	0.539	0.072	0.001	0.044	0	0.001	0.012	0.042	0.075	0.058	0.318
640	0.466	0.028	0.005	0.025	0	0.004	0.013	0.016	0.020	0.024	0.317
1424	0.425	0.044	0	0.032	0	0.002	0.018	0.027	0.038	0.036	0.339
372	0.413	0.043	0.001	0.048	0	0.001	0.001	0.001	0.031	0.057	0.356
108	0.362	0.029	0.001	0.025	0	0.001	0.006	0.012	0.0260	0.028	0.191

6.4.3. Feature Testing

The scatter matrix, displayed in Figure 56 (all features have been abbreviated in the graph labels) visualises the relationship between the features to predict the most appropriate for the LOF classification. The scatter matrix displays the positive and negative correlation between the features. In this case, from the visual inspection, the majority of features have a positive correlation. However, based on Figure 56, the consideration would be to remove the feature Frequency for each Unique Identifier (FUID) for the UserID, Routine and Device Interaction classification but retain it for PatientID. Referring to the Routine and Device Interaction, the data collected relates predominately to unique routine combinations, so logically the FUID feature is less significant, as confirmed by the scatter matrix.



a) Scatter Matrix of extracted features for UserID

b) Scatter Matrix of extracted features for DeviceID

c) Scatter Matrix of extracted features for PatientID

d) Scatter Matrix of extracted features for Routine

Figure 56 - Scatter Matrix of extracted features

6.4.4. LOF: User, Patient and Device ID

A LOF score is calculated for the 45 combinations (of the 10 features) for each of the 90,385 unique IDs in the dataset. Therefore 4,067,325 unique LOF scores are calculated in total. An average is then taken to assign an anomaly score. In Table 34, LOF identifies anomalous User IDs, Patient IDs and Device IDs. The neighbourhood radius is defined in stage 3 of the LOF algorithm (Section 3.2.3.1), the density score is defined in stage 4, and the anomaly score is the final LOF value, as defined in stage 5.

Table 34 - LOF (Mean) Anomaly Scores for User ID, Patient ID and Device ID

User ID	Density Score	Anomaly Score	Neigbourhood Radius	Patient ID	Density Score	Anomaly Score	Neigbourhood Radius	Device ID	Density Score	Anomaly Score	Neigbourhood Radius
685	3.03	4.36	0.546	**35888**	371.74	9.41	0.006	**2258**	374.85	4.86	0.003
260	5.73	3.54	0.518	**19327**	175.92	8.81	0.018	**1082**	2.26	4.75	0.730
1037	69.14	2.80	0.251	**58816**	794.70	8.58	0.003	**1557**	168.84	2.92	0.009
1002	46.81	2.61	0.051	**69053**	51.59	8.55	0.053	**729**	5.26	2.80	0.303
1401	16.55	2.56	0.153	**51280**	765.21	7.61	0.003	**499**	29.58	2.52	0.048
707	19.05	2.28	0.207	**41306**	150.01	7.53	0.014	**527**	6.84	2.43	0.206
1311	83.73	2.23	0.016	**46695**	647.07	7.32	0.008	**896**	10.35	2.32	0.170
242	77.78	2.13	0.024	**13704**	1315.64	6.99	0.002	**2014**	6.75	2.29	0.216
1493	47.75	2.03	0.134	**34419**	23.12	6.97	0.101	**1104**	107.50	2.28	0.033
507	28.66	2.00	0.103	**56428**	2570.47	6.94	0.003	**523**	17.38	2.25	0.077

Within the User ID range, the most notable ID is #685, with an anomaly score of 4.36. There are 10 User IDs with an anomaly score above 2. Therefore, LOF has indicated that 0.66% of the User IDs are anomalous. Similarly, the most notable Patient ID is #35888, with an anomaly score of 9.41. There are 122 Patient IDs with an anomaly score above 2; indicating 0.17% of the Patient IDs are anomalous. Finally, the most notable Device ID is #2258, with an anomaly score of 4.86. There are 12 Device IDs with an anomaly score above 2, indicating that 0.53% of the Device IDs are irregular. Overall therefore, LOF identifies 0.45% of IDs as anomalous, which would be highlighted to a patient privacy officer for investigation.

Examples of audit log data classified as inlier, outlier and abnormal data for User ID are presented in Table 35. Audit log data classified as an inlier within the dense region (<1) is User ID 571, with a LOF score of 0.95. Audit log data classified as an outlier within the normal region (>1 and <2) is User ID 1486, with a LOF score of 1.12. Audit log data classified as an outlier within the abnormal region (>2) is User ID 707, with a LOF score of 2.28.

Table 35 - EPR Audit Log Data Examples for Inlier, Outlier and Abnormal Data Points

Date & Time	Device ID	User ID	Routine Description	Patient ID	Duration (sec)	Adm Date	Dis Date
08/03/17 01:32	2046	571	Visit History	33727	28	08/03/2017	08/03/2017
07/08/17 15:37	396	1485	Current Medication Orders | Pharmacy Orders	62584	58	16/10/2001	16/10/2001
30/05/16 11:09	936	707	Visit History | Radiology Reports | Maternity Data | Cancelled Account.UK.Letter | Cancelled Account.UK.Scheduling UK.View Orders	28160	385	26/01/2016	26/01/2016

The results presented here demonstrate a technique for uncovering anomalous or irregular behavioural patterns from a complex dataset that would otherwise not be possible from either a visual inspection/visualisation of the whole dataset.

6.4.5. LOF: Routine ID

However, the LOF technique cannot be applied as effectively to the Routine ID. Table 36 presents a sample of the highest LOF anomaly scores for the Routine ID dataset.

Table 36 - LOF (Mean) Anomaly Scores for Routine ID

Routine Set Description	Density Score	Anomaly Score	Neigbourhood Radius
Assessment Forms | Maternity Data | Care-Area Administrative Data | Admissions Demographic Data	1043.094	13.34	0.003
*** | UK.View Orders | Admissions Demographic Data | Pharmacy Orders	1649.703	11.64	0.005
*** | Cancelled Account.UK.Letter | Admissions Demographic Data	2213.821	11.41	0.004
Maternity Data | Theatre Management | Assessment Forms | Visit History	581.246	11.35	0.004
Theatre Management | Cancelled Account.UK.Letter | Cancelled Account.UK.Scheduling | Admissions Demographic Data	632.774	9.70	0.005
Recent Clinical Results | Recent Clinical Results:(Departmental Reports) | Pharmacy.Medication Order History | UK.View Orders	70.561	9.54	0.035
Assessment Forms | Admissions Demographic Data | Visit History | Alerts	601.429	9.29	0.004
Cancelled Account.UK.Letter | Pharmacy Orders | Admissions Demographic Data	470.423	8.81	0.005
Assessment Forms | Cancelled Account.UK.Letter | Cancelled Account.UK.Scheduling | Medication Order History	646.410	8.32	0.006
Internet Access | Alerts | Assessment Forms	693.934	8.22	0.005

The EPR audit logs calculate a string of routines performed on the same patient as a unique Routine ID. The differing routines are delimited with a pipe (|). Therefore, there are 13,722 Routine IDs in the dataset, whereas there are more accurately approximately 100 unique routines a user could perform.

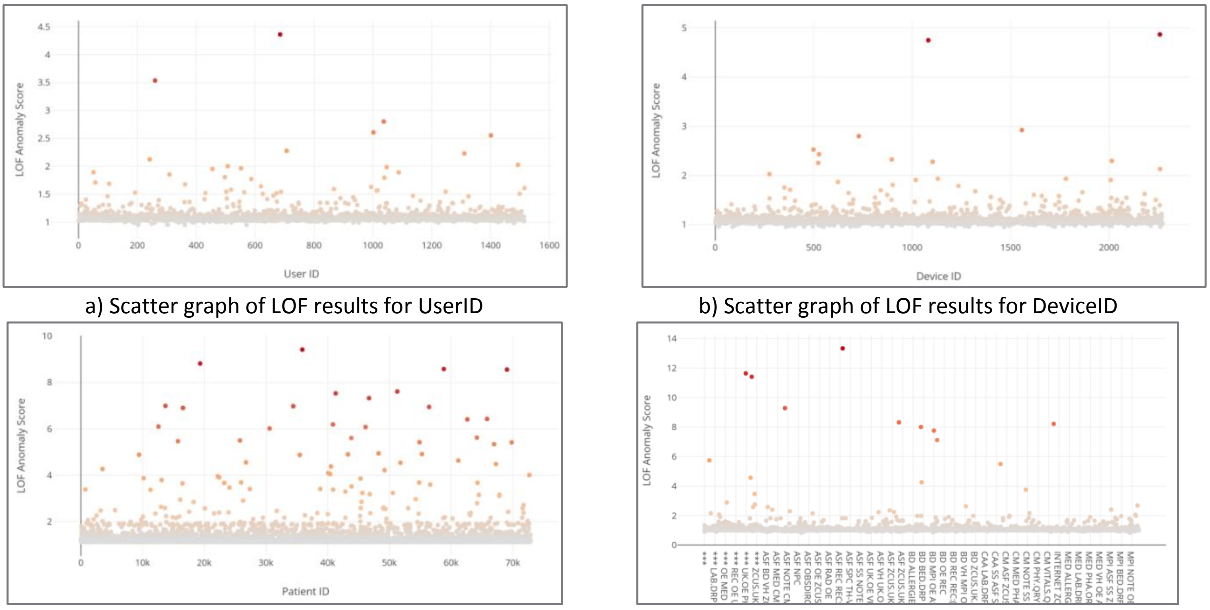
There are 102 routine sets with an anomaly score above 2. Therefore, LOF has indicated that 0.74% of the routine sets are anomalous. The most notable routine set is the combination 'Assessment Forms | Maternity Data | Care-Area Administrative Data | Admissions Demographic Data', with an anomaly score of 13.34. This specific routine combination only occurs twice in the audit logs of over 1,000,000 rows.

6.4.6. Quantifying LOF

The data is cleaned as per the process outlined in chapter 4. NAN and Inf values are replaced with 1 and 2 respectively, whilst missing or null values are assigned the median value for their feature class.

6.4.7. Visualisation of LOF Results

A visualisation of the LOF results for each ID is presented in Figure 57.



a) Scatter graph of LOF results for UserID

b) Scatter graph of LOF results for DeviceID

c) Scatter graph of LOF results for PatientID

d) Scatter graph of LOF results for Routine

Figure 57 – Scatter graph of LOF results

6.4.8. Anomaly Score Ensemble Averaging

A sample of EPR data with a calculated ensemble average LOF anomaly score (including Routine ID) is presented in Table 37. The table is ordered by the highest LOF anomaly scores. Within the date range, the most notable audit log occurred on 26th Sep 2016 at 17:02. User #435 accessed Patient #71272 on Device #1284 performing the following Routine combination ‘Assessment Forms Maternity Data Care-Area Administrative Data Admissions Demographic Data’, with an anomaly score of 4.139. There are 145 audit logs with an anomaly score above 2. Therefore, PARISS has indicated that 0.014% of the EPR Audit Logs are anomalous.

Table 37 - EPR Audit Data with Ensemble Averaging Applied to LOF Anomaly Score (including Routine)

Date & Time	Device	Device Anomaly Score	User ID	User Anomaly Score	Routine ID	Routine Anomaly Score	Patient ID	Patient Anomaly Score	Duration (sec)	Adm Date	Dis Date	Ensemble Averaging Anomaly Score (inc Routine)
26/09/16 17:02	1284	1.050195	435	1.086618	ASF SPC CAA MPI	13.33895	71272	1.080687	853	15/03 /2016	15/03 /2016	**4.139**
25/11/16 03:39	102	1.084235	1487	1.043842	ASF SPC CAA MPI	13.33895	29971	1.047444	901	02/09 /2015	02/09 /2015	**4.129**
15/08/16 20:56	531	1.161338	358	1.051894	*** UK.OE MPI PHA.ORDS	11.64312	23637	1.066424	1180	08/01 /2016	08/01 /2016	**3.730**
21/11/16 21:46	369	1.087683	1021	1.125426	SPC SS ASF VH	11.34951	41661	1.089557	970	24/01 /1997	24/01 /1997	**3.663**
09/08/17 11:39	1537	1.122767	77	1.048351	SPC SS ASF VH	11.34951	57108	1.029903	1041	30/12 /2015	30/12 /2015	**3.638**
21/11/16 17:38	1052	1.094431	809	1.08748	SS ZCUS.UK.LETTER ZCUS.UK.SCH MPI	9.700566	43065	1.053911	723	29/12 /1997	29/12 /1997	**3.234**
01/04/16 01:12	49	1.151319	117	1.03098	SS ZCUS.UK.LETTER ZCUS.UK.SCH MPI	9.700566	52200	1.028481	861	29/09 /2015	29/09 /2015	**3.228**
19/12/16 20:03	293	1.066586	992	1.090164	REC REC:(DRP) PHA.MEDS UK.OE	9.538052	41375	1.054439	2454	28/11 /2016	28/11 /2016	**3.187**
07/02/17 00:18	566	1.164282	262	1.073656	ZCUS.UK.LETTER	1.084454	35888	9.413876	1182	18/04 /2013	18/04 /2013	**3.184**
27/12/16 18:50	293	1.066586	992	1.090164	REC REC:(DRP) PHA.MEDS UK.OE	9.538052	46862	1.01959	1691	07/12 /2016	07/12 /2016	**3.179**

A sample of EPR data with a calculated ensemble average LOF anomaly score (excluding Routine ID) is presented in Table 38. The table is ordered by the highest LOF anomaly scores. Within the date range, the most notable audit log occurred on 7th Feb 2017 at 00:18. User #262 accessed Patient #35888 on Device #566 performing the following Routine combination ‘ZCUS.UK.Letter’, with an anomaly score of 3.884. There are 156 audit logs with an anomaly score above 2. Therefore, PARISS has indicated that 0.015% of the EPR Audit Logs are anomalous.

Table 38 - EPR Audit Data with Ensemble Averaging Applied to LOF Anomaly Score (excluding Routine)

Date & Time	Device	Device Anomaly Score	User ID	User Anomaly Score	Routine ID	Routine Anomaly Score	Patient ID	Patient Anomaly Score	Duration (sec)	Adm Date	Dis Date	Ensemble Averaging Anomaly Score (exc Routine)
07/02/01 00:18	566	1.164	262	1.074	ZCUS.UK.LETTER	1.084	35888	9.414	1182	18/04/2013	18/04/2013	**3.884**
26/07/16 23:36	594	1.141	262	1.074	ZCUS.UK.LETTER	1.084	35888	9.414	1342	18/04/2013	18/04/2013	**3.876**
08/11/16 22:37	566	1.164	262	1.074	ZCUS.UK.LETTER	1.084	19327	8.814	2352	06/04/1994	06/04/1994	**3.684**
29/09/16 18:44	1050	1.153	262	1.074	ZCUS.UK.LETTER	1.084	19327	8.814	2845	06/04/1994	06/04/1994	**3.680**
27/07/16 15:46	286	1.071	809	1.087	ZCUS.UK.LETTER	1.084	69053	8.552	4251	11/02/2003	11/02/2003	**3.570**
20/06/17 18:43	1970	1.113	117	1.031	ZCUS.UK.LETTER	1.084	69053	8.552	2680	11/02/2003	11/02/2003	**3.565**
18/01/17 01:57	347	1.041	1439	1.070	ZCUS.UK.LETTER	1.084	58816	8.577	844	N/A	N/A	**3.563**
01/03/16 16:46	497	1.055	181	1.028	ZCUS.UK.LETTER	1.084	58816	8.577	924	N/A	N/A	**3.553**
14/09/16 02:52	890	1.070	740	1.080	ZCUS.UK.LETTER	1.084	51280	7.608	884	31/12/2014	31/12/2014	**3.253**
11/01/17 02:40	1967	1.085	1241	1.036	ZCUS.UK.LETTER	1.084	51280	7.608	814	31/12/2014	31/12/2014	**3.243**

In comparison, the results in Table 37 and Table 38 are dissimilar. The only audit log that appears in both is the most anomalous audit log in Table 38 (7th Feb 2017 00:18) which is the ninth most anomalous audit log in Table 37. This is due to both the Patient ID and Routine IDs having statistically high anomaly scores within the data (with anomaly scores over 5). Therefore in Table 37 the audit logs that include anomalous routines are prioritised for the attention of the analyst, whereas in Table 38 the audit logs that include anomalous patient IDs are displayed instead. The audit log on the 7th Feb 2017 has a high enough patient anomaly score to appear in both tables. Ultimately, Table 38 is a more useful indicator of anomalous audit logs.

