Anomaly Detection Beyond the Research Setting

An exploration of the use of statistics and machine learning to detect cyber attacks

G.D. Sæmundsson
Faculty of Technology, Policy & Management
Delft University of Technology
Jaffalaan 5, 2628 BX Delft, The Netherlands
Email: g.d.saemundsson@student.tudelft.nl

***Abstract*—In this paper we approach the problem of deploying anomaly detection techniques for detecting cyber attacks in an organisational environment. Anomaly detection has been an active research area for almost three decades with promising results. However, few such systems have been successfully implemented in an operational environment for improving cyber security. Researchers have attempted to identify the reasons for this gap between research and operational success, and provide guidelines on how to overcome it. In this work we use these guidelines to guide us in the exploration of how business organisations approach anomaly detection. We compare the insights from practice with theory in an effort to better understand the main discrepancies between the two settings.**

Keywords

Anomaly detection, cyber security, cyber attacks, intrusion detection, usability, organisational challenges, technical challenges

I. Introduction

Governments and businesses around the world operate complex and interconnected information systems and networks. In recent years organisations have made their networks increasingly open to the outside world. For instance, many organisations allow partner organisations to access parts of their inner network, customers to directly interact with a company's databases in e-commerce transactions, or allow employees to access their private networks from home. This increase in external access has made today's networks more susceptible to attacks [1].

In the past, cyber attacks were generally perceived to be the work of the stereotypical lone hacker. In recent times they're increasingly observed to be the work of disgruntled employees, hacktivists, vandals and script kiddies, organised crime, terrorist organisations, and state actors [2], [3].

These developments, along with increasingly sophisticated methods to attack being more accessible, are considered some of the reasons why cyber attacks are on the rise [1]. In recent times there have been several highly publicised attacks, e.g. when an unknown group of hackers reportedly stole $300 million from banks [4], published private pictures of celebrities after hacking cloud services [5] and stole sensitive personal data on over 22 million U.S. government employees [6].

Most of the methods and systems currently in use for detecting cyber attacks are rule-based, i.e. pre-existing knowledge is used to define monitoring rules. For example, a rule might generate an alert if the system observes 5 failed login attempts for a given user within a period of 30 minutes. With a well-defined set of rules, the main advantages of such systems is reliable detection and low false alarm rate. However, the main disadvantage of such systems is that without the relevant monitoring rules they can not detect unknown or novel attacks [1], [7].

Anomaly detection refers to finding patterns or instances in data that do not conform to what is normal and expected, i.e. anomalies are rare and different from the norm. Anomalous patterns found in organisational system and application data can reveal attacks against an organisation without the need for pre-existing rules for each attack [8].

Anomaly detection for the detection of cyber attacks has been extensively researched by academia since it was originally proposed in 1987 [9]. However, few such systems have been successfully implemented in an operational environment for improving cyber security [7], [10]. One of the reasons for limited operational success is that technologies failed to provide the environment needed to do anomaly detection. For a long time it was not economically feasible to store and retain the needed (and vast) amounts of data and build the computational capabilities needed. With the emergence of new tools and technologies that are capable of handling and analysing large amounts of data (e.g. distributed technologies like Hadoop and Mapreduce) it is becoming increasingly feasible to do anomaly detection on a large scale [11].

In this work we approach this problem from an organisational perspective and try to identify promising practices (e.g. regarding data choices, output expectations, performance requirements) for deploying anomaly detection for cyber security purposes in a business environment.

While anomaly detection is used in other problem domains (e.g. detecting anomalies in medical imagery and damage in industrial machinery [8]) we use the term 'anomaly detection' to refer to its application in cyber security.

II. Research methodology

Anomaly detection has promising applications for cyber security. Mainly, for detecting unknown (or undefined) attacks that traditional rule-based detection can not do [12], [13]. However, there are many practical challenges that apply to this problem in particular. For example, a high rate of false positives, difficulties with modelling normal activity, and costly evaluation [7], [10]. In addition, organisational issues play an important role in the preparation, execution, and success of anomaly detection projects. Gaining access the necessary access to data and experts, and privacy and legal concerns are organisational issues that commonly affect anomaly detection projects [7], [14], [15], [16], [8], [17] Furthermore, the combination of a high false positive rates and alerts that are generally less interpretable than alerts from rule-based systems make usability especially challenging [7], [10], [18].

Some research exists (e.g. [7], [10]) where researchers provide a set of guidelines, i.e. best practices, that are designed to improve the approach of researchers or others that want to deploy operationally sound and effective anomaly detection techniques for detecting cyber attacks. These are general guidelines on how to avoid common pitfalls and address potential problems when applying anomaly detection. However, the guidelines are still too abstract to be directly applied for addressing specific cyber security tasks (e.g. attack detection, profiling user groups, traffic classification) in business environments. Moreover, the given guidelines are general, i.e. they do not have specific types of organisations (e.g. size, industry), tasks (e.g. classification, outlier detection), or resources (e.g. data sources, human expertise) in mind. Therefore, it is a challenging task for an organisation to create its own pathway for deploying a usable anomaly detection approach.

As mentioned before anomaly detection has been extensively researched for almost three decades while it has seen relatively little operational success. We argue that we have to better understand this gap in order to move research and practice towards more operational and usable anomaly detection.

Based on the information, we have defined scientific research questions that will answered in this article:

Main research question:

> **What are the core discrepancies between theoretical guidelines and operational approaches when using anomaly detection in business organisations?**

Research questions:

- **RQ 1** - What is the state of the art of research on anomaly detection for detecting cyber attacks?
- **RQ 2** - What methodology for anomaly detection does academic research propose for business organisations?
- **RQ 3** - How do business organisations approach anomaly detection?
- **RQ 4** - How does a theoretical methodology compare to practical approaches in business organisations?

This research follows an *exploratory* approach. As is common in exploratory research, this work uses a literature review is to form propositions [19]. The propositions are the main findings of the literature review that form the building blocks of the theoretical methodology for doing anomaly detection. These are provisional tools for advancing the research with the aim of leading to the discovery of new insights or facts. Thereafter evidence is collected that may or may not support the the propositions [19].

In this work we collect evidence in a case study at two business organisations, a financial institution (**FI**) and an internet service provider (**NET**), that are doing anomaly detection in practice. By means of interviews and practical work in these organisations we explore how they approach the problem of making anomaly detection operational. Moreover, we explore how well the propositions of academic theory hold in their intended environment.

III. Theoretical foundation

A. State of the art

We took a broad look at the research developments and promising applications of anomaly detection for cyber attacks. We now highlight a few conclusions. Firstly, anomaly detection has the potential of detecting attacks without the prior knowledge needed for constructing conventional detection rules [12], [13]. Moreover, the knowledge gained from using these methods can be used for defining and refining conventional detection rules [13], [20], [21]. Secondly, unsupervised (and semi-supervised) machine learning techniques seem to be the most applicable since labelled training data is often unavailable and hard to obtain [8], [22].

B. Technical challenges

We explored the challenges of operationalising anomaly detection. We looked at the challenges from two different perspectives. Firstly, they lie in assumptions that are commonly made in anomaly detection research. However, these assumptions may not hold for anomaly detection an operational environment. Secondly, in unique problems of using machine learning techniques for detecting cyber attacks.

The challenges discussed cover a wide range of topics. For instance, that not all attacks are anomalies, and anomalies are not necessarily malicious [4], [7], [10], [23], [24], [25], [26], [27]. Moreover, real-life network and system activity is much more irregular and unpredictable than most people expect [7]. A high variability in the short term and on the long run leads to difficulties in defining normal behaviour needed for detecting anomalies [7], [10]. Furthermore, the high cost of both false positives and false negatives make it hard to tune the model. Evaluation is also costly as it can take a long time to verify each alert produced by anomaly detectors . Moreover, the evaluation normally has to be done by expensive and busy experts[7], [10], [18].

Research on anomaly detection sometimes overlooks these issues important for operational success [7], [10]. For example, how to cope with constantly changing notion of normal activity, or whether the alerts produced are manageable and useful for the experts that verify them.

C. Organisational challenges

We explored the organisational aspect of anomaly detection (and data mining) projects. A good understanding of the problem from a business perspective and translating this understanding into concrete parts of the project definition is crucial [7], [28], [29], [30], [14]. Other success factors include data quality, organisational support and integration of results into business processes [29], [30]. Moreover, smaller anomaly detection projects (scope, resource dependency, time) tend to be more successful [29]. In addition, we outline human and organisational issues that can hinder the progress of anomaly detection projects. For instance, common problems with access to necessary data, documentation, and experts with relevant domain knowledge [14], [15], [16]. Lastly, we explore how issues such as privacy and legal concerns, and security policies written in legal language can affect anomaly detection [7], [8], [10], [15], [16], [17].