6.4.9. Visualisation of Audit Log Results

In Figure 58, a visualisation of LOF results of calculated ensemble averaged anomaly scores (including the Routine ID anomaly score) is displayed for all 1,007,727 audit logs. The x-axis displays the date, and the y-axis displays the calculated ensemble average anomaly score.

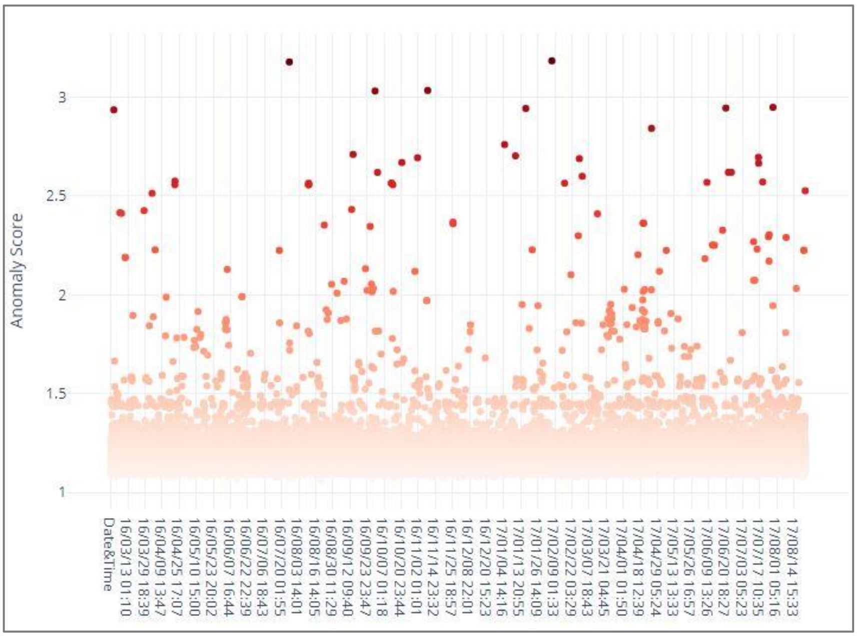


Figure 58 - Visualisation of LOF Results for Ensemble Averaged Anomaly Scores (inc Routine)

In Figure 59, a visualisation of LOF results of calculated ensemble averaged anomaly scores (excluding the Routine ID anomaly score) is shown.

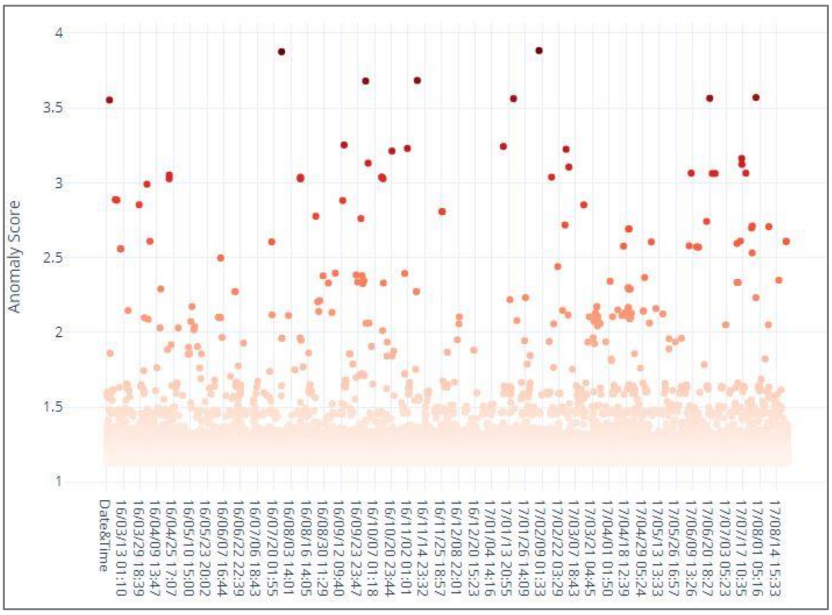
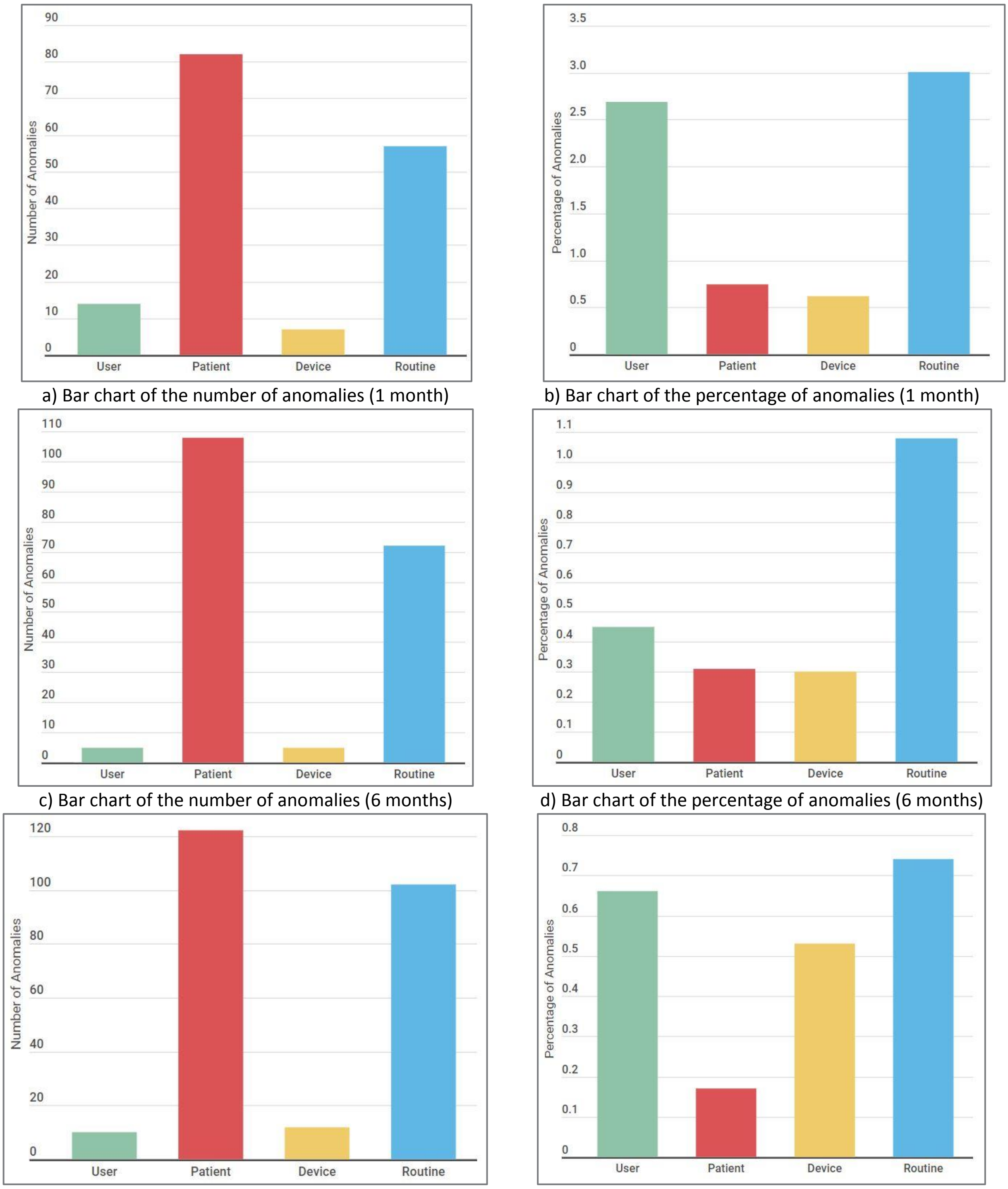


Figure 59 - Visualisation of LOF Results for Ensemble Averaged Anomaly Scores (exc Routine)

6.5. Discussion

A discussion of the three datasets is presented in this section. A bar chart comparing the number and percentage of anomalies for each of the ID types in the one-month dataset is presented in Figure 60(a) and Figure 60(b). A bar chart comparing the number and percentage of anomalies for each of the ID types in the six-month dataset is presented in Figure 60(c) and Figure 60(d). A bar chart comparing the number and percentage of anomalies for each of the ID types in the eighteen-month dataset is presented in Figure 60(e) and Figure 60(f).



a) Bar chart of the number of anomalies (1 month)

b) Bar chart of the percentage of anomalies (1 month)

c) Bar chart of the number of anomalies (6 months)

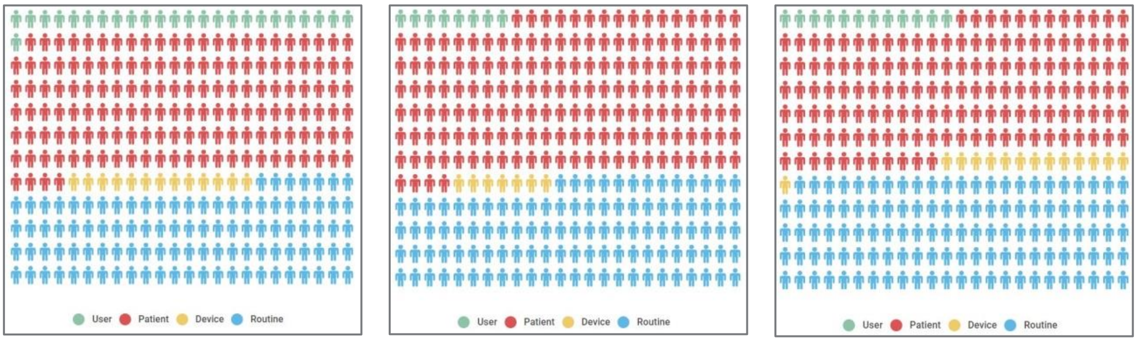
d) Bar chart of the percentage of anomalies (6 months)

e) Bar chart of the number of anomalies (18 months)

f) Bar chart of the percentage of anomalies (18 months)

Figure 60 – Bar chart of the number and percentage of anomalies for the 1 month, 6 month and 18 month datasets

Pictorial showing the anomalies in the data is presented in Figure 61. User ID is green, Patient ID is red, Device ID is yellow and Routine ID is blue.



(a) Number of anomalies in the dataset for 1 month

(b) Number of anomalies in the dataset for 6 months

(c) Number of anomalies in the dataset for 18 months

Figure 61 – Pictorial of the number of anomalies in the dataset for 1 month, 6 months, and 18 months

To compare if any audit logs appear as anomalous, the top 100 anomalous audit logs are included with their anomaly score (including routine ID). In Table 39 the matching records are presented, along with their overall rank and ensemble average within the results for the respective dataset. The full top 100 is in the appendix.

Table 39 – Matching records in the Top 100 Anomalous Audit Logs (inc Routine ID) for 6 month and 18 months

Date & Time (Jul16-Dec16)	Rank	Ensemble Average	Date & Time (Feb16-Aug17)	Rank	Ensemble Average
16/12/07 20:17	2	3.535	16/12/07 20:17	85	2.327
16/11/21 13:55	3	3.535	16/11/21 13:55	83	2.331
16/11/08 22:37	12	3.099	16/11/08 22:37	14	3.034
16/09/29 18:44	16	3.058	16/09/29 18:44	15	3.031
16/12/05 19:19	17	3.043	16/12/05 19:19	44	2.646
16/10/20 18:35	18	3.043	16/10/20 18:35	45	2.645
16/11/26 16:42	35	2.582	16/11/26 16:42	77	2.368
16/11/26 16:49	37	2.518	16/11/26 16:49	80	2.360
16/10/21 15:28	62	2.129	16/10/21 15:28	42	2.670
16/11/01 19:13	63	2.115	16/11/01 19:13	40	2.694
16/12/27 18:50	65	2.095	16/12/27 18:50	10	3.179
16/12/19 20:03	66	2.090	16/12/19 20:03	8	3.187
16/09/27 02:42	68	2.040	16/09/27 02:42	82	2.346
16/08/23 06:42	74	2.018	16/08/23 06:42	81	2.353

There are no matching records between the one-month dataset and the other datasets in the top 100 audit logs, this is because data within May 2017 does not appear in the top 100 audit logs for the other datasets. However, there are fourteen pairs of matching audit logs between the six month and eighteen-month datasets. For example, the audit log on the 7th December 2016 at 20:17 is the 2nd most anomalous data point in the 6-month dataset and the 85th most anomalous data point in the eighteen-month dataset. However, there are some unusual observations within the dataset. For example, the datapoint of 26th September 2016 at 17:02 is the most anomalous for the eighteen-month dataset, and falls within the date range of the six-month dataset. It would therefore be expected that this data point would also appear as one of the notable data points in the six-month dataset but it does not. The anomaly score for the eighteen-month dataset has a routine ID (ASF SPC CAA MPI) of 13.339 whereas the anomaly score for the same routine in the six-month dataset is 5.789, which is the 21st most anomalous routine in the initial six month dataset. This indicates that when more data is provided to PARISS, the most anomalous datapoints can be reanalysed as more typical behaviour, and data points which are of some interest in the six month analysis become of high interest when more data is provided and the datapoint becomes more unusual. The full dataset is available for reference in the appendix.

A limitation of this work is that HILML is not applied to these results as due to the tokenisation of the data they would not return meaningful results. The results determined audit logs that were detected as displaying unusual activity, however this could not be validated as the data could not be untokenised for further investigation. Therefore, applying a HILML score for these audit logs would not reflect a genuine analysis of the audit log, as this was not possible. As discussed in Chapter 7, an aim of future work would be validating the system framework on a real-world dataset which has not been anonymised.

6.6. Summary

In this chapter, the real-world dataset described in chapter 5 is used to validate the system framework detailed in chapter 4. The framework is applied to three sub-sets of the dataset, 1 month, 6 months and 18 months. The results are then compared and discussed. In chapter 7, the thesis conclusion is presented, along with thoughts on the direction of potential future work.

7. Conclusion and Future Work

7.1. Introduction

Electronic Patient Record systems represent a fundamental shift for healthcare through increasing availability of healthcare data to providers. However, this ubiquity of data is causing privacy concerns among patients who feel their data is less secure electronically. Current procedure-based models are insufficient and most information security incidents are detected by the patient, or staff member, whose privacy has been violated, causing reputational damage to the hospital. Therefore, this thesis represents research towards a system to ensure confidentiality and privacy of EPR systems. Through the use of machine learning techniques which employ human-in-the-loop and density-based outlier detection techniques, proactive monitoring of EPR audit logs is achieved. Proactive monitoring allows for inappropriate behaviour to be detected and managed, in addition to prompting a cultural shift among employees to refrain from such behaviour in future.

7.2. Thesis Summary

Patient Privacy within healthcare infrastructures is a key concern for hospitals. Data breaches can have the unintended consequence of patients losing trust in hospitals and being selective about the information they share, which can affect care. Access to EPRs is typically managed through access control and reactive auditing. However, a proactive monitoring approach is required to maintain patient trust. Visualisation and machine learning techniques are leveraged to utilise audit logs and actively monitor EPR accesses for inappropriate usage.

In this section, an overview of the research presented in this thesis is discussed, with a summary of each chapter provided.

7.2.1. Contribution to Knowledge

The aim of the project was to develop a novel system framework capable of autonomously detecting unusual data behaviour within an EPR and presenting them to an analyst for review. To achieve this aim, a literature review of healthcare infrastructures was performed, in addition to a review of machine learning and visualisation techniques. A novel system framework was defined and developed and validated using a real-world dataset.

The research presented in this thesis offers a significant contribution in patient privacy monitoring. Proactive monitoring of audit logs is required to achieve comprehensive situational awareness of activity within an EPR. The system framework uses an unsupervised machine learning algorithm LOF to detect unusual data patterns. The system framework can be utilised in any hospital to identify anomalous behaviours and over time becomes more adapted to each instance through HILML. By combining unsupervised machine learning and HILML, the system framework provides an analyst with a holistic and tailored view of patient privacy violations within their Trust. This is a novel approach to patient privacy monitoring. The system framework uses a unique visualisation tool (the User Interface) which ranks events by severity, enabling an at-a-glance view of the number of flags to be reviewed by the analyst, in order of priority.

The framework presented in this thesis is novel through the unique combination of existing methods to the context of EPR patient privacy. The LOF density-based outlier detection algorithm is well-established as an outlier detection tool but has not previously been applied to EPRs. In doing so, the framework enables hidden patterns of data to be discerned and investigated, which current procedure-based solutions cannot detect. Similarly, human-in-the-loop machine learning techniques have previously been applied in cyber security contexts, but are applied to EPRs here. Finally, previous research in outlier detection within EPRs focuses primarily on the algorithms and techniques to detect anomalies, whereas in the novel framework presented here, consideration is given to the visualisation of these results in order to aid situational awareness of the patient privacy analyst.