D. Usability

We took a closer look at two important usability issues of anomaly detection in practice: a manageable amount of false positives, and alerts that convey meaning to the end-user (i.e. actionable alerts). These issues are important as we have seen in the literature looking at unique challenges of anomaly detection for in an operational environment. According to papers discussing actionable alerts we see that flexible presentation (e.g. visualisation, reports) of information and facilitation further investigation are important factors [31], [32], [33], [34], [35]. To achieve a manageable number of alerts we found that a very low rate of false alarms is needed to maintain users' trust in the results. In addition, we see researchers stating that a low false positive rate is of greater importance than a high true positive rate [36], [37]. Lastly, we identified best practices for working with usability: In early phases, use relevant expert-based usability knowledge to guide the process of deploying anomaly detectors. In later phases, verify usability in practice with end users and problem owners [38], [39], [40].

IV. Core findings

A. Theoretical methodology

Based on literature we constructed a theoretical methodology (summarised in Figure 1) using CRISP-DM [28] as underlying structure.

CRISP-DM helps us build a methodology that covers the whole life-cycle of organisational anomaly detection project, just as CRISP-DM is designed to do for data mining. We divide the phases of the framework into three 'bins' each holds five key propositions from the literature.

The theoretical methodology is revisited in the following sections where we present some the core findings of this work of testing the theoretical methodology by comparing it to practice.

B. Testing the theoretical methodology

In this section we present an overview of the core findings of the case studies by linking the data from the case studies to test the theoretical methodology.

In Table I we provide an overview of the findings where for each proposition we give an indication of whether the data collected during the case studies support (+), strongly support (+ +), contradict (−), or strongly contradict (− −) the propositions, or whether the interviews gave an alternative perspective ($\star$) or no answer/data (n.a.) was provided/available.

In the following subsections discuss the propositions partially supported or rejected based on the findings of the case study.

1) Phases 1-2: Business understanding and Data understanding:

P1 Business organisations are motivated to deploy anomaly detection for detecting new/unknown attacks and for automatic refinement of detection rules.[12], [13], [20], [21].
Partially supported

From the case studies we see that organisations are motivated by the detection capabilities that anomaly detection promises (detection of unknown/undefined attacks). However, there was no explicit mentioning of a motivation being to use anomaly detection to make the refinement of detection rules more efficient. While not a motivating factor, both organisations have (or intend to) create detection rules based on the output of anomaly detector, e.g. to use in real-time detection systems.

An important motivation factor for security experts of both organisations is that anomaly detection will increase their understanding of what happens on their networks. Furthermore, both FI and NET both mention their organisational situation. As a financial organisation, FI states it is a likely target of attackers and welcomes any additional tool for detecting cyber attacks. As an ISP, it is important for NET to provide their customers with a safe and reliable network.

P2 It is important that the problem owner defines clear objectives, requirements, constraints, and success criteria for the anomaly detector in early phases.[7], [14], [28], [29], [30]
Not supported

Anomaly detection is a new tool for most organisations, most techniques and information comes from academic research, the anomaly detection market is not mature, and there are few commercial solutions available. From the case studies we understand that it may not be realistic to expect problem owners to set clear and specific criteria for the anomaly detectors, in particular in an early stage. An example of specific criteria would be requiring a 0.1% false positive rate before starting the data preparation and modelling.

At FI, general high-level criteria, objectives and requirements were defined in an early stage. At NET, the deployment of anomaly detection does not happen within a clearly defined project, rather as a part of their day-to-day operations.

We see that business organisations want the empowering capabilities that anomaly detection tools promise. In addition, they want to use the tools to increase their understanding of what happens on their networks and systems. For both organisations, these criteria are general in the begin, based on broad goals, that will evolve into specific ones as deployment progresses. In brief, the case study organisations know what capabilities they want, but the problem and solutions are not known well enough to be able to define specific criteria in early stages.

2) Phases 3-4: Data preparation and Modelling:

P6 It is important to reduce the variability and dimensionality of the data. For instance, by filtering and aggregating it as much as possible around the scope of the anomaly detector.[7], [10], [41]
Partially supported

The literature is clear about the problems associated with the variability of organisational data. While both organisations agree that a narrow scope (**P3**) is the right approach, the case studies reveal interesting differences between the two organisations.

As discussed before, NET's approach for anomaly detection is to have specific attacks in mind and only use narrowly scoped subsets of the data each time. Moreover, it was clear that they have not been successful with anomaly detectors that apply to their whole network, nor when applied to individual customers.

FI is currently working on anomaly detectors with the goal to detect specific type of malicious activity, e.g. scans and worm infections. The tools and techniques used come from academic literature, for instance in the form of specific features that can be built from NetFlow that may indicate scans. However, they currently do not apply anomaly detection to specific subsets of data (e.g. ports, protocols).

The differences in approach for the two organisations may be explained with the fact that they are not looking for the same type of attacks. For NET, the main goal is to detect DDoS attacks where they know of specific vulnerable ports/protocols they choose to focus on. At FI, they are looking for activity within their network that indicates advanced (and unknown) attackers or targeted attacks. Still being in an exploratory stage with the goal of detecting 'unknown' attacks it is perhaps premature to expect a clearly defined subset of data.

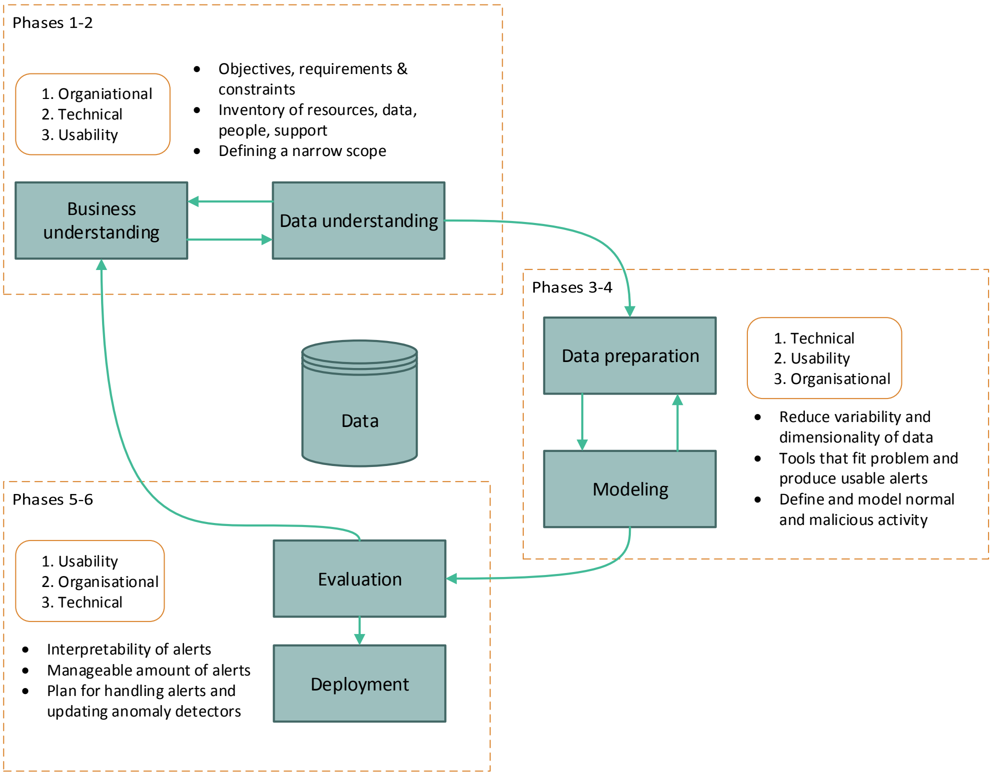
Phases 1-2
1. Organiational
2. Technical
3. Usability
Objectives, requirements & constraints
Inventory of resources, data, people, support
Defining a narrow scope
Business understanding
Data understanding
Phases 3-4
Data preparation
1. Technical
2. Usability
3. Organisational
Reduce variability and dimensionality of data
Tools that fit problem and produce usable alerts
Define and model normal and malicious activity
Modeling
Data
Phases 5-6
1. Usability
2. Organisational
3. Technical
Evaluation
Deployment
Interpretability of alerts
Manageable amount of alerts
Plan for handling alerts and updating anomaly detectors

Figure 1. The conceptual framework of this research

P9 Clearly argue why anomaly detection is suitable for a particular problem or for detecting certain activities, and how the features in the data (selected or extracted) are significant for detecting the activity.[7], [10]
Not supported

On a general level, both organisations are motivated to deploy anomaly detection for its promising capability of detecting unknown/novel attacks. In that sense they can argue why anomaly detection is suitable for a particular problem, e.g. FI will not detect advanced targeted attacks by matching internal data with threat intelligence feeds of known attackers, however, statistical anomalies on the internal network traffic may reveal such attacks.

On a detailed level, these arguments come from academic literature, studies where anomaly detection is applied to a particular problem. Both FI and NET use academic research to guide their efforts or use commercial tools offered as a solution to their problems.

From the case studies we conclude that it is not realistic to expect organisations to have a clear and detailed argument for tool and data choices. For both organisations, in the early phases these clear and specific arguments come from external sources (research papers, commercial tools). Moreover, these efforts are driven by cyber security teams that will go through a phase of trial and error on their path of deploying these tools. In essence, both organisations trying to solve a business problem with innovative tools rather than doing scientific research.

However, as we see from the case study at NET this capability can grow as their maturity level and experience with anomaly detection increases.