The dataset used for this research is provided by a specialist Liverpool-based hospital. PARISS is able to detect 144 anomalous behaviours (0.014%) in an unlabelled dataset of 1,007,727 audit logs. This includes 0.66% of the total users on the system, 0.17% of patient record accesses, 0.74% of routine accesses, and 0.53% of the devices used.

7.3. Future Work

As PARISS is embedded into hospitals, there is scope for further work to enhance its use as a patient privacy monitoring framework. PARISS was developed using a Waterfall development model. Future work will involve iterative iterations to the design and will therefore incorporate an Agile development model in future work. In doing so, stakeholder feedback will be incorporated to continually refine the system design. In this section, possible further research to continue to develop PARISS is discussed.

7.3.1. Live Data

The datasets were provided for this work on condition of tokenisation, so that patients and staff could not be identified. Once a fully-fledged concept was developed, the hospital indicated a willingness to deploy the system within the hospital. Therefore, future work will involve gathering feedback and testing the system with information security analysts on un-anonymised live data in a hospital. This will validate the concept on real world non-anonymised data. Using non-anonymised data will allow for other factors to be taken into consideration to determine motivational indicators. For example, determining a user's role may provide valuable insight. Admin staff and doctors may both have access to the EPR. If an admin staff member is accessing clinical data, this would achieve a higher anomaly score than a doctor, and may indicate a breach. Additionally, a patient's characteristics, such as a VIP or a relation to the patient, may provide context to determine whether a patient's confidentiality has been breached. Accounting for additional factors such as these will continuously improve the system.

The features discussed in the thesis compare every activity performed associated with each ID, but without detail. For example, for each User it compares the duration of all actions performed for that user. This can broadly identify anomalous behaviour, but for a more nuanced approach, other factors can be taken into consideration. For example, how long a user typically spends performing a certain task, or accesses a specific device, or with a particular patient. By calculating the local outlier factor for these behaviours, and assigning each a weighted score, these can be factored together to provide data-driven insight of potential EPR misuse. Additionally, currently inputting new data to calculate their LOF values is a manual process and not in real-time. This will be explored further with an aim to automate this and improve update efficiency within the big data context of EPR audit logs.

7.3.2. Visualisation and Virtual Reality

Future work will also build on the visualisation approach undertaken in the research case studies presented. The Virtual Hospital option presents the user another screen where a 3D virtual layout of the hospital is displayed. This enables the user to view recreations of events (through using the Date & Time, User and Device fields. The user is asked which time period they wish to view, and a simulation can be run to aid situational awareness. The Virtual Reality option enables the user to view the virtual hospital, enabling a more immersive experience. Figure 62 displays the virtual hospital and virtual reality sections of PARISS.

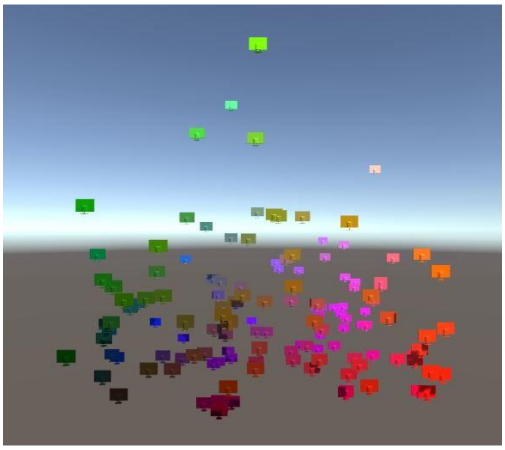


Figure 62 - PARISS User Interface - Virtual Hospital VR

The device height represents its anomaly score. These devices can be mapped to their location within a 3D virtual hospital, allowing an at a glance view of where anomalous activity is happening. Additionally, several anomalous data behaviours can be viewed over a time period, allowing a reconstruction to be created of user behaviour in order to aid the patient privacy officer’s investigation.

7.3.3. Bespoke Anomaly Scores for Routine IDs

Future Work will involve normalising the data further with a case study of the routine ‘Pharmacy Orders’. This routine accounts for approximately 21.27% of the actions performed on the EPR. It is therefore possible to use this as a case study to understand user roles within the dataset and compare similar actions, in order to identify anomalous behaviours. Factors other than solely the duration of the routine (such as the date and time an action is performed) will be considered. Additionally, a quantitative model-based approach that takes into account the duration and the sequence of events during the interaction of the user with the EPR will be explored.

7.3.4. Multiple Hospitals

The incorporation of further supervised learning models into PARISS can be achieved through leveraging signature detection methods. If the system is deployed at more than one hospital, a database of known patient privacy violations can be accrued and shared. For example, as an analyst in one hospital interacts with PARISS, roles within the hospital (such as Doctor/Nurse/Admin) will be defined. This role-based profile can be applied when deploying the system at another hospital. This process is outlined in Figure 63.

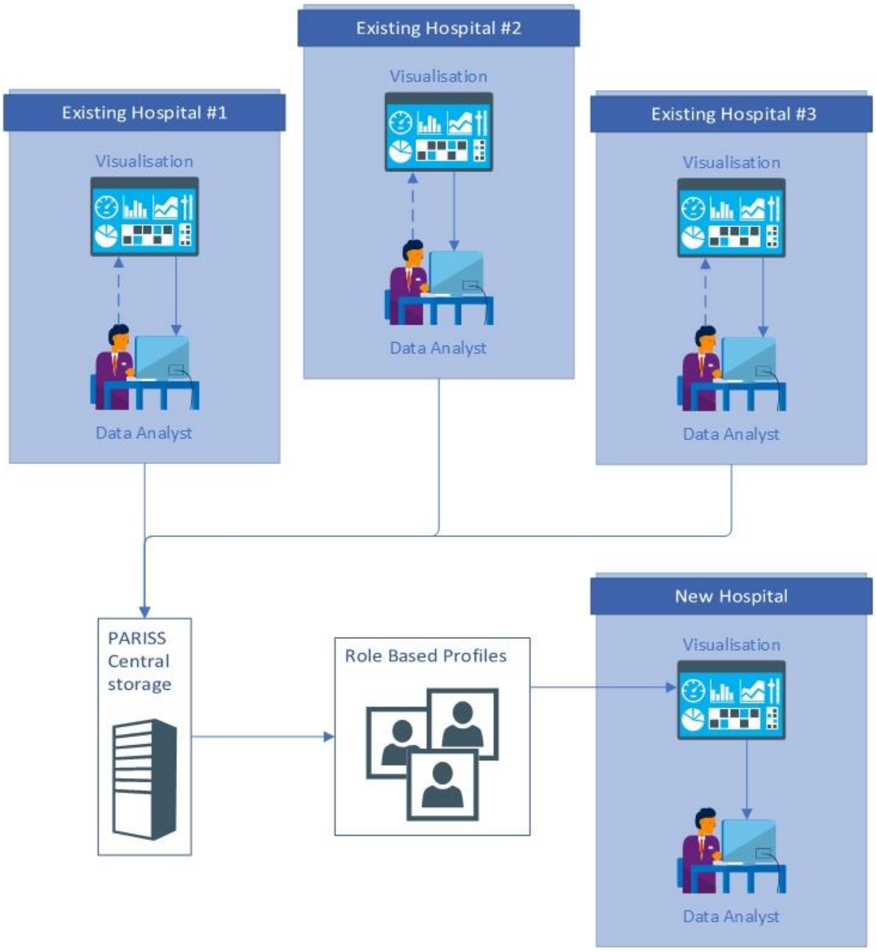


Figure 63 - PARISS in Multiple Hospitals

7.4. Limitations of Work

The key limitations of this work were the use of a tokenised dataset to validate the system framework. Due to this limitation, the results could not be verified as anomalous by the partner hospital. Additionally, the audit log data itself concatenated Routine behaviour into a single audit log, resulting in nuanced duration details being lost within the dataset.

Under a more formal project scope with the partner hospital, a more agile system development approach would have been used to develop the system framework, rather than the waterfall model used due to the limited contact time available.

7.5. Concluding Remarks

The far-reaching consequences of this work are illustrated with a prediction: This research project will increase the situational awareness of data flow and actively address the issue of data misuse. Machine learning algorithms have the capability to observe and learn patterns of data and profile users' behaviour, which can then be represented visually. The work will result in the development of a system that can be used by healthcare practitioners to increase the protection of their EPR records. This will make the UK, not only one of the safest places to conduct business, but also one of the most secure in protecting patient privacy in healthcare systems.

References

[1] G. Martin, S. Ghafur, J. Kinross, C. Hankin, and A. Darzi, “WannaCry - A year on,” *BMJ (Online)*, vol. 361. BMJ Publishing Group, 2018.

[2] M. Honeyman, P. Dunn, and H. Mckenna, “A digital NHS?,” 2016.

[3] D. Kruger and T. Anschutz, “A new approach to IT security: information object-level controls have the potential to better protect hospitals from data breaches by building security controls into the information itself.,” in *Healthcare Financial Management*, 2013, p. 104+.

[4] R. Vargheese, “Dynamic Protection for Critical Health Care Systems Using Cisco CWS: Unleashing the Power of Big Data Analytics,” in *2014 Fifth International Conference on Computing for Geospatial Research and Application*, 2014, pp. 77–81.

[5] P. S. Mathew and A. S. Pillai, “Big Data solutions in Healthcare: Problems and perspectives,” in *2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, 2015, pp. 1–6.

[6] S. M. Riazul Islam, M. Humaun Kabir, and M. Hossain, “The Internet of Things for Health Care: A Comprehensive Survey,” *IEEE Access*, vol. 3, pp. 678–708, 2015.

[7] W. Hurst and C. Dobbins, “Guest Editorial Special Issue on: Big Data Analytics in Intelligent Systems,” *J. Comput. Sci. Appl. Spec. Issue Big Data Anal. Intell. Syst.*, vol. 3, no. 3A, pp. 1–9, 2015.

[8] K. Chui *et al.*, “Disease Diagnosis in Smart Healthcare: Innovation, Technologies and Applications,” *Sustainability*, vol. 9, no. 12, p. 2309, Dec. 2017.

[9] D. Arney, K. K. Venkatasubramanian, O. Sokolsky, and I. Lee, “Biomedical devices and systems security.,” *Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 2011, pp. 2376–9, Jan. 2011.

[10] 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)*, 2011, pp. 81–85.

[11] S. Mansfield-Devine, “Leaks and ransoms – the key threats to healthcare organisations,” *Netw. Secur.*, vol. 2017, no. 6, pp. 14–19, Jun. 2017.

[12] H. Government, “National Cyber Security Strategy 2016-2021,” 2016.

[13] L. Ayala, *Cybersecurity for Hospitals and Healthcare Facilities: A Guide to Detection and Prevention*. 2016.

[14] I. N. Shu and H. Jahankhani, “The Impact of the new European General Data Protection Regulation (GDPR) on the Information Governance Toolkit in Health and Social Care with Special Reference to Primary Care in England,” in *Proceedings - 2017 Cybersecurity and Cyberforensics Conference, CCC 2017*, 2018, vol. 2018-Septe, pp. 31–37.

[15] ICO, “Data security incident trends,” 2016. [Online]. Available: https://ico.org.uk/action-weve-taken/data-security-incident-trends/. [Accessed: 02-Oct-2017].

[16] C. Czeschik, “Black Market Value of Patient Data,” in *Digital Marketplaces Unleashed*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2018, pp. 883–893.

[17] 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.

[18] M. Honeyman, D. Dunn, and H. McKenna, “A digital NHS? : : an introduction to the digital agenda and plans for implementation.,” *Brief. ;*, no. September, 2016.

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

[20] J. E. Rekdal, 12hmisa -Forensics Pieter, and B. Ruthven, “Advanced Persistent Threat (APT): Beyond the hype,” 2013.

[21] H. M. Chao, C. M. Hsu, and S. G. Miaou, “A data-hiding technique with authentication, integration, and confidentiality for electronic patient records,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 6, no. 1, pp. 46–53, Mar. 2002.

[22] F. Bilimleri Dergisi, A. J. Rashidi, K. D. Ahmadi, and M. Heidarpour, “Cyber Situational Awareness using Intelligent Information Fusion Engine (IIFE),” *Cumhur. Univ. Fac. Sci. Sci. J.*, vol. 36, no. 3, p. 36, 2015.

[23] P. Nuttachot, M. Anirach, S. Supaporn, and N. Nati, “Multi-Dimensional Visualization for Network Forensic Analysis,” *Int. J. Adv. Comput. Technol.*, vol. 4, no. 5, pp. 222–232, Mar. 2012.

[24] J. Blocki and N. Christin, “Audit Mechanisms for Privacy Protection in Healthcare Environments,” *Proc. 2nd USENIX Conf. Heal. Secur. Priv.*, pp. 4–5, 2011.

[25] A. Appari and M. E. Johnson, “Information security and privacy in healthcare: current state of research,” *Int. J. Internet Enterp. Manag.*, vol. 6, no. 4, p. 279, 2010.

[26] J. Lee, B. Kang, and S. H. Kang, “Integrating independent component analysis and local outlier factor for plant-wide process monitoring,” *J. Process Control*, vol. 21, no. 7, pp. 1011–1021, Aug. 2011.

[27] A. Boddy, W. Hurst, M. Mackay, and A. El Rhalibi, “Securing Health Care Information Systems using Visualisation Techniques,” in *21st UKAIS Conference on Information Systems*, 2016.

[28] 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.*, 2016.

[29] 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.

[30] A. Boddy, W. Hurst, M. Mackay, A. El Rhalibi, and M. Mwansa, “Data Analysis Techniques to Visualise Accesses to Patient Records in Healthcare Infrastructures,” 2018, pp. 65–70.

[31] A. Boddy, W. Hurst, M. Mackay, and A. El Rhalibi, “A Hybrid Density-Based Outlier Detection Model for Privacy in Electronic Patient Record Systems,” in *International Conference on Information Management 2019*, 2019.

[32] A. Boddy, W. Hurst, M. Mackay, and A. El Rhalibi, “Establishing Situational Awareness for Securing Healthcare Patient Records,” in *HEALTHINFO2019*, 2019.

[33] A. Boddy *et al.*, “An Investigation into Healthcare-Data Patterns,” *MDPI Futur. Internet*, vol. 11, p. 30, Jan. 2019.

[34] A. Boddy, W. Hurst, M. Mackay, and A. El Rhalibi, “Density-Based Outlier Detection for Safeguarding Electronic Patient Record Systems (January 2019),” *IEEE Access*, vol. PP, p. 1, Mar. 2019.

[35] BBC, “Medical devices vulnerable to hackers - BBC News.” [Online]. Available: http://www.bbc.co.uk/news/technology-34390165. [Accessed: 16-Dec-2015].

[36] P. A. H. Williams and V. McCauley, “Always connected: The security challenges of the healthcare Internet of Things,” in *2016 IEEE 3rd World Forum on Internet of Things, WF-IoT 2016*, 2017, pp. 30–35.

[37] J. M. Holroyd-Leduc, D. Lorenzetti, S. E. Straus, L. Sykes, and H. Quan, “The impact of the electronic medical record on structure, process, and outcomes within primary care: A systematic review of the evidence,” *Journal of the American Medical Informatics Association*, vol. 18, no. 6. pp. 732–737, 2011.

[38] A. Arabo, “Cyber Security Challenges within the Connected Home Ecosystem Futures,” in *Procedia Computer Science*, 2015, vol. 61, pp. 227–232.

[39] M. Rong, C. Han, and L. Liu, “Critical Infrastructure Failure Interdependencies in the 2008 Chinese Winter Storms,” in *2010 International Conference on Management and Service Science*, 2010, pp. 1–4.

[40] R. Skowyra, S. Bahargam, and A. Bestavros, “Software-Defined IDS for securing embedded mobile devices,” in *2013 IEEE High Performance Extreme Computing Conference (HPEC)*, 2013, pp. 1–7.

[41] C. Chalmers, W. Hurst, M. Mackay, and P. Fergus, “Smart Meter Profiling For Health Applications,” in *The Internal Joint Conference on Neural Networks*, 2015, pp. 1–7.

[42] A. Sawand, S. Djahel, Z. Zhang, and F. Nait-Abdesselam, “Multidisciplinary approaches to achieving efficient and trustworthy eHealth monitoring systems,” in *2014 IEEE/CIC International Conference on Communications in China (ICCC)*, 2014, pp. 187–192.

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

[44] D. Rose and N. Joshi, *The Digital Doctor: Hope, Hype, and Harm at the Dawn of Medicine’s Computer Age*, vol. 71, no. 1. 2018.

[45] L. B. O. Jans, J. M. L. Bosmans, K. L. Verstraete, and R. Achten, “Optimizing communication between the radiologist and the general practitioner.,” *JBR-BTR*, vol. 96, no. 6, pp. 388–90.

[46] J. Hippisley-Cox, “Validity and completeness of the NHS Number in primary and secondary care Electronic data in England 1991-2013,” 1991.

[47] O. O. Akinsanya, M. Papadaki, and L. Sun, “Current cybersecurity maturity models: How effective in healthcare cloud?,” in *CEUR Workshop Proceedings*, 2019, vol. 2348, pp. 211–222.

[48] M. Patel and J. Wang, “Applications, challenges, and prospective in emerging body area networking technologies,” *IEEE Wirel. Commun.*, vol. 17, no. 1, pp. 80–88, Feb. 2010.

[49] S. Walker-Roberts, M. Hammoudeh, and A. Dehghantanha, “A Systematic Review of the Availability and Efficacy of Countermeasures to Internal Threats in Healthcare Critical Infrastructure,” *IEEE Access*, vol. 6. Institute of Electrical and Electronics Engineers Inc., pp. 25167–25177, 19-Mar-2018.

[50] O. Kocabas, T. Soyata, and M. K. Aktas, “Emerging Security Mechanisms for Medical Cyber Physical Systems,” *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 13, no. 3, pp. 401–416, May 2016.

[51] R. Mitchell and I.-R. Chen, “Behavior Rule Specification-Based Intrusion Detection for Safety Critical Medical Cyber Physical Systems,” *IEEE Trans. Dependable Secur. Comput.*, vol. 12, no. 1, pp. 16–30, Jan. 2015.

[52] Drager, “Derriford Hospital , Plymouth,” *Drager: Derriford Hospital: One patient, one monitor, one acute care episode*, 2012. [Online]. Available: https://www.draeger.com/Products/Content/derriford_hospital_cs_9066449_en.pdf. [Accessed: 22-Jun-2018].

[53] R. Ogie, “Bring Your Own Device: An overview of risk assessment.,” *IEEE Consum. Electron. Mag.*, vol. 5, no. 1, pp. 114–119, Jan. 2016.

[54] Bosheng Zhou, A. Marshall, and Tsung-Han Lee, “Wireless Security Issues in Pervasive Computing,” in *2010 Fourth International Conference on Genetic and Evolutionary Computing*, 2010, pp. 509–512.

[55] M. Meingast, T. Roosta, and S. Sastry, “Security and privacy issues with health care information technology,” in *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 2006, pp. 5453–5458.

[56] G. Zhao and D. W. Chadwick, “On the modeling of Bell-LaPadula security policies using RBAC,” in *Proceedings of the Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises, WETICE*, 2008, pp. 257–262.

[57] Fair Warning, “How Privacy Considerations Drive Patient Decisions and Impact Patient Care Outcomes Purpose of the Study and Executive Overview Report,” 2011.

[58] 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.

[59] S. D. Castilho, E. P. Godoy, T. W. L. Castilho, and A. F. Salmen, “Proposed model to implement high-level Information Security in Internet of Things,” in *2017 Second International Conference on Fog and Mobile Edge Computing (FMEC)*, 2017, pp. 165–170.

[60] K. Vangury, “Bring your own device security issues and challenges,” in *2014 IEEE 11th Consumer Communications and Networking Conference (CCNC)*, 2014, pp. 80–85.

[61] Q. Shafi, “Cyber Physical Systems Security: A Brief Survey,” in *2012 12th International Conference on Computational Science and Its Applications*, 2012, pp. 146–150.

[62] D. B. Kramer *et al.*, “Security and privacy qualities of medical devices: an analysis of FDA postmarket surveillance.,” *PLoS One*, vol. 7, no. 7, p. e40200, Jan. 2012.

[63] Q. Chen and J. Lambright, “Towards Realizing a Self-Protecting Healthcare Information System,” in *2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC)*, 2016, pp. 687–690.

[64] P. Scott and R. Worden, “Semantic mapping to simplify deployment of HL7 v3 Clinical Document Architecture,” *J. Biomed. Inform.*, vol. 45, no. 4, pp. 697–702, Aug. 2012.

[65] Microsoft, “Netstat,” *TechNet*. [Online]. Available: https://technet.microsoft.com/en-us/library/bb490947.aspx. [Accessed: 08-Dec-2016].

[66] M. Jacomy *et al.*, “ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software,” *PLoS One*, vol. 9, no. 6, p. e98679, Jun. 2014.

[67] Y. Hu, “Efficient, High-Quality Force-Directed Graph Drawing,” *Math. J.*, vol. 10, no. 1, pp. 37–71, 2005.

[68] T. M. J. Fruchterman and E. M. Reingold, “Graph drawing by force-directed placement,” *Softw. Pract. Exp.*, vol. 21, no. 11, pp. 1129–1164, 1991.

[69] S. G. Kobourov, “Force-directed drawing algorithms,” in *Handbook of Graph Drawing and Visualization (Discrete Mathematics and Its Applications)*, 2013, pp. 383–408.

[70] S. Martin, W. M. Brown, R. Klavans, and K. W. Boyack, “OpenOrd: an open-source toolbox for large graph layout,” 2011, vol. 7868, p. 786806.

[71] K. W. Boyack, B. N. Wylie, and G. S. Davidson, “Domain visualization using VxInsight® for science and technology management,” *J. Am. Soc. Inf. Sci. Technol.*, vol. 53, no. 9, pp. 764–774, Jul. 2002.

[72] J. Barnes and P. Hut, “A hierarchical O(N log N) force-calculation algorithm,” *Nature*, vol. 324, no. 6096, pp. 446–449, Dec. 1986.

[73] N. Promrit, M. Merabti, A. Mingkhwan, and W. Hurst, “Advanced Feature Extraction for Evaluating Host Behaviour in a Network,” *15th Annu. Conf. Converg. Telecommun. Netw. Broadcast.*, 2014.

[74] W. H. Loob, “The Safe Medical Devices Act of 1990,” *J. Clin. Eng.*, vol. 16, no. 1, pp. 35–38, 1991.

[75] U. F. and D. Administration, “Manufacturer and user facility device experience (MAUDE),” 1995.

[76] H. Yin, D. Song, M. Egele, C. Kruegel, and E. Kirda, “Panorama: Capturing system-wide information flow for malware detection and analysis,” in *Proceedings of the ACM Conference on Computer and Communications Security*, 2007, pp. 116–127.

[77] FDA, “MERGE HEALTHCARE MERGE HEMO PROGRAMMABLE DIAGNOSTIC COMPUTER,” *MAUDE Adverse Event Report*, 2016. [Online]. Available: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/detail.cfm?mdrfoi__id=5487204. [Accessed: 08-Dec-2016].

[78] E. E. Schultz, “RPC in Windows systems: What you don’t know could hurt you,” *Netw. Secur.*, vol. 2004, no. 6, pp. 5–8, 2004.

[79] W. M. Tierney *et al.*, “Provider Responses to Patients Controlling Access to their Electronic Health Records: A Prospective Cohort Study in Primary Care,” *J. Gen. Intern. Med.*, vol. 30, no. 1, pp. 31–37, 2015.

[80] S. Chapman and S. Thomas, “Falls and the rise of the GP contract: An EMIS web protocol and template to help identify frail patients,” in *British Journal of Community Nursing*, 2017, vol. 22, no. 11, pp. 554–556.

[81] M. Gibbs, “Evaluating Cyber Threats to the United Kingdom’s National Health Service (NHS) Spine Network,” in *Advances in Intelligent Systems and Computing*, 2018, vol. 738, pp. 39–42.

[82] M. M. Gunal and M. Pidd, “Interconnected DES models of emergency, outpatient, and inpatient departments of a hospital,” in *Proceedings - Winter Simulation Conference*, 2007, pp. 1461–1466.

[83] Y. Chen, N. Lorenzi, S. Nyemba, J. S. Schildcrout, and B. Malin, “We work with them? Healthcare workers interpretation of organizational relations mined from electronic health records,” *Int. J. Med. Inform.*, vol. 83, no. 7, pp. 495–506, Jul. 2014.

[84] T. A. Shamliyan, S. Duval, J. Du, and R. L. Kane, "Just what the doctor ordered. Review of the evidence of the impact of computerized physician order entry system on medication errors," *Health Serv. Res.*, vol. 43, no. 1 P1, pp. 32–53, Jun. 2008.

[85] C. Chen, T. Garrido, D. Chock, G. Okawa, and L. Liang, "The Kaiser Permanente electronic health record: Transforming and streamlining modalities of care," *Health Aff.*, vol. 28, no. 2, pp. 323–333, Mar. 2009.

[86] J. A. Zlabek, J. W. Wickus, and M. A. Mathiason, "Early cost and safety benefits of an inpatient electronic health record," *J. Am. Med. Informatics Assoc.*, vol. 18, no. 2, pp. 169–172, Mar. 2011.

[87] N. USENIX Association., D. ACM SIGMOBILE., M. ACM Digital Library., and S. Moulton, "A sensor-based, web service-enabled, emergency medical response system," in *Proceedings of the 2005 workshop on End-to-end, sense-and-respond systems, applications and services*, 2005, p. 48.

[88] J. Sheather and S. Brannan, "Patient confidentiality in a time of care.data," *BMJ (Online)*, vol. 347, British Medical Journal Publishing Group, p. f7042, 27-Nov-2013.

[89] S. A. Hameed, H. Yuchoh, and W. F. Al-Khateeb, "A model for ensuring data confidentiality: In healthcare and medical emergency," in *2011 4th International Conference on Mechatronics: Integrated Engineering for Industrial and Societal Development, ICOM'11 - Conference Proceedings*, 2011, pp. 1–5.

[90] C. Papoutsi, J. E. Reed, C. Marston, R. Lewis, A. Majeed, and D. Bell, "Patient and public views about the security and privacy of Electronic Health Records (EHRs) in the UK: results from a mixed methods study," *BMC Med. Inform. Decis. Mak.*, vol. 15, no. 1, p. 86, Oct. 2015.

[91] G. Laurie, M. L. Stevens, K. H. Jones, and C. Dobbs, "A Review of Evidence Relating to Harms Resulting from Uses of Health and Biomedical Data," Nuffield Council on Bioethics, Feb. 2014.

[92] C. Papoutsi, J. E. Reed, C. Marston, R. Lewis, A. Majeed, and D. Bell, "Patient and public views about the security and privacy of Electronic Health Records (EHRs) in the UK: results from a mixed methods study," *BMC Med. Inform. Decis. Mak.*, vol. 15, no. 1, p. 86, Oct. 2015.

[93] L. Presser, M. Hruskova, H. Rowbottom, and J. Kancir, "Care data and access to UK health records: patient privacy and public trust," *Technol. Sci.*, pp. 1–31, 2015.

[94] 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.

[95] 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.

[96] 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.

[97] A. Ferreira, R. Cruz-Correia, L. Antunes, and D. Chadwick, "Access control: how can it improve patients' healthcare?," *Stud. Health Technol. Inform.*, vol. 127, pp. 65–76, 2007.

[98] J. Salazar-Kish, D. Tate, P. D. Hall, and K. Homa, "Development of CPR security using impact analysis.," *Proceedings. AMIA Symp.*, pp. 749–53, 2000.

[99] P. V Asaro, R. L. Herting, A. C. Roth, and M. R. Barnes, "Effective audit trails--a taxonomy for determination of information requirements.," *Proceedings. AMIA Symp.*, pp. 663–5, 1999.

[100] B. Malin and E. Airoldi, "Confidentiality preserving audits of electronic medical record access.," *Stud. Health Technol. Inform.*, vol. 129, no. Pt 1, pp. 320–4, 2007.

[101] 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.

[102] F. . F. Rezaeibagha, K. T. . K. T. Win, and W. . W. Susilo, "A systematic literature review on security and privacy of electronic health record systems : technical perspectives," *Heal. Inf.*, vol. 44, no. 3, pp. 1–16, 2015.

[103] J. Salazar-Kish, D. Tate, P. D. Hall, and K. Homa, "Development of CPR security using impact analysis," *Proc AMIA Symp*, pp. 749–753, 2000.

[104] R. Clarke and T. Youngstein, "Cyberattack on Britain's national health service — A wake-up call for modern medicine," *New England Journal of Medicine*, vol. 377, no. 5. pp. 409–411, 2017.

[105] B. Malin, S. Nyemba, and J. Paulett, "Learning relational policies from electronic health record access logs," *J. Biomed. Inform.*, vol. 44, no. 2, pp. 333–342, Apr. 2011.

[106] R. J. Gallagher, D. Ph, S. Sengupta, G. Hripcsak, R. C. Barrows, and P. D. Clayton, "An Audit Server for Monitoring Usage of Clinical Information Systems," *Proc. AMIA Symp.*, p. 10032, 1997.

[107] L. S. Sulmasy, A. M. López, and C. A. Horwitch, "Ethical Implications of the Electronic Health Record: In the Service of the Patient," *J. Gen. Intern. Med.*, vol. 32, no. 8, pp. 935–939, Aug. 2017.

[108] N. Shen *et al.*, "Understanding the patient privacy perspective on health information exchange: A systematic review," *International Journal of Medical Informatics*, vol. 125. Elsevier Ireland Ltd, pp. 1–12, 01-May-2019.

[109] C. Esposito, A. De Santis, G. Tortora, H. Chang, and K. K. R. Choo, "Blockchain: A Panacea for Healthcare Cloud-Based Data Security and Privacy?," *IEEE Cloud Comput.*, vol. 5, no. 1, pp. 31–37, Jan. 2018.

[110] D. Birnbaum, K. Gretsinger, M. G. Antonio, E. Loewen, and P. Lacroix, “Revisiting public health informatics: patient privacy concerns,” *Int. J. Heal. Gov.*, vol. 23, no. 2, pp. 149–159, 2018.

[111] K. Abouelmehdi, A. Beni-Hessane, and H. Khaloufi, “Big healthcare data: preserving security and privacy,” *J. Big Data*, vol. 5, no. 1, Dec. 2018.

[112] T. Glenn and S. Monteith, “Privacy in the Digital World: Medical and Health Data Outside of HIPAA Protections,” *Current Psychiatry Reports*, vol. 16, no. 11. 2014.

[113] I. T. Agaku, A. O. Adisa, O. A. Ayo-Yusuf, and G. N. Connolly, “Concern about security and privacy, and perceived control over collection and use of health information are related to withholding of health information from healthcare providers,” *J. Am. Med. Informatics Assoc.*, vol. 21, no. 2, pp. 374–378, Mar. 2014.

[114] J. S. Ancker, M. Silver, M. C. Miller, and R. Kaushal, “Consumer experience with and attitudes toward health information technology: A nationwide survey,” *J. Am. Med. Informatics Assoc.*, vol. 20, no. 1, pp. 152–156, Jan. 2013.

[115] . The National Partnership for Women & Families, “Making IT Meaningful : How Consumers Value and Trust Health IT,” no. February, pp. 1–278, 2012.

[116] P. H. Keckley, “2012 Survey of U.S. Health Care Consumers: The performance of the health care system and health care reform,” 2012.

[117] Lake Research Partners, “Topline Results From a National Consumer Survey on HIT,” 2010.

[118] M. Brodie *et al.*, “Family Foundation/Harvard School of Public Health.”

[119] by Ruth Helman, M. Greenwald, and P. Fronstin, “The 2008 Health Confidence Survey: Rising Costs Continue to Change the Way Americans Use the Health Care System,” 2008.

[120] L. R. Partners and A. Viewpoint, “Survey Finds Americans Want Electronic Personal Health Information to Improve Own Health Care,” *Suite*, vol. 294, no. November, pp. 1–4, 2006.

[121] N. London and C. December, “Canada : How Privacy Considerations Drive Patient Decisions and Impact Patient Care Outcomes,” 2011.

[122] Harris Interactive Inc, “Many U.S. Adults Are Satisfied with Use of Their Personal Health Information,” *Harris Poll*. 2007.

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

[124] 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.

[125] 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, 1997, pp. 345–360.

[126] 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.

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

[128] 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.

[129] 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.

[130] 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.

[131] 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.

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

[133] 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.

[134] 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*, 2001, pp. 21–27.

[135] J. King, B. Smith, and L. Williams, “Modifying without a trace: general audit guidelines are inadequate for open-source electronic health record audit mechanisms,” *Proc. 2nd ACM SIGHIT Int. Heal. Informatics Symp.*, pp. 305–314, 2012.

[136] S. Nunn, “Managing audit trails,” *J. AHIMA*, vol. 80, no. 9, pp. 44–45, Sep. 2009.

[137] R. Cruz-Correia *et al.*, “Analysis of the quality of hospital information systems audit trails,” *BMC Med. Inform. Decis. Mak.*, vol. 13, no. 1, p. 84, Aug. 2013.