3) Phases 5-6: Evaluation and Deployment:

P11 A low false positive rate is essential for operational usability of anomaly detectors. As a result, organisations should prioritise a low false positive rate over a high detection rate.[36], [37]
Partially supported

From the literature study we found that the high false alarm rates are a barrier for deploying a usable anomaly detector within business organisations. Looking at this issue in the case studies showed that the two organisations had a more relaxed view on false positives.

At FI, most acknowledge that there will be a large number of false positives, particularly in the beginning of deployment. The vision is that when an anomaly detector is deployed an improvement and learning process begins. Alerts generated that prove to be false positives will further the experts' understanding of the network. This understanding is then used to explain why an event is a false positive and should not be detected again. More specifically, this process of tuning the models, and understanding the false positives brings valuable knowledge. However, this process should eventually lead to a model that is usable. As a result, false alarms are not a great concern for them in the early phases.

Anomaly detectors that require a lot of time and resources to tune will eventually be abandoned as we saw at NET. From their experience, having a good (combination of) tools to analyse the alerts and find out what is happening is essential. For example, if these tools can reduce the time it takes to investigate some alerts from 2 hours to 2-3 minutes, like in the case of NET, the cost of a false alarm is reduced.

In summary, the experts do agree with the fact that high rates of false positives is a problem. However, there are many ways to reduce the impact of them, e.g. having the tools to quickly investigate. Moreover, the insights a false alert brings can be of value for the security experts as it enhances their understanding of what type of (strange) activity takes place on their networks. More specifically, for the security experts it is a priority to better know their networks.

P14 It is important to address the issue of updating operational anomaly detectors as the notion of normal traffic tends to change over time.[10]
Not supported

The case studies did not reveal any decisive statements of support for this proposition. The team at FI have not reached the stage where these issues are addressed. As for NET, without access to the data and tools at the time it is difficult to say whether the long time spent on tuning models was only due to the poor performance of the algorithms or whether what should be considered as normal was changing during the same time. Hence, this proposition is not supported as neither organisation has consciously dealt with this problem.

More importantly, it is unclear whether the organisations would be able to determine whether false positive rates are caused by a model that needs to be updated because 'normal' is changing, or by a model that performs poorly or does not fit the problem. Moreover, before deployment it may prove difficult to decide a good frequency for updating an anomaly detector.

V. Incompleteness of the theoretical methodology

During the case studies we conducted semi-structured interviews with open questions. The questions purposely covered the broad contents of the theoretical methodology and allowed the freedom of discussing other issues relevant for business organisations. In this section we outline some of the issues observed from practice that were either understated or overlooked in the the propositions. In short, these are issues missing from the methodology or more important in practice. With regard to **RQ 4** this section looks at the *completeness* of the theoretical methodology while in section IV we examined its *soundness*.

In the following subsections we list (in no particular order) and summarise some of these issues. Furthermore, we present relevant (and anonymous) quotes from interviews that capture the issues discussed. Subsequently, we elaborate on them in section VI.

A. Data governance

> *"It is hard to get the data, everything is hard about it, technically and organisationally speaking."*

In one proposition (**P4**) of the theoretical methodology we mention the importance of securing access to data and other resources from within the organisation. Considering the insights from the case studies it can be concluded that the theoretical methodology understates this issue.

Even when a security department is allowed to use the data, getting the data is not as simple as showing up and asking for it. It can still be an immense technical and organisational challenge to actually retrieve the data. For instance, the data may be owned by other departments that have neither the interest nor capability to collect and deliver network logs.

Once the data is collected it is has to be made accessible to the data scientist. Furthermore, the raw data is not always enough, metadata describing the data is also needed, e.g. a network/asset model that describes the topology of a particular network segment, and helps identify the actors on the network. Again, this information may belong to other departments that have to agree to deliver this meta-information.

There are many other challenges related to data, including the storage of large amounts of data, managing performance simultaneously for data scientists and security experts, building a system that is compliant to an organisation's security and privacy policy, annotating the data with contextual information, and dealing with poor data quality.

In summary, the theoretical methodology does not emphasise the issues of getting and working with data with the same weight as these issues appear in the case studies. The problems with data may arise at different times from different directions, but all directly affect the anomaly detector.

B. Laws and regulations

> *"There are laws in some countries that prohibit logs from leaving the country."*

In the theoretical methodology, the only challenge related to laws and regulation has to do with whether it is allowed to use live data (**P5**), i.e. simulated data is often used for legal and privacy reason.