[138] J. D. Halamka, P. Szolovits, D. Rind, and C. Safran, “A WWW Implementation of National Recommendations for Protecting Electronic Health Information,” *J. Am. Med. Informatics Assoc.*, vol. 4, no. 6, pp. 458–464, 1997.

[139] P. V Asaro, R. L. Herting, A. C. Roth, M. R. Barnes, and M. R. Barnes, “Effective audit

trails--a taxonomy for determination of information requirements.," *Proceedings. AMIA Symp.*, pp. 663–5, 1999.

[140] N. R. C. (US) C. on M. P. and S. in H. C. A. of the N. I. Infrastructure, *For the Record*. National Academies Press (US), 1997.

[141] 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*, 2016, pp. 49–54.

[142] J. Kaplan, *Artificial Intelligence: What everyone needs to know*. 2016.

[143] T. M. Mitchell, "The Discipline of Machine Learning," *Mach. Learn.*, vol. 17, no. July, pp. 1–7, 2006.

[144] P. Domingos, "A few useful things to know about machine learning," *Commun. ACM*, vol. 55, no. 10, p. 78, Oct. 2012.

[145] W. Samek, T. Wiegand, and K.-R. Müller, "Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models," Aug. 2017.

[146] A. De Mauro, M. Greco, and M. Grimaldi, "What is big data? A consensual definition and a review of key research topics," in *AIP Conference Proceedings*, 2015, vol. 1644, pp. 97–104.

[147] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *J. Big Data*, vol. 2, no. 1, p. 1, Dec. 2015.

[148] 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.

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

[150] 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*, 2004, p. 486.

[151] 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, 2005, pp. 381–389.

[152] M. H. Ozçelik, E. Duman, M. Isik, and T. Cevik, "Improving a credit card fraud detection system using genetic algorithm," *2010 Int. Conf. Netw. Inf. Technol.*, pp. 436–440, Jun. 2010.

[153] K. K. Venkatasubramanian, A. Banerjee, and S. K. S. Gupta, "Green and Sustainable

Cyber-Physical Security Solutions for Body Area Networks," in *2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks*, 2009, pp. 240–245.

[154] 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.

[155] V. M. Gélvez-Ordoñez, F. Mendoza-Galvis, and J. O. Delgado, "Efecto del tratamiento con ultrasonido sobre algunas propiedades funcionales de la clara de huevo," *Rev. Cient. la Fac. Ciencias Vet. la Univ. del Zulia*, vol. 19, no. 1, pp. 71–76, Jan. 2009.

[156] N. Abe, B. Zadrozny, and J. Langford, "Outlier detection by active learning," in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '06*, 2006, p. 504.

[157] J. Li, K. Y. Huang, J. Jin, and J. Shi, "A survey on statistical methods for health care fraud detection," *Health Care Management Science*, vol. 11, no. 3. Springer US, pp. 275–287, 10-Sep-2008.

[158] L. Buitinck *et al.*, "API design for machine learning software: experiences from the scikit-learn project," *arXiv Prepr. arXiv …*, pp. 1–15, Sep. 2013.

[159] Y. MisnevsIrina, Boriss, Yatskiv, "Data Science: Professional Requirements and Competence Evaluation," in *Baltic J.Modern Computing, Vol. 4*, 2016, pp. 441–453.

[160] 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*, 2007, pp. 1–4.

[161] 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.

[162] W. Li and A. W. Moore, "A machine learning approach for efficient traffic classification," in *IEEE International Workshop on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems - Proceedings*, 2007, pp. 310–317.

[163] N. Promrit and A. Mingkhwan, "Traffic flow classification and visualization for network forensic analysis," *Adv. Inf. Netw.*, 2015.

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

[165] W. Lee, S. J. Stolfo, and K. W. Mok, "A data mining framework for building intrusion detection models," in *Proceedings - IEEE Symposium on Security and Privacy*, 1999, vol. 1999-Janua, pp. 120–132.

[166] W. Hurst, M. Merabti, and P. Fergus, "Behavioural Observation for Critical Infrastructure Security Support | PROTECT Centre," in *Modelling Symposium (EMS), 2013 European*, 2013, pp. 36–41.

[167] G. James, D. Witten, and T. Hastie, “An Introduction to Statistical Learning: With Applications in R.,” 2014.

[168] R. P. W. Duin *et al.*, “PRTools4 A Matlab Toolbox for Pattern Recognition.”

[169] L. I. Kuncheva, *Combining classifiers: ideas and methods*. 2004.

[170] S. Huh and D. Lee, “Linear discriminant analysis for signatures.,” *IEEE Trans. Neural Netw.*, vol. 21, no. 12, pp. 1990–1996, 2010.

[171] D. Whitley, “A genetic algorithm tutorial,” *Stat. Comput.*, vol. 4, no. 2, pp. 65–85, Jun. 1994.

[172] “Use of K-Nearest Neighbor classifier for intrusion detection1,” *Comput. Secur.*, vol. 21, no. 5, pp. 439–448, Oct. 2002.

[173] 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*, 2008, p. 169.

[174] “An overview of anomaly detection techniques: Existing solutions and latest technological trends,” *Comput. Networks*, vol. 51, no. 12, pp. 3448–3470, Aug. 2007.

[175] 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.

[176] M. M. Breunig *et al.*, “LOF,” in *Proceedings of the 2000 ACM SIGMOD international conference on Management of data - SIGMOD ’00*, 2000, vol. 29, no. 2, pp. 93–104.

[177] H.-P. Kriegel, E. Schubert, and A. Zimek, *LoOP: Local Outlier Probabilities*. 2009.

[178] J. Vaidya, J. Kim, A. K. Menon, X. Jiang, and L. Ohno-Machado, “Detecting inappropriate access to electronic health records using collaborative filtering,” *Mach. Learn.*, vol. 95, no. 1, pp. 87–101, Apr. 2013.

[179] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise,” 1996.

[180] G. O. Campos *et al.*, “On the evaluation of unsupervised outlier detection: measures, datasets, and an empirical study,” *Data Min. Knowl. Discov.*, vol. 30, no. 4, pp. 891–927, Jul. 2016.

[181] P. Maestro, M. Chagny, P. P. Jobert, H. Van Damme, and S. Berthier, “Optics,” in *Nanomaterials and Nanochemistry*, vol. 28, no. 2, New York, New York, USA: ACM Press, 2007, pp. 633–659.

[182] J. S. Mihael Ankerst, Markus M. Breunig, Hans-Peter Kriegel, “OPTICS: Ordering Points To Identify the Clustering Structure,” in *ACM SIGMOD’99 Int. Conf. on Management of Data*, 1999.

[183] N. S. Altman, “An introduction to kernel and nearest-neighbor nonparametric regression,” *Am. Stat.*, vol. 46, no. 3, pp. 175–185, 1992.

[184] S. Ramaswamy, R. Rastogi, and K. Shim, "Efficient algorithms for mining outliers from large data sets," *ACM SIGMOD Rec.*, vol. 29, no. 2, pp. 427–438, 2005.

[185] S. Robertson, "Understanding inverse document frequency: on theoretical arguments for IDF," *J. Doc.*, vol. 60, no. 5, pp. 503–520, Oct. 2004.

[186] A. Mishra and S. Vishwakarma, "Analysis of TF-IDF Model and its Variant for Document Retrieval," in *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*, 2015, pp. 772–776.

[187] L. Wei, Y. Yang, R. M. Nishikawa, and Y. Jiang, "A study on several machine-learning methods for classification of malignant and benign clustered microcalcifications," *IEEE Trans. Med. Imaging*, vol. 24, no. 2, pp. 371–380, 2005.

[188] S. Kundu, R. Rangaswami, A. Gulati, M. Zhao, and K. Dutta, "Modeling virtualized applications using machine learning techniques," *ACM SIGPLAN Not.*, vol. 47, no. 7, p. 3, 2012.

[189] U. Naftaly†, N. Intrator‡, and D. Horn§, "Optimal ensemble averaging of neural networks," *Netw. Comput. Neural Syst.*, vol. 8, no. 3, pp. 283–296, 1997.

[190] S. Hashem, "Optimal linear combinations of neural networks," *Neural Networks*, vol. 10, no. 4, pp. 599–614, Jun. 1997.

[191] A. Lazarevic and V. Kumar, "Feature bagging for outlier detection," in *Proceeding of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining - KDD '05*, 2005, p. 157.

[192] A. Ghoting *et al.*, "SystemML: Declarative machine learning on MapReduce," in *Proceedings - International Conference on Data Engineering*, 2011, pp. 231–242.

[193] D. Xin, L. Ma, J. Liu, S. Macke, S. Song, and A. Parameswaran, "Accelerating Human-in-the-loop Machine Learning: Challenges and Opportunities," 2018.

[194] A. Ozgur, J. Srivastava, V. Kumar, L. Ertoz, and A. Lazarevic, "A Comparative Study of Anomaly Detection Schemes in Network Intrusion Detection," 2013, pp. 25–36.

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

[196] F. Ullah, M. A. Habib, M. Farhan, S. Khalid, M. Y. Durrani, and S. Jabbar, "Semantic interoperability for big-data in heterogeneous IoT infrastructure for healthcare," *Sustain. Cities Soc.*, vol. 34, pp. 90–96, Oct. 2017.

[197] H. Koike, K. Ohno, and K. Koizumi, "Visualizing cyber attacks using IP matrix," in *IEEE Workshop on Visualization for Computer Security, 2005. (VizSEC 05).*, 2005, pp. 91–98.

[198] M. C. Wright, "Objective measures of situation awareness in a simulated medical environment," *Qual. Saf. Heal. Care*, vol. 13, no. suppl_1, pp. i65–i71, Oct. 2004.

[199] S. Jajodia, P. Liu, V. Swarup, and C. Wang, Eds., *Cyber Situational Awareness*, vol. 46. Boston, MA: Springer US, 2010.

[200] R. N. Landers, “Developing a Theory of Gamified Learning: Linking Serious Games and Gamification of Learning,” *Simul. Gaming*, vol. 45, no. 6, pp. 752–768, Dec. 2014.

[201] J. Hamari, J. Koivisto, and H. Sarsa, “Does gamification work? - A literature review of empirical studies on gamification,” in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2014, pp. 3025–3034.

[202] R. M. Baron and D. A. Kenny, “The Moderator-Mediator Variable Distinction in Social The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations,” *J. Pers. Soc. Psychol.*, vol. 51, no. 6, pp. 1173–1182, 1986.

[203] C. Croux and G. Haesbroeck, “Influence Function and Efficiency of the Minimum Covariance Determinant Scatter Matrix Estimator,” *J. Multivar. Anal.*, vol. 71, no. 2, pp. 161–190, Nov. 1999.

[204] H. Oja, S. Sirkiä, and J. Eriksson, “Scatter Matrices and Independent Component Analysis,” *Austrian J. Stat.*, vol. 35, no. 2&3, pp. 175–189–175–189, 2006.

[205] C. Di Sarno, V. Formicola, M. Sicuranza, and G. Paragliola, “Addressing security issues of electronic health record systems through enhanced SIEM technology,” in *Proceedings - 2013 International Conference on Availability, Reliability and Security, ARES 2013*, 2013, pp. 646–653.

[206] PatternEx, “White Paper: The PatternEx Virtual Analyst Platform,” 2017.

[207] S. More and A. Rohela, “Vulnerability Assessment and Penetration Testing through Artificial Intelligence.”

[208] B. Esmael, A. Arnaout, R. K. Fruhwirth, and G. Thonhauser, “A Statistical Feature-Based Approach for Operations Recognition in Drilling Time Series,” *Int. J. Comput. Inf. Syst. Ind. Manag. Appl.*, vol. 5, pp. 2150–7988, 2013.

[209] A. Begoyan, “AN OVERVIEW OF INTEROPERABILITY STANDARDS FOR ELECTRONIC HEALTH RECORDS,” 2007.

Appendix

Python Code

Feature Scaling

```
import pandas as pd

import numpy as np

df=pd.io.parsers.read_csv('Data\Data for Scaling.csv')

from sklearn import preprocessing

minmax_scale = preprocessing.MinMaxScaler().fit(df[['Frequency of User UID', 'Mean Duration of User ID']])

df_minmax = minmax_scale.transform(df[['Frequency of User UID', 'Mean Duration of User ID']])

print('Mean after standardization:\nFrequency of User UID={:.2f}, Mean Duration of User ID={:.2f}'

    .format(df_std[:,0].mean(), df_std[:,1].mean()))

print('\nStandard deviation after standardization:\nFrequency of User UID={:.2f}, Mean Duration of User ID={:.2f}'

    .format(df_std[:,0].std(), df_std[:,1].std()))

print('Min-value after min-max scaling:\nFrequency of User UID={:.2f}, Mean Duration of User ID={:.2f}'

    .format(df_minmax[:,0].min(), df_minmax[:,1].min()))

print('\nMax-value after min-max scaling:\nFrequency of User UID={:.2f}, Mean Duration of User ID={:.2f}'

    .format(df_minmax[:,0].max(), df_minmax[:,1].max()))

#Plot Data

%matplotlib inline

from matplotlib import pyplot as plt

def plot():

   plt.figure(figsize=(10,10))

   x1,x2,y1,y2 = plt.axis()

   plt.axis((0,1.2,0,1.2))
```

```
    plt.scatter(df_minmax[:,0], df_minmax[:,1],
            color='blue', label='Min-max scaled', alpha=0.3)
    plt.xlabel('Frequency of User ID')
    plt.ylabel('Mean Duration of User ID')
    plt.legend(loc='upper left')
    plt.grid()
    #plt.tight_layout()
plot()
plt.show()
#Scale data and extract
minmax_scale = preprocessing.MinMaxScaler().fit(df[['Frequency of User UID', 'Mean Duration of User ID', 'Mode Duration of User ID', 'STD Duration of User ID', 'Min Duration of User ID', '5th Percentile of User ID', '25th Percentile  of User ID', 'Median Duration of User ID', '75th Percentile  of User ID', ' 95th Percentile  of User ID', 'Max Duration of User ID', 'Frequency of Patient ID', 'Mean Duration of Patient ID', 'Mode Duration of Patient ID', 'STD Duration of Patient ID', 'Min Duration of Patient ID', '5th Percentile of Patient ID', '25th Percentile of Patient ID', 'Median Duration of Patient ID', '75th Percentile of Patient ID', ' 95th Percentile of Patient ID', 'Max Duration of Patient ID', 'Frequency of Device ID', 'Mean Duration of Device ID', 'Mode Duration of Device ID', 'STD Duration of Device ID', 'Min Duration of Device ID', '5th Percentile of Device ID', '25th Percentile of Device ID', 'Median Duration of Device ID', '75th Percentile of Device ID', ' 95th Percentile of Device ID', 'Max Duration of Device ID', 'Frequency of Routine', 'Mean Duration of Routine ID', 'Mode Duration of Routine ID', 'STD Duration of Routine ID', 'Min Duration of Routine ID', '5th Percentile of Routine ID', '25th Percentile of Routine ID', 'Median Duration of Routine ID', '75th Percentile of Routine ID', ' 95th Percentile of Routine ID', 'Max Duration of Routine ID']])
df_minmax = minmax_scale.transform(df[['Frequency of User UID', 'Mean Duration of User ID', 'Mode Duration of User ID', 'STD Duration of User ID', 'Min Duration of User ID', '5th Percentile of User ID', '25th Percentile  of User ID', 'Median Duration of User ID', '75th Percentile  of User ID', ' 95th Percentile  of User ID', 'Max Duration of User ID', 'Frequency of Patient ID', 'Mean Duration of Patient ID', 'Mode Duration of Patient ID', 'STD Duration of Patient ID', 'Min Duration of Patient ID', '5th Percentile of Patient ID', '25th Percentile of Patient ID', 'Median Duration of Patient ID', '75th Percentile of Patient ID', ' 95th Percentile of Patient ID', 'Max Duration of Patient ID', 'Frequency of Device ID', 'Mean Duration of Device ID', 'Mode Duration of Device ID', 'STD Duration of Device ID', 'Min Duration of Device ID', '5th Percentile of Device ID', '25th Percentile of Device ID', 'Median Duration of Device ID', '75th Percentile of Device ID', ' 95th Percentile of Device ID', 'Max Duration of Device ID', 'Frequency of Routine', 'Mean Duration of Routine ID', 'Mode Duration of Routine ID', 'STD Duration of Routine ID', 'Min Duration of Routine ID', '5th Percentile of Routine ID', '25th Percentile of Routine ID', 'Median Duration of Routine ID', '75th Percentile of Routine ID', ' 95th Percentile of Routine ID', 'Max Duration of Routine ID']])
pd.DataFrame(df_minmax).to_csv("file/save.csv")
```

Feature Testing

```
#data preprocessing

import pandas as pd

from IPython.display import display

 # Read data and drop redundant column.

data = pd.read_csv('C:/Users/cmpwhurs/Desktop/SixMonthRID.csv')

 display(data.head())

 from pandas.tools.plotting import scatter_matrix

 scatter_matrix(data[['RID', 'RMID', 'RMoID', 'RSTDID', 'RMiID', 'R5thID', 'R25thID', 'RMeID', 'R75thID', 'R95thID', 'RMaID']], figsize=(10,10))
```

LOF - UserID

```
#Process repeated for each of the ID Types – User, Patient, Device, Routine

import graphlab as gl

UserID = gl.SFrame.read_csv(‘Data\Normalised Features UserID.csv', header=True)

features = ['Frequency of User UID', 'Mean Duration of User ID', 'Mode Duration of User ID', 'STD Duration of User ID', 'Min Duration of User ID', '5th Percentile of User ID', '25th Percentile of User ID', 'Median Duration of User ID', '75th Percentile of User ID', '95th Percentile of User ID', 'Max Duration of User ID']

UserID = UserID[['User ID'] + features]

UserID.print_rows(5)

UIDMeanMode = gl.anomaly_detection.local_outlier_factor.create(UserID,

                              features=['Mean Duration of User ID', 'Mode Duration of User ID'])

UIDMeanMode['scores'].export_csv('LOF Results\User ID\UIDMeanMode.csv', delimiter=', ', line_terminator='\n', header=True, quote_level=2, double_quote=True, escape_char='\\', quote_char='"', na_rep='', file_header='', file_footer='', line_prefix='', _no_prefix_on_first_value=False)

#This section of code is repeated for all 45 combinations of features.

# All related csv files are then merged into one using cmd commands (copy *.csv combine.csv), and ordered by ID. Then an AVERAGEIF Statement is used to calculate the average LOF value for each ID.
```

Turi – Local Outlier Factor code

Below is the Class definition and utilities for the Local Outlier Factor tool taken from the Turi Code.