In practice, these issues have a wide impact and can introduce complexities to an anomaly detection project. Firstly, everything from departmental policies to international law has to be considered. Secondly, data collection and transfer is subject to many laws. For instance, at FI we saw that banking secrecy laws of one country dictates whether a certain data source is available, and data protection laws in another country forbids logs ever leaving the country. Thirdly, laws regarding privacy are not always clear, and subject to interpretation, meaning lawyers can interpret the same law differently. Fourthly, a deployed and operational anomaly detector has to be in compliance with company regulations. Reaching the point of compliance can be costly in terms of time spent doing the necessary documentation and analysis. Furthermore, as observed at FI, requirements and constraints for these types of projects often come from people working with risk management, compliance or legal departments.

C. Understanding the network

> *"We do basic monitoring, we know our baseline, and we know our network. Without this knowledge, and without tools for doing and checking your baseline you cannot effectively do anomaly detection."*

The propositions of the theoretical methodology are focused on the end goal, deploying an operational and usable anomaly detector. The methodology is about defining a clear scope and requirements, apply statistics or machine learning techniques, generate usable alerts, and deploy the technology in the organisation.

From the case studies we see that the respondents, and particularly the security experts, want to use these tools to better understand what is happening on their networks and systems, i.e. gain situational awareness. Considering the detailed problem definition and narrow scope the theoretical methodology ask of organisations we argue that it assumes this situational awareness is already at hand.

However, we see that the tools that support business organisations in developing a good understanding of the networks are being deployed simultaneously as those that allow them to do anomaly detection, e.g. Hadoop and Mapreduce frameworks, visualisation tools. Furthermore, a good understanding of the network and tools that support exploring the data helps with investigating of alerts.

In brief, the theoretical methodology does not sufficiently address the need of security experts, or those who verify alerts, to have a good understanding of what goes on in the organisation's networks and systems.

D. Finding and retaining talent

> *"In the current labour market it is difficult to get the right people for a project like ours."*

The theoretical methodology implicitly assumes that the business organisation already has all the experts needed to deploy anomaly detection. In reality, finding and retaining talent is a significant challenge that affects anomaly detection projects in various ways.

For instance, we saw that the case study organisations need three main types of experts: data scientists, security experts, and data engineer (Hadoop specialists, infrastructure developers). Reportedly, these are all in great demand and in short supply, making them difficult to find and retain in the current labour market.

To get the necessary expertise they have had to partially rely on external employees. Naturally, this is a temporary arrangement and in the end of the project the external employees leave and take their knowledge with them. Moreover, the organisations have invited students to deploy anomaly detection models, e.g. as a part of their graduation. Similarly, the students are likely to leave with valuable knowledge, leaving the internal employees with the difficult task of tuning the deployed models.

As for the internal employees, they gain valuable experience from working on these 'cutting edge' projects, e.g. building a threat management system and deploying anomaly detection. This increases their value on the labour market which may encourage them to leave for other opportunities. Furthermore, it may be hard to provide talent with an environment where they can enjoy their work. For example, providing good flexibility and freedom while working with such sensitive information.

The issues with talent can affect various parts of anomaly detection projects, like deciding what kind of tools to use.

E. The open source community and the commercial market

> *"Some models seemed promising, with many academic papers about them, but in practice they did not work."*

Another important issue missed by the theoretical methodology is the decision whether to go an open source route, a commercial one, or combining the two. This is not an easy decision, both types have their strengths, weaknesses, and accompanying requirements.

Open source tools, including the use of scientific research, are often the only choice available as the vendors of commercial solutions have not been offering many solutions in recent years. However, these tools introduce a high level uncertainty and dependency on highly skilled people into anomaly detection projects. Moreover, they often lack functionalities that are required in a corporate environment that have to be built, e.g. authentication and user management.

As for commercial tools, though the market has not offered many solutions, a few vendors and commercial anomaly detection tools have become increasingly mature and sophisticated in the recent years. The main drawback of the more sophisticated tools is their high price. However, they come with less uncertainty than the open source tools, they are likely to come with a range of functionalities organisations require, and have less demand on talent.

VI. Discussion

A. Gathering resources and support

The topic of gathering the necessary resources and support was explored to some extent in the literature review. During the case studies we saw that these issues are critical. For example, in one of the case studies all eight respondents said that getting data from within the organisation was the biggest (or one of the biggest) challenge they have encountered. In the following paragraphs we elaborate on these issues in light of what we observed in the case studies (summarised in Figure 2 and Figure 3).

a) Data.: In large organisations, getting data is not as simple as showing up and asking for it. Obviously it is vital to get data, and in one organisation's view this was the most difficult part of their project.

First, network logs are owned or managed by the various different departments. They often need some convincing, as they do not necessarily understand the usefulness of the data or only use it for operational purposes, e.g. store it for 2 weeks or collect only samples of the data.