"""

Create a :class:`LocalOutlierFactorModel`. This mode contains local outlier factor (LOF) scores for the training data passed to this model, and can predict the LOF score for new observations.

The LOF method scores each data instance by computing the ratio of the average densities of the instance's neighbors to the density of the instance itself. The higher the score, the more likely the instance is to be an outlier *relative to its neighbors*. A score of 1 or less means that an instance has a density similar (or higher) to its neighbors and is unlikely to be an outlier.

The model created by this function contains an SFrame called 'scores' that contains the computed local outlier factors. The `scores` SFrame has four columns:

- *row_id*: the row index of the instance in the input dataset. If a label column is passed, the labels (and the label name) are passed through to this column in the output.

- *density*: the density of instance as estimated by the LOF procedure.

- *neighborhood_radius*: the distance from the instance to its furthest neighbor (defined by 'num_neighbors', and used for predicting the LOF for new points).

- *anomaly_score*: the local outlier factor.

Parameters

dataset : SFrame

Input dataset. The 'dataset' SFrame must include the features specified in the 'features' or 'distance' parameter (additional columns are ignored).

features : list[string], optional

Names of feature columns. 'None' (the default) indicates that all columns should be used. Each column can be one of the following types:

- *Numeric*: values of numeric type integer or float.

- *Array*: array of numeric (integer or float) values. Each array element is treated as a separate variable in the model.

- *Dictionary*: key-value pairs with numeric (integer or float) values. Each key indicates a separate variable in the model.

- *String*: string values.

Please note: if 'distance' is specified as a composite distance, then that parameter controls which features are used in the model. Also note that the column of row labels is automatically removed from the features, if there is a conflict.

label : str, optional

Name of the input column containing row labels. The values in this column must be integers or strings. If not specified, row numbers are used by default.

distance : string or list[list], optional

Function to measure the distance between any two input data rows. If left unspecified, a distance function is automatically constructed based on the feature types. The distance may be specified by either a string or composite distance:

- *String*: the name of a standard distance function. One of 'euclidean', 'squared_euclidean', 'manhattan', 'levenshtein', 'jaccard', 'weighted_jaccard', 'cosine', or 'dot_product'.

- *Composite distance*: the weighted sum of several standard distance functions applied to various features. This is specified as a list of distance components, each of which is itself a list containing three

items:

1. list or tuple of feature names (strings)

2. standard distance name (string)

3. scaling factor (int or float)

num_neighbors : int, optional

Number of neighbors to consider for each point.

threshold_distances : bool, optional

If True (the default), the distance between two points is thresholded. This reduces noise and can improve the quality of results, but at the cost of slower computation.

verbose : bool, optional

If True, print progress updates and model details.

Returns

model : LocalOutlierFactorModel

A trained :class:`LocalOutlierFactorModel`, which contains an SFrame called 'scores' that includes the 'anomaly score' for each input instance.

LocalOutlierFactorModel, graphlab.toolkits.nearest_neighbors

The LOF method scores each data instance by computing the ratio of the average densities of the instance's neighbors to the density of the instance itself. According to the LOF method, the estimated density of a point :math:`p` is the number of :math:`p`'s neighbors divided by the sum of distances to the instance's neighbors. In the following, suppose

:math:`N(p)` is the set of neighbors of point

:math:`p`, :math:`k` is the number of points in this set (i.e. the 'num_neighbors' parameter), and :math:`d(p, x)` is the distance between points :math:`p` and :math:`x` (also based on a user-specified distance function).

.. math:: \hat{f}(p) = \\frac{k}{\sum_{x \in N(p)} d(p, x)}

- The LOF score for point :math:`p` is then the ratio of :math:`p`'s density to the average densities of :math:`p`'s neighbors:

.. math:: LOF(p) = \\frac{\\frac{1}{k} \sum_{x \in N(p)} \hat{f}(x)}{\hat{f}(p)}

- If the 'threshold_distances' flag is set to True, exact distances are replaced by "thresholded" distances. Let Suppose :math:`r_k(x)` is the distance from :math:`x` to its :math:`k`'th nearest neighbor. Then the thresholded distance from point :math:`p` to point :math:`x_i` is

.. math:: d^*(p, x) = \max\{r_k(x), d(p, x)\}

This adaptive thresholding is used in the original LOF paper to reduce noise in the computed distances and improve the quality of the final LOF scores.

- For features that all have the same type, the distance parameter may be a single standard distance function name (e.g. "euclidean"). In the model, however, all distances are first converted to composite distance functions; as a result, the 'distance' field in the model is always a composite distance.

- Standardizing features is often a good idea with distance-based methods, but this model does *not* standardize features.

- If there are several observations located at an identical position, the LOF values can be undefined. An LOF score of "nan" means that a point is either in or near a set of co-located points.

- This implementation of LOF forces the neighborhood of each data instance to contain exactly 'num_neighbors' points, breaking ties arbitrarily.

This differs from the original LOF paper, which allows neighborhoods to expand if there are multiple neighbors at exactly the same distance from an instance.

## Start the training time clock and instantiate an empty model

## Validate the input dataset

## Validate the number of neighbors, mostly to make the error message use the right parameter name.

## Validate the row label against the features *using the nearest neighbors tool with only one row of data.

## Compute the similarity graph based on k and radius, without self-edges, but keep it in the form of an SFrame. Do this *without* the row label, because I need to sort on the row number, and row labels that aren't already in order will be screwed up.

## Bias the distances by making them at least equal to the *reference* point's k'th neighbor radius. This is "reach-distance" in the original paper.

## Find the sum of distances from each point to its neighborhood, then compute the "local reachability density (LRD)". This is not remotely a valid density estimate, but it does have the form

of mass / volume, where the mass is estimated by the number of neighbors in point x's neighborhood, and the volume is estimated by the sum of the distances between x and its neighbors.

## NOTE: if a vertex is co-located with all of its neighbors, the sum of distances will be 0, in which case the inverse distance sum value is 'inf'.

## Join the density of each point back to the nearest neighbors results, then get the average density of each point's neighbors' densities.

## Combine each point's density and average neighbor density into one SFrame, then compute the local outlier factor (LOF).

## Add each point's neighborhood radius to the output SFrame.

## Remove the extraneous columns from the output SFrame and format.

## Substitute in the row labels.

## Post-processing and formatting

class LocalOutlierFactorModel(_CustomModel, _ProxyBasedModel):

"""

Local outlier factor model. The LocalOutlierFactorModel contains the local outlier factor scores for training data passed to the 'create' function, as well as a 'predict' method for scoring new data. Outliers are determined by comparing the probability density estimate of each point to the density estimates of its neighbors.

"""

Compute local outlier factors for new data. The LOF scores for new data instances are based on the neighborhood statistics for the data used when the model was created. Each new point is scored independently.

Parameters

----

dataset : SFrame

Dataset of new points to score with LOF against the training data already stored in the model.

verbose : bool, optional

If True, print progress updates and model details.

Returns

----

out : SArray

LOF score for each new point. The output SArray is sorted to match the order of the 'dataset' input to this method.

## Query the knn model with the new points.

## Join the reference data's neighborhood statistics to the nearest neighbors results.

# Compute reachability distance for each new point and its neighborhood.

## Find the sum of distances from each point to its neighborhood, then compute the "local reachability density" for each query point.

## Find the average density for each query point's neighbors.

## Join the point densities and average neighbor densities into a single SFrame and compute the local outlier factor.

## Remove extraneous columns and format.

Netstat Data

In Table 40 a sample of analysed netstat data is shown displaying firstly, the connection type, secondly the IP source connecting to the DC, thirdly the target of the IP address (the DC server), and fourthly the state of the connection. All data presented is anonymised. The data is a single snapshot of the domain controller server and comprises of 590 established connections of 5688 total ports. In Table 41 a sample of the netstat data is shown, displaying the connection type, the IP source connecting to the EP, the target of the IP address (the EP server) and the state of the connection. The data is a single snapshot of the domain controller server and comprises of 18 established connections of 88 total ports. In Table 42 a sample of the netstat data displays, the connection type, the IP source connecting to the PAS, the target of the IP address (the PAS server) and the connection state. The data is a single snapshot of the domain controller server and comprises of 93 established connections of 173 total ports.

Table 40. Active Directory Domain Controller – TCP Socket Connections Sample Data (Anonymised)

Proto	Local Address	Foreign Address	State
TCP	0.0.0.0:*****	0.0.0.0:0	LISTENING
TCP	**.**.***.16:53	0.0.0.0:0	LISTENING
TCP	**.**.***.16:135	**.**.**.148:53173	ESTABLISHED
TCP	**.**.***.16:135	**.**.***.51:63068	ESTABLISHED

Table 41. Electronic Prescribing System – TCP Socket Connections Sample Data (Anonymised)

Proto	Local Address	Foreign Address	State
TCP	0.0.0.0:*****	0.0.0.0:0	LISTENING
TCP	**.**.***.197:139	0.0.0.0:0	LISTENING
TCP	**.**.***.197:8194	**.**.***.133:50176	ESTABLISHED
TCP	**.**.***.197:8194	**.**.***.133:50326	ESTABLISHED

Table 42. Patient Administration System – TCP Socket Connections Sample Data (Anonymised)

Proto	Local Address	Foreign Address	State
TCP	0.0.0.0:*****	0.0.0.0:0	LISTENING
TCP	**.**.***.16:53	0.0.0.0:0	LISTENING
TCP	**.**.***.16:135	**.**.**.148:53173	ESTABLISHED
TCP	**.**.***.16:135	**.**.***.51:63068	ESTABLISHED

Results sample data

Table 43. DBSCAN Results

User ID	User ID cluster_id	User ID type	Patient ID	Patient ID cluster_id	Patient ID type	Device ID row_id	Device ID cluster_id	Device ID type	Routine ID row_id	Routine ID cluster_id	Routine ID type
119	n/a	noise	**803**	n/a	noise	**1**	n/a	noise	**MPI ZCUS.UK.SCH ZCUS.UK.LETTER**	n/a	noise
126	n/a	noise	**804**	n/a	noise	**2**	n/a	noise	**ZCUS.UK.LETTER VH SPC OE**	n/a	noise
144	n/a	noise	**805**	n/a	noise	**3**	n/a	noise	***** ASF**	6	core
203	n/a	noise	**806**	n/a	noise	**4**	n/a	noise	***** ASF MPI**	6	core
226	n/a	noise	**807**	n/a	noise	**5**	n/a	noise	***** ASF NOTE ZCUS.UK.LETTER**	6	core
297	n/a	noise	**4764**	n/a	noise	**6**	n/a	noise	***** ASF NPC**	6	core
359	n/a	noise	**4765**	n/a	noise	**7**	n/a	noise	***** ASF SPC VH ZCUS.UK.SCH SPCUS**	6	core
404	n/a	noise	**4766**	n/a	noise	**8**	n/a	noise	***** ASF SS ZCUS.UK.SCH**	6	core
432	n/a	noise	**4767**	n/a	noise	**9**	n/a	noise	***** ASF ZCUS.UK.LETTER**	6	core
442	n/a	noise	**4768**	n/a	noise	**10**	n/a	noise	***** ASF ZCUS.UK.LETTER ZCUS.UK.SCH**	6	core
526	n/a	noise	**6674**	n/a	noise	**11**	n/a	noise	***** ASF ZCUS.UK.SCH BD**	6	core
770	n/a	noise	**8763**	n/a	noise	**12**	n/a	noise	***** BD**	6	core
775	n/a	noise	**8764**	n/a	noise	**13**	n/a	noise	***** BD CM NOTE**	6	core
793	n/a	noise	**6**	19	core	**299**	n/a	noise	***** BD UK.OE VH OE**	6	core
795	n/a	noise	**7**	19	core	**300**	n/a	noise	***** CM**	6	core
1	3	core	**8**	19	core	**301**	n/a	noise	***** CM PHA.ORDS**	6	core
6	6	core	**10**	19	core	**302**	n/a	noise	***** LAB.DRP**	6	core
14	6	core	**11**	19	core	**303**	n/a	noise	***** LAB.DRP UK.OE REC**	6	core
18	7	core	**12**	19	core	**304**	n/a	noise	***** MED**	6	core
28	0	core	**13**	19	core	**305**	n/a	noise	***** MED ASF**	6	core
31	0	core	**14**	19	core	**306**	n/a	noise	***** MED CM NOTE**	6	core
37	1	core	**15**	19	core	**307**	n/a	noise	***** MPI**	6	core
38	1	core	**17**	19	core	**308**	n/a	noise	***** MPI SPC**	6	core
47	6	core	**18**	19	core	**309**	n/a	noise	***** MPI ZCUS.UK.LETTER**	0	core
48	6	core	**19**	19	core	**310**	n/a	noise	***** MPI ZCUS.UK.SCH**	6	core
52	6	core	**20**	19	core	**311**	n/a	noise	***** NOTE**	6	core
55	6	core	**22**	19	core	**312**	n/a	noise	***** NOTE PHA.ORDS**	6	core
58	6	core	**24**	19	core	**313**	n/a	noise	***** NOTE SS**	6	core
60	6	core	**25**	19	core	**314**	n/a	noise	***** NOTE ZCUS.UK.LETTER**	6	core
64	6	core	**27**	19	core	**315**	n/a	noise	***** OE**	6	core
68	4	core	**29**	19	core	**409**	n/a	noise	***** OE MPI**	6	core
72	7	core	**30**	19	core	**410**	n/a	noise	***** OE REC REC:(DRP) UK.OE**	6	core

77	7	core	32	19	core	411	n/a	noise	*** OE ZCUS.UK.LETTER	6	core
80	7	core	34	19	core	412	n/a	noise	*** PHA.ORDS	0	core
82	7	core	36	19	core	413	n/a	noise	*** PHA.ORDS CM	6	core
85	7	core	39	19	core	414	n/a	noise	*** PHA.ORDS LAB.DRP	6	core
97	5	core	41	19	core	415	n/a	noise	*** PHA.ORDS REC REC:(DRP)	6	core
98	5	core	43	19	core	416	n/a	noise	*** PHA.ORDS REC REC:(DRP) OE UK.OE	6	core
100	5	core	44	19	core	417	n/a	noise	*** PHA.ORDS REC REC:(DRP) UK.OE	6	core
101	5	core	45	19	core	418	n/a	noise	*** RAD.DRP	6	core
102	5	core	46	19	core	419	n/a	noise	*** RAD.DRP LAB.DRP	6	core
107	0	core	48	19	core	420	n/a	noise	*** REC ASF BD	6	core
109	0	core	51	19	core	421	n/a	noise	*** REC OE	6	core
111	0	core	53	19	core	422	n/a	noise	*** REC OE UK.OE	6	core
115	0	core	56	19	core	423	n/a	noise	*** REC PHA.ORDS	6	core
118	9	core	57	19	core	424	n/a	noise	*** REC REC:(DRP)	6	core
120	1	core	58	19	core	425	n/a	noise	*** REC REC:(DRP) OE UK.OE	6	core
122	1	core	61	19	core	426	n/a	noise	*** REC REC:(DRP) OE UK.OE LAB.DRP	6	core
124	3	core	63	19	core	445	n/a	noise	*** REC REC:(DRP) PHA.ORDS LAB.DRP OE UK.OE	6	core
129	6	core	64	19	core	446	n/a	noise	*** REC REC:(DRP) UK.OE	6	core
133	4	core	66	19	core	447	n/a	noise	*** REC REC:(DRP) UK.OE OE	6	core
135	7	core	68	19	core	448	n/a	noise	*** REC UK.OE	6	core
136	7	core	69	19	core	449	n/a	noise	*** REC VH LAB.DRP	6	core
138	7	core	71	19	core	450	n/a	noise	*** SPC	6	core
141	7	core	73	19	core	451	n/a	noise	*** SPCUS ASF	6	core
149	5	core	74	19	core	452	n/a	noise	*** SS	6	core
152	5	core	75	19	core	474	n/a	noise	*** SS ASF ZCUS.UK.LETTER	6	core
154	5	core	76	19	core	808	n/a	noise	*** SS ZCUS.UK.SCH	2	core
158	0	core	77	19	core	809	n/a	noise	*** UK.OE	6	core
163	3	core	78	19	core	810	n/a	noise	*** UK.OE OE	6	core
168	6	core	79	19	core	811	n/a	noise	*** UK.OE PHA.ORDS	6	core
169	6	core	80	19	core	812	n/a	noise	*** UK.OE REC	6	core
170	6	core	81	19	core	813	n/a	noise	*** UK.OE REC REC:(DRP)	6	core
171	6	core	83	19	core	814	n/a	noise	*** UK.OE ZCUS.UK.LETTER	6	core
172	6	core	84	19	core	815	n/a	noise	*** UK.OE ZCUS.UK.SCH	1	core
173	6	core	85	19	core	816	n/a	noise	*** VH MPI ZCUS.UK.SCH OE	0	core
175	6	core	88	19	core	817	n/a	noise	*** VH UK.OE	6	core

177	6	core	**89**	19	core	**818**	n/a	noise	***** VH ZCUS.UK.LETTER**	6	core
178	6	core	**90**	19	core	**819**	n/a	noise	***** ZCUS.UK.LETTER ASF**	6	core
180	4	core	**93**	19	core	**820**	n/a	noise	***** ZCUS.UK.LETTER ASF VH**	6	core
181	7	core	**94**	19	core	**821**	n/a	noise	***** ZCUS.UK.LETTER NOTE ASF ZCUS.UK.SCH MPI**	6	core
186	7	core	**95**	19	core	**822**	n/a	noise	***** ZCUS.UK.LETTER SS**	6	core
187	7	core	**97**	19	core	**823**	n/a	noise	***** ZCUS.UK.LETTER UK.OE**	1	core
188	10	core	**98**	19	core	**824**	n/a	noise	***** ZCUS.UK.LETTER ZCUS.UK.SCH**	6	core
190	5	core	**99**	19	core	**825**	n/a	noise	***** ZCUS.UK.LETTER ZCUS.UK.SCH MPI**	6	core
192	5	core	**100**	19	core	**826**	n/a	noise	***** ZCUS.UK.LETTER ZCUS.UK.SCH SPC**	6	core
196	0	core	**101**	19	core	**827**	n/a	noise	***** ZCUS.UK.LETTER ZCUS.UK.SCH SS WL**	6	core
197	0	core	**103**	19	core	**947**	n/a	noise	***** ZCUS.UK.SCH**	6	core
199	9	core	**105**	19	core	**948**	n/a	noise	***** ZCUS.UK.SCH ASF SS ZCUS.UK.LETTER**	6	core
200	3	core	**106**	19	core	**949**	n/a	noise	***** ZCUS.UK.SCH ASF ZCUS.UK.LETTER**	6	core
204	8	core	**108**	19	core	**950**	n/a	noise	***** ZCUS.UK.SCH SS**	6	core
206	6	core	**110**	19	core	**951**	n/a	noise	***** ZCUS.UK.SCH UK.OE VH OE ZCUS.UK.LETTER**	6	core
209	6	core	**111**	19	core	**952**	n/a	noise	***** ZCUS.UK.SCH ZCUS.UK.LETTER SS**	6	core
211	4	core	**113**	19	core	**953**	n/a	noise	**ALLERGIES BD MPI**	2	core
215	7	core	**115**	19	core	**954**	n/a	noise	**ALLERGIES BD PHA.ORDS**	0	core
218	5	core	**116**	19	core	**955**	n/a	noise	**ALLERGIES CM**	6	core
219	5	core	**117**	19	core	**1054**	n/a	noise	**ALLERGIES MPI**	6	core
222	0	core	**118**	19	core	**1123**	n/a	noise	**ASF BD**	1	core
225	3	core	**119**	19	core	**458**	0	core	**ASF BD CM**	6	core

Table 44. Ensemble LOF Audit Logs

Date & Time	Device	Device Anomaly Score	User ID	User Anomaly Score	Routine ID	Routine Anomaly Score	Patient ID	Patient Anomaly Score	Duration (sec)	Ensemble Averaging Anomaly Score
26/09/16 17:02	1284	1.05	435	1.087	ASF SPC CAA MPI	13.339	71272	1.081	853	**4.139**
26/09/16 17:02	1284	1.050	435	1.087	ASF SPC CAA MPI	13.339	71272	1.081	853	**4.139**
25/11/16 03:39	102	1.084	1487	1.044	ASF SPC CAA MPI	13.339	29971	1.047	901	**4.129**
15/08/16 20:56	531	1.161	358	1.052	*** UK.OE MPI PHA.ORDS	11.643	23637	1.066	1180	**3.731**
21/11/16 21:46	369	1.088	1021	1.125	SPC SS ASF VH	11.350	41661	1.090	970	**3.663**
09/08/17 11:39	1537	1.123	77	1.048	SPC SS ASF VH	11.350	57108	1.030	1041	**3.638**
21/11/16 17:38	1052	1.094	809	1.087	SS ZCUS.UK.LETTER ZCUS.UK.SCH MPI	9.701	43065	1.054	723	**3.234**
01/04/16 01:12	49	1.151	117	1.031	SS ZCUS.UK.LETTER ZCUS.UK.SCH MPI	9.701	52200	1.028	861	**3.228**
19/12/16 20:03	293	1.067	992	1.090	REC REC:(DRP) PHA.MEDS UK.OE	9.538	41375	1.054	2454	**3.187**
07/02/17 00:18	566	1.164	262	1.074	ZCUS.UK.LETTER	1.084	35888	9.414	1182	**3.184**
27/12/16 18:50	293	1.067	992	1.090	REC REC:(DRP) PHA.MEDS UK.OE	9.538	46862	1.020	1691	**3.179**
26/07/16 23:36	594	1.141	262	1.074	ZCUS.UK.LETTER	1.084	35888	9.414	1342	**3.178**
04/04/17 23:56	161	1.123	639	1.122	ASF MPI VH BD	9.287	31633	1.033	890	**3.141**
11/08/16 15:49	1064	1.084	36	1.051	ASF MPI VH BD	9.287	8178	1.060	964	**3.121**
08/11/16 22:37	566	1.164	262	1.074	ZCUS.UK.LETTER	1.084	19327	8.814	2352	**3.034**
29/09/16 18:44	1050	1.153	262	1.074	ZCUS.UK.LETTER	1.084	19327	8.814	2845	**3.031**
13/07/16 01:56	49	1.151	117	1.031	ZCUS.UK.LETTER PHA.ORDS MPI	8.812	22427	1.030	792	**3.006**
05/05/17 23:32	2138	1.020	472	1.042	ZCUS.UK.LETTER PHA.ORDS MPI	8.812	39506	1.044	944	**2.979**
27/07/17 15:46	286	1.071	809	1.087	ZCUS.UK.LETTER	1.084	69053	8.552	4251	**2.949**

Table 45. User ID

UserID	UserID Mean Density	UserID Mean Anomaly Score	UserID Mean Neighbourhood Radius
1	1043.798	1.027673	0.002242
2	410.4643	1.149913	0.005742
3	1107.154	1.02388	0.002242
4	411.4143	1.097459	0.004856
5	1404.795	1.027075	0.000934
6	1992.614	1.298183	0.001338
7	1059.464	1.024369	0.00188
8	966.7279	1.061966	0.00278
9	1836.34	1.058019	0.001218
10	20.60531	1.335073	0.115581
11	829.4562	1.036223	0.002875
12	280.8375	1.117352	0.009614
13	1045.505	1.095059	0.002041
14	505.6614	1.097874	0.003836
15	1048.693	1.122646	0.002028
16	503.9542	1.148246	0.008533
17	1132.574	1.095685	0.00174
18	324.0562	1.093005	0.006633
19	1281.272	1.030559	0.001436
20	574.2201	1.105305	0.005684
21	341.9207	1.196402	0.004584
22	358.5714	1.040309	0.007083
23	830.0152	1.285939	0.001662
24	680.3369	1.127738	0.007377
25	199.4606	1.40477	0.009852
26	1604.986	1.093321	0.000747
27	927.0253	1.053835	0.001627
28	729.0049	1.11716	0.013151
29	959.3541	1.05737	0.002119
30	515.686	1.242591	0.030363
31	265.017	1.1775	0.033147
32	884.469	1.072201	0.002139
33	1252.326	1.054169	0.001652
34	293.599	1.05082	0.00701
35	886.6879	1.10547	0.005313
36	583.6535	1.05124	0.003058
37	995.0664	1.030046	0.002155
38	451.1214	1.093259	0.006491
39	2110.942	1.152689	0.001599
40	347.9017	1.139919	0.008658
41	578.6834	1.117277	0.016201
42	1313.021	1.046276	0.001184
43	1654.547	1.113666	0.002374
44	1361.999	1.039389	0.001183
45	975.0377	1.057005	0.001299
46	840.3081	1.086663	0.002689
47	770.5982	1.07576	0.011974
48	1775.245	1.19275	0.000724
49	770.1085	1.068873	0.003674
50	1233.655	1.089465	0.002676
51	40.51377	1.894182	0.043856
52	1485.645	1.03328	0.001091
53	263.6294	1.136582	0.00846
54	220.6689	1.06799	0.016424
55	534.9696	1.221609	0.023388
56	1197.878	1.070044	0.00122
57	125.9688	1.712649	0.014847
58	2537.062	1.057756	0.000941
59	422.9958	1.04958	0.004849
60	493.3177	1.078579	0.0024
61	247.2395	1.146779	0.006096
62	163.9963	1.163406	0.011082
63	661.6243	1.036495	0.002593
64	1237.226	1.070512	0.001099
65	1055.389	1.053929	0.003331
66	783.5505	1.066146	0.002603
67	546.5981	1.108777	0.005848
68	1349.482	1.069858	0.001467
69	733.4366	1.277643	0.012178
70	1069.543	1.103969	0.001752
71	606.9516	1.088872	0.006271
72	982.9848	1.139018	0.001601
73	864.7124	1.04592	0.00513
74	1241.164	1.049223	0.001905
75	961.1611	1.061536	0.002025
76	943.9458	1.074482	0.002206
77	591.93	1.048351	0.004095
78	1141.941	1.068552	0.001406
79	276.7341	1.143636	0.005377
80	1928.322	1.1276	0.001246
81	165.3009	1.163549	0.027191
82	1389.1	1.041385	0.002741
83	139.0991	1.312776	0.023174
84	967.3163	1.048901	0.006339
85	1115.1	1.027863	0.00225

Table 46. Device ID

DeviceID	DeviceID Mean Density	DeviceID Mean Anomaly Score	DeviceID Mean Neighbourhood Radius
1	2858.934	1.033719	0.000386
2	2656.22	1.092214	0.000775
3	2150.651	1.032036	0.000921
4	2955.04	1.123626	0.003081
5	1397.68	1.13256	0.001655
6	1319.408	1.093516	0.00808
7	1991.14	1.060952	0.000812
8	1624.906	1.102338	0.00127
9	1249.741	1.180543	0.001222
10	188.7924	1.278305	0.008672
11	3044.181	1.095267	0.001619
12	2287.773	1.089656	0.001209
13	3584.672	1.092371	0.001078
14	2804.52	1.077334	0.000932
15	1078.192	1.089531	0.004487
16	1666.605	1.097122	0.003032
17	2052.178	1.08645	0.002593
18	950.1689	0.960641	0.001285
19	2161.774	1.079007	0.001348
20	1855.666	1.136736	0.003641
21	1102.481	1.051255	0.002067
22	1003.194	1.152335	0.002774
23	2565.335	1.064898	0.001183
24	3523.11	1.226143	0.003727
25	391.3256	1.127998	0.003428
26	2184.298	1.090676	0.004121
27	868.1916	1.172174	0.001668
28	1134.046	1.094787	0.002864
29	686.7189	1.144315	0.004906
30	1864.846	1.072197	0.003751
31	9981.286	1.034232	0.000824
32	2300.671	1.044591	0.00086
33	2457.688	1.119069	0.000925
34	2304.553	1.10126	0.000865
35	1762.515	1.060756	0.001135
36	2334.817	1.064558	0.003101
37	1716.771	1.08667	0.001499
38	1541.66	1.100596	0.001948
39	1598.653	1.134652	0.002227
40	2054.984	1.149397	0.00096
41	2021.987	1.051669	0.00204
42	2071.922	1.046377	0.001919
43	3995.073	1.092616	0.002453
44	2676.568	1.06876	0.00146
45	2754.871	1.026911	0.000711
46	2893.927	1.09326	0.000888
47	228.7059	1.025623	0.005027
48	1064.573	1.054493	0.003412
49	873.4354	1.151319	0.005191
50	5115.861	1.0533	0.000812
51	2625.263	1.078016	0.000806
52	997.9633	1.037486	0.003188
53	1523.635	1.037883	0.001758
54	1629.545	1.176478	0.001324
55	1072.677	1.289804	0.001482
56	1753.89	1.05992	0.002427
57	977.9016	1.154851	0.006188
58	1771.538	1.03749	0.001974
59	1816.97	1.021814	0.002063
60	1332.155	1.126619	0.002195
61	2317.439	1.068483	0.000771
62	2000.792	1.152663	0.00182
63	1639.034	1.149167	0.001633
64	1850.995	1.142319	0.001103
65	5158.694	1.049373	0.000844
66	431.7613	1.188906	0.003057
67	1026.989	1.089064	0.001604
68	1306.299	1.095637	0.00148
69	497.6144	1.065403	0.003154
70	1156.222	1.068798	0.001504
71	2157.89	1.092324	0.001202
72	2553.632	1.117734	0.001032
73	655.1366	1.232578	0.007781
74	638.1413	1.250553	0.007538
75	198.5485	1.197098	0.006633
76	261.9671	1.254643	0.009892
77	1725.735	1.09643	0.00083
78	2233.137	1.073705	0.000911
79	2769.067	1.077599	0.001118
80	1962.593	1.068115	0.002264
81	2300.771	1.065968	0.00193
82	2406.798	1.412366	0.000881
83	2145.968	1.116279	0.001223
84	3261.986	1.056026	0.001349
85	1440.747	1.098618	0.001336