Second, capabilities like generating NetFlow, and delivering it, do not necessarily exist in all these department, meaning they may need support in the form of advice on devices to collect data, best practices, or funding.

Third, legal and privacy laws on a local, national and international level make it difficult to collect some data, e.g. some laws prohibit logs leaving the country of origin.

Fourth, once the data is delivered it does not come with any quality assurance. This can cause a problem for the anomaly detection if the models are being built with faulty data.

Fifth, it is challenging to be context-aware in such a large setting, with data from so many sources. Without a system of collecting and providing these contextual information in an organised way it is difficult to understand or explain the data, e.g. for verifying alerts.

Sixth, as the amount of data grows the scale of the infrastructure has to grow as well. It is difficult to know whether the decisions that make the system run today will not impede future growth.

b) Project team.: From the case studies we observed that within the project team the system is highly dependent on people, the talent running the system. In one case study organisation, once operational, the threat management system (including anomaly detection) needs talented data scientists, experienced security people, developers, and data engineers (e.g. Hadoop specialists).

c) Organisational support and dependencies.: Within the organisation we see that these projects are dependent on several different departments and stakeholders.

Firstly, there is a critical dependency on local IT- and network departments for getting new data sources into the system.

Secondly, due to privacy sensitive data is used, people from risk management, data protection, legal, and compliance departments have to be actively involved. Moreover, these parties have to be consulted before using certain data or expanding the scope of the system to incorporate more data sources.

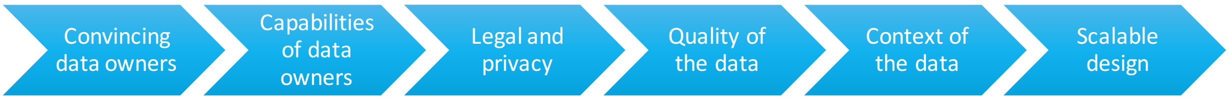

Figure 2. Summary of main challenges with getting data

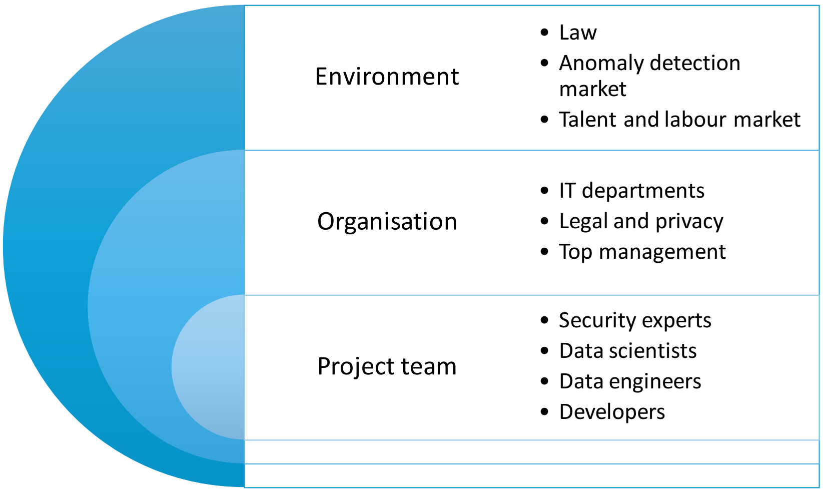

Figure 3. Summary of main organisational dependencies

Thirdly, the (continued) support and funding from senior level management, where some of them may have other priorities than cyber security in mind.

d) Environment.: Outside the organisation anomaly detection projects are mainly dependent on three factors.

Firstly, the system is highly dependant on various (privacy, data protection, banking) rules policies and laws on different levels (local, national, international) that affect how organisations can work with data. For example, some countries have banking secrecy laws that affect whether and how it is allowed to extract certain data. Moreover, others have laws prohibiting the logs from leaving the country.

Secondly, projects are dependent on the outside anomaly detection 'market' consisting of academic research on the topic, open source communities and tools, vendors of commercial solutions, other organisations pursuing anomaly detection, and freely available or commercial models.

Currently, doing an anomaly detection project is essentially a research project as they have to rely on academic papers and research. One challenge is that there is limited amount of (open source or commercial) readily available tools and models for doing anomaly detection. However, the project team has observed the outside world becoming more mature in the two years of working on the project, both vendors of commercial software and other organisations pursuing similar projects.

Thirdly, the labour market. Finding and retaining talent for such projects is difficult. The expertise needed for the building, maintaining and operating the system are all hard to find in the current labour market. Moreover, external employees hired to fill the gap in such projects leave with valuable knowledge, and internal employees that work on these projects increase in value and may leave as a consequence.

While none of these issues are unique to anomaly detection we have observed all of them directly influence such efforts. Most of the issues are hard to solve and come from outside the core team involved in deploying anomaly detection. In essence, the combination of challenges is unfortunate for the potential success of these tools in a practical setting.

B. Defining and modelling normal and anomalous

One of the fundamental statement from the literature study about the challenges of anomaly detection is that it is hard to define normal with operational data, and that non-malicious and non-interesting anomalies are common. In the following paragraphs we discuss how the case study organisations looked at the topic (summarised in Figure 4 and Figure 5).

a) Defining and modelling normal: Business organisations deploy anomaly detection to detect malicious activity. More specifically, an anomaly in and of itself is not automatically considered malicious. Before making that distinction malicious and anomalous, organisations have to define what an anomaly is and the only way to do that is to define what is normal traffic.

From the case studies we observe that the definition of normal in statistics is not the same as in cyber security (Figure 4). First, from a data science perspective, 'normal' is when an event looks like most others, and has many neighbours around. Second, from an organisational cyber security perspective, the definition is more complex. Normal is when the organisation is not being attacked, and there are no threats of interruptions by internal or external parties.

From both theory and practice we know that it is not possible to define normal unless using live data, i.e. simulated data cannot be used to model and distinguish normal and anomalous. However, it is hard to define normal traffic because organisations do not know if they have normal traffic, i.e. live operational data may or may not contain malicious activity. The only way to overcome this challenge would be to use supervised machine learning on a training dataset that has been manually checked for normal and malicious behaviour, but on real-life scale that is not feasible.

We observed some real life challenges that organisations have had with defining normal. For instance, two similar customers of an ISP (e.g. universities in the same country) do not share a common definition of normal traffic patterns.

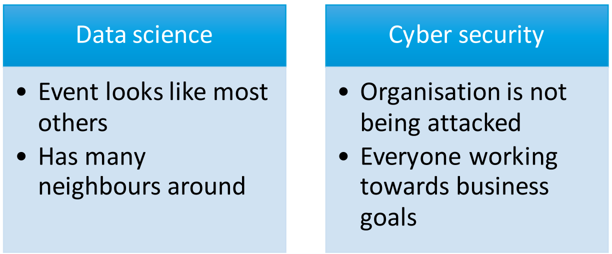
Data science
Event looks like most others
Has many neighbours around
Cyber security
Organisation is not being attacked
Everyone working towards business goals

Figure 4. 'Normal' according to data science and cyber security

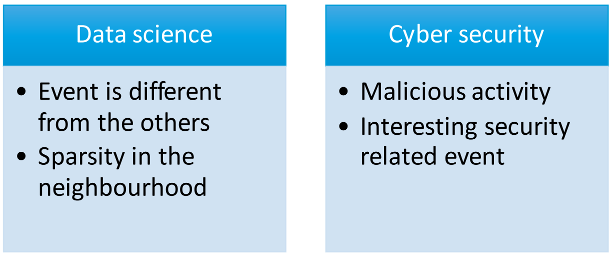
Data science
Event is different from the others
Sparsity in the neighbourhood
Cyber security
Malicious activity
Interesting security related event

Figure 5. 'Anomaly' according to data science and cyber security

Moreover, tools based on available research designed to find statistical anomalies in traffic data simply fail to find patterns in the data.

The limited success we observed only happened in cases where the scope of the detector is small. For instance, in one organisation they were able to effectively define normal traffic when limiting the scope to a single protocol and port (e.g. UDP traffic on the DNS port). In their experience, only in narrowly defined subsets of traffic they observe a relatively stable baseline.

b) Defining and modelling anomalies.: As mentioned before, anomaly detection is about detecting malicious activity, and similarly to the problems with the definition of 'normal', 'anomaly' in data science and cyber security is not the same (Figure 5). From a data science perspective, you are anomalous when an event is different than most others. The anomalies that organisations are looking for are not statistical anomalies as such, instead malicious activity, i.e. an attack on the organisation. In addition, organisations are interested in finding non-malicious but security related anomalies, e.g. data leakage or misconfiguration.

The case studies produced both promising and alarming insight into the issue of detecting anomalies in practice.

One one hand, it is important to keep in mind that interesting anomalies do not only come from the use of advanced statistical models. The case studies revealed a variety of different ways anomalies have been detected when building a detection platform or working with data. First, working with data on other parts of the system (e.g. data ingestion). On several occasions the developers handling the data have stumbled upon some strange events and patterns in the data. Second, doing basic visualisation and overview of the data can reveal interesting fluctuations that the team wants to investigate. Third, using basic statistics, looking at extreme values in the network traffic logs one data scientist did detect (quick) port scans by internal employees. Fourth