Table 47. Patient ID

PatientID	PatientID Mean Density	PatientID Mean Anomaly Score	PatientID Mean Neighbourhood Radius
1	25146.61	1.008021	9.27E-05
2	21847.27	1.099227	9.31E-05
3	11924.72	1.0324	6.28E-05
4	12157.76	1.06249	9.50E-05
5	6467.031	1.042467	0.000184
6	1562.59	1.060335	0.001455
7	4119.074	1.046545	0.000569
8	13032.63	1.042264	0.000191
9	8288.565	1.071476	0.000225
10	6273.409	1.000661	0.000186
11	2	1	0
12	6672.44	1.081824	0.000505
13	17177.09	1.025434	5.47E-05
14	1479.414	1.040299	0.000842
15	24071.95	1.277255	4.46E-05
16	564.8539	1.108071	0.004238
17	27537.06	1.035388	1.63E-05
18	2101.907	1.243641	0.000871
19	25286.69	1.074709	0.000136
20	2	1	0
21	8825.066	1.058902	0.000604
22	2	1	0
23	7461.91	1.017446	0.000171
24	11821.52	1.034918	6.56E-05
25	49811.44	1.102173	2.22E-05
26	2	1	0
27	10548.69	1.06291	0.000183
28	9172.058	1.032271	0.000148
29	28969.53	1.115197	2.29E-05
30	9239.153	1.067164	0.000117
31	2	1	0
32	33095.88	1.16714	2.42E-05
33	2	1	0
34	2	1	0
35	10231.15	1.101179	0.000263
36	11676.7	1.061805	0.00028
37	26464.88	1.04918	2.52E-05
38	2	1	0
39	16190.96	1.091126	9.84E-05
40	19094.33	1.053864	9.41E-05
41	9034.347	1.060244	0.00071
42	2	1	0
43	16629.93	1.087424	0.000339
44	2	1	0
45	2	1	0
46	15820.67	1.024007	0.000184
47	27131.21	1.030738	7.12E-05
48	2093.142	1.129269	0.001105
49	2	1	0
50	20083.65	1.025113	4.45E-05
51	19823.76	1.067118	0.000188
52	2	1	0
53	2	1	0
54	2	1	0
55	44288.79	1.151934	1.76E-05
56	13189.62	1.071704	0.000123
57	16081.07	1.092298	6.68E-05
58	17222.06	1.067191	7.27E-05
59	10909.62	1.041247	0.000615
60	9586.836	1.020053	0.000191
61	9378.382	1.032494	0.000228
62	2	1	0
63	2	1	0
64	5456.331	1.049474	0.000316
65	17417	1.057319	9.95E-05
66	1083.232	1.048929	0.001406
67	2	1	0
68	2	1	0
69	2	1	0
70	19847.73	1.06386	7.74E-05
71	15548.4	1.027631	9.92E-05
72	18222.32	1.007681	8.38E-05
73	2	1	0
74	6762.828	1.006501	0.000152
75	7700.662	1.048497	0.000131
76	4224.482	1.05453	0.00024
77	2	1	0
78	15024.2	1.089634	6.82E-05
79	15558.53	1.140467	3.62E-05
80	22075.27	1.048835	3.53E-05
81	19533.65	1.040046	6.69E-05
82	2	1	0
83	6830.069	1.041488	0.000866
84	5768.007	0.999322	0.00013
85	14497.44	1.020985	4.89E-05

Table 48. Routine ID

RoutineID	RoutineID Mean Density	RoutineID Mean Anomaly Score	RoutineID Mean Neighbourhood Radius
***	2039.149	1.34573	0.001508
*** ALLERGIES MPI ZCUS.UK.LETTER	1912.238	0.909873	0.000601
*** ALLERGIES PHA.ORDS	7166.723	1.019241	0.000342
*** ASF	2009.382	1.117673	0.00137
*** ASF BD	2158.983	1.048254	0.000939
*** ASF CAA MPI	2973.4	1.021882	0.000735
*** ASF CM	540.1888	1.022932	0.000223
*** ASF CM PHA.ORDS	596.8578	1.005121	0.000197
*** ASF MED NOTE ZCUS.UK.LETTER	1774.526	1.036549	0.000595
*** ASF MPI	1728.243	1.062767	0.001111
*** ASF MPI ZCUS.UK.LETTER	585.1347	1.007937	0.000555
*** ASF NOTE	650.1179	1.006881	0.000219
*** ASF NOTE MED ZCUS.UK.LETTER	2	1	0
*** ASF NOTE SS PHA.ORDS	2	1	0
*** ASF NOTE ZCUS.UK.LETTER	1981.007	1.0287	0.001768
*** ASF NPC	3235.509	1.059358	0.000623
*** ASF NPC MED SS	2	1	0
*** ASF NPC SPC	2297.035	1.062682	0.000672
*** ASF OBSDIRC.DRP SPCUS ZCUS.UK.SCH	2	1	0
*** ASF OE	3214.874	1.054969	0.000528
*** ASF OE NPC NOTE SPC ZCUS.UK.SCH	2	1	0
*** ASF OE SPC	732.3954	1.031354	0.000802
*** ASF OE VH SPC NPC	2	1	0
*** ASF PHA.ORDS	2104.9	1.051077	0.004452
*** ASF PHY.QRY ZCUS.UK.SCH ZCUS.UK.LETTER	2	1	0
*** ASF RAD ZCUS.UK.LETTER	512.5214	1.013389	0.000658
*** ASF RAD.DRP	1059.255	1.214022	0.017684
*** ASF SPC	1816.373	1.053864	0.000689
*** ASF SPC MED PHA.MEDS PHA.ORDS	1747.807	1.195802	0.000604
*** ASF SPC PHA.ORDS	4769.808	1.024346	0.000472
*** ASF SPC VH ZCUS.UK.SCH SPCUS	2	1	0
*** ASF SPCUS	1029.991	1.011576	0.000181
*** ASF SPCUS ZCUS.UK.SCH	6344.499	1.004587	0.001992
*** ASF SS	2584.86	1.198148	0.000813
*** ASF SS NOTE	460.9117	1.033476	0.000407
*** ASF SS SPC PHA.ORDS	708.8701	1.011379	0.000184
*** ASF SS SPCUS ZCUS.UK.SCH RAD.DRP MPI OE	2	1	0
*** ASF SS WL	677.3306	1.006596	0.000223
*** ASF SS ZCUS.UK.LETTER	347.9728	1.030623	0.000349
*** ASF SS ZCUS.UK.SCH	1441.508	1.00779	0.000428
*** ASF SS ZCUS.UK.SCH MPI UK.OE	1767.674	1.007101	0.00059
*** ASF UK.OE PHA.ORDS	115.7811	1.071519	0.000656
*** ASF UK.OE ZCUS.UK.SCH SPC	2	1	0
*** ASF VH	3291.566	1.085844	0.000455
*** ASF VH OE	3067.544	0.99938	0.000221
*** ASF VH SPC NPC	2	1	0
*** ASF VH SS	516.0939	1.019133	0.001241
*** ASF VH UK.OE	4678.902	1.174131	0.006304
*** ASF VH ZCUS.UK.LETTER	594.8677	0.996964	0.000168
*** ASF VH ZCUS.UK.SCH	1918.062	1.063186	0.000912
*** ASF WL	7655.482	1.046246	0.000371
*** ASF WL SS	716.3716	1.009277	0.000191
*** ASF ZCUS.UK.LETTER	2698.567	1.025048	0.001191
*** ASF ZCUS.UK.LETTER OE	93.79351	1.024572	0.000437
*** ASF ZCUS.UK.LETTER PHA.ORDS	242.2059	1.006914	0.000147
*** ASF ZCUS.UK.LETTER SPCUS SPC VH	2	1	0
*** ASF ZCUS.UK.LETTER SS	1927.678	5.746574	0.126766
*** ASF ZCUS.UK.LETTER ZCUS.UK.SCH	1561.53	1.052312	0.001623
*** ASF ZCUS.UK.SCH	2531.265	1.047235	0.001129
*** ASF ZCUS.UK.SCH BD	1501.992	1.005335	0.000266
*** ASF ZCUS.UK.SCH MPI ZCUS.UK.LETTER	2	1	0
*** ASF ZCUS.UK.SCH OE	552.4111	0.998563	0.000199
*** ASF ZCUS.UK.SCH RAD.DRP OE SS WL	2	1	0
*** ASF ZCUS.UK.SCH ZCUS.UK.LETTER	2236.572	1.046826	0.001414
*** ASF ZCUS.UK.SCH ZCUS.UK.LETTER MPI	2	1	0
*** BD	1431.465	1.095381	0.001595
*** BD ASF	2063.547	1.14231	0.000824

Table 49. Top 100 Anomalous Audit Logs

Date & Time (May 2017)	Ensemble Average	Date & Time (Jul16-Dec16)	Ensemble Average	Date & Time (Feb16-Aug17)	Ensemble Average
17/05/16 02:46	3.180	16/07/15 22:10	5.136	16/09/26 17:02	4.139
17/05/16 01:40	3.161	16/12/07 20:17	3.535	16/11/25 03:39	4.129
17/05/26 03:20	2.894	16/11/21 13:55	3.535	16/08/15 20:56	3.731
17/05/26 15:48	2.894	16/09/09 17:35	3.300	16/11/21 21:46	3.663
17/05/01 13:33	2.875	16/10/07 15:45	3.201	17/08/09 11:39	3.638
17/05/08 20:48	2.870	16/07/08 22:01	3.180	16/11/21 17:38	3.234
17/05/19 23:50	2.817	16/07/13 06:01	3.159	16/04/01 01:12	3.228
17/05/30 17:20	2.777	16/12/29 17:27	3.157	16/12/19 20:03	3.187
17/05/17 15:59	2.660	16/10/11 20:00	3.128	17/02/07 00:18	3.184
17/05/25 20:34	2.659	16/10/11 01:17	3.118	16/12/27 18:50	3.179
17/05/11 20:22	2.652	16/11/17 19:05	3.100	16/07/26 23:36	3.178
17/05/11 21:49	2.651	16/11/08 22:37	3.099	17/04/04 23:56	3.141
17/05/15 21:44	2.650	16/10/06 09:49	3.099	16/08/11 15:49	3.121
17/05/24 15:55	2.650	16/09/28 01:24	3.090	16/11/08 22:37	3.034
17/05/09 01:35	2.646	16/10/25 02:46	3.077	16/09/29 18:44	3.031
17/05/03 14:14	2.641	16/09/29 18:44	3.058	16/07/13 01:56	3.006
17/05/10 17:30	2.638	16/12/05 19:19	3.043	17/05/05 23:32	2.979
17/05/24 03:04	2.636	16/10/20 18:35	3.043	17/07/27 15:46	2.949
17/05/24 04:22	2.636	16/10/04 18:32	2.899	17/06/20 18:43	2.945
17/05/03 15:34	2.636	16/10/03 19:06	2.892	17/01/18 01:57	2.943
17/05/22 18:50	2.634	16/08/24 17:34	2.866	16/03/01 16:46	2.936
17/05/10 17:01	2.625	16/12/20 18:49	2.859	17/08/11 17:06	2.876
17/05/23 15:57	2.624	16/12/22 14:00	2.842	17/07/31 20:58	2.870
17/05/11 00:08	2.622	16/10/24 23:58	2.837	17/01/16 23:28	2.870
17/05/23 12:39	2.581	16/08/24 13:41	2.772	17/04/25 21:47	2.842
17/05/10 20:06	2.542	16/07/19 19:06	2.756	16/06/21 10:38	2.838
17/05/08 02:59	2.524	16/09/24 05:08	2.753	16/06/03 01:16	2.824
17/05/02 23:48	2.516	16/12/30 15:17	2.753	16/04/14 19:18	2.811
17/05/09 02:53	2.478	16/09/04 23:52	2.753	16/04/29 15:14	2.773
17/05/17 04:34	2.475	16/08/24 23:40	2.707	17/01/17 01:42	2.772
17/05/15 16:41	2.467	16/10/17 19:02	2.702	17/01/04 22:29	2.760
17/05/04 18:20	2.466	16/10/31 01:03	2.691	16/05/27 18:40	2.721
17/05/03 16:43	2.461	16/12/19 20:11	2.691	16/07/13 01:37	2.720
17/05/04 15:45	2.455	16/08/24 01:01	2.676	17/05/09 21:49	2.715
17/05/08 17:28	2.446	16/11/26 16:42	2.582	16/09/14 02:52	2.711
17/05/22 20:55	2.445	16/10/31 11:54	2.566	16/08/25 18:54	2.710
17/05/05 06:29	2.441	16/11/26 16:49	2.518	16/06/17 15:47	2.707
17/05/25 01:57	2.439	16/09/27 18:31	2.512	17/01/11 02:40	2.703
17/05/15 23:38	2.434	16/10/18 20:39	2.498	17/07/15 03:25	2.696
17/05/15 18:29	2.433	16/08/17 13:10	2.452	16/11/01 19:13	2.694
17/05/10 15:36	2.432	16/08/05 17:04	2.451	17/02/28 18:36	2.689
17/05/17 19:14	2.412	16/08/16 21:56	2.432	16/10/21 15:28	2.670
17/05/25 01:02	2.394	16/12/21 01:35	2.426	17/07/15 07:54	2.666
17/05/05 01:45	2.388	16/09/01 21:44	2.417	16/12/05 19:19	2.646
17/05/18 03:17	2.379	16/11/30 23:58	2.369	16/10/20 18:35	2.645
17/05/11 03:22	2.358	16/07/11 17:14	2.363	17/06/22 00:00	2.620
17/05/02 17:15	2.347	16/11/09 16:51	2.352	17/06/23 14:11	2.620
17/05/02 00:54	2.316	16/11/04 18:43	2.343	16/10/03 00:06	2.620
17/05/22 19:17	2.305	16/07/01 18:30	2.343	16/10/06 19:04	2.610
17/05/09 07:42	2.295	16/11/08 01:25	2.329	17/03/02 15:33	2.600
17/05/25 17:13	2.293	16/10/31 17:17	2.293	17/06/15 16:55	2.596
17/05/17 23:59	2.292	16/09/05 17:19	2.268	16/04/21 06:56	2.576
17/05/02 03:34	2.291	16/10/21 17:38	2.218	17/07/18 22:04	2.571
17/05/17 03:28	2.288	16/11/08 15:07	2.210	16/04/19 01:47	2.569
17/05/31 19:06	2.274	16/11/08 15:17	2.210	17/06/08 15:36	2.569
17/05/02 16:53	2.268	16/08/26 21:55	2.173	16/09/01 22:36	2.567
17/05/24 22:24	2.253	16/09/26 22:31	2.161	16/08/11 01:24	2.565
17/05/18 23:25	2.247	16/09/21 03:03	2.151	16/10/13 06:56	2.565

Date & Time (May 2017)	Ensemble Average	Date & Time (Jul16-Dec16)	Ensemble Average	Date & Time (Feb16-Aug17)	Ensemble Average
17/05/18 23:33	2.247	16/10/05 15:49	2.149	16/10/13 08:24	2.565
17/05/23 14:09	2.246	16/09/29 01:42	2.142	17/02/16 06:50	2.565
17/05/25 20:46	2.243	16/12/15 01:33	2.141	17/01/25 07:59	2.561
17/05/17 14:14	2.243	16/10/21 15:28	2.129	17/06/29 15:59	2.560
17/05/22 18:55	2.228	16/11/01 19:13	2.115	16/04/21 06:55	2.557
17/05/22 19:04	2.228	16/11/08 04:51	2.097	16/08/11 01:20	2.557
17/05/22 19:06	2.223	16/12/27 18:50	2.095	16/08/11 01:35	2.557
17/05/04 23:24	2.218	16/12/19 20:03	2.090	16/10/14 00:03	2.557
17/05/09 22:36	2.197	16/11/08 17:38	2.087	17/08/21 22:19	2.527
17/05/18 02:51	2.194	16/09/27 02:42	2.040	16/04/04 01:01	2.514
17/05/15 23:25	2.192	16/10/27 00:10	2.030	17/07/24 21:54	2.489
17/05/15 01:34	2.190	16/10/07 15:53	2.027	16/09/13 01:31	2.432
17/05/09 04:56	2.187	16/09/12 22:00	2.024	16/03/29 13:31	2.426
17/05/24 23:38	2.187	16/09/12 22:16	2.022	16/03/07 09:39	2.416
17/05/05 14:08	2.186	16/07/08 22:26	2.020	16/03/08 12:20	2.413

Date & Time (May 2017)	Ensemble Average	Date & Time (Jul16-Dec16)	Ensemble Average	Date & Time (Feb16-Aug17)	Ensemble Average
17/05/18 03:43	2.183	16/08/23 06:42	2.018	17/03/14 01:54	2.410
17/05/25 22:35	2.175	16/07/21 22:21	2.017	16/09/28 20:49	2.380
17/05/09 10:07	2.175	16/09/12 14:19	2.014	17/07/25 01:37	2.378
17/05/18 04:53	2.174	16/07/28 01:43	2.011	16/11/26 16:42	2.368
17/05/16 02:55	2.163	16/07/04 15:53	2.009	17/04/19 15:33	2.363
17/05/17 01:38	2.161	16/07/20 04:30	2.009	17/04/19 20:21	2.363
17/05/16 20:55	2.160	16/08/05 21:43	2.000	16/11/26 16:49	2.360
17/05/11 18:39	2.155	16/07/08 16:46	1.998	16/08/23 06:42	2.353
17/05/09 17:34	2.151	16/09/13 14:06	1.995	16/09/27 02:42	2.346
17/05/25 02:59	2.150	16/09/13 14:07	1.995	16/11/21 13:55	2.331
17/05/23 02:36	2.149	16/08/10 00:13	1.988	17/06/19 01:52	2.327
17/05/26 18:25	2.148	16/08/01 20:55	1.983	16/12/07 20:17	2.327
17/05/21 14:34	2.147	16/09/07 01:39	1.982	17/05/26 17:34	2.315
17/05/08 03:04	2.146	16/08/22 15:04	1.982	16/11/22 16:45	2.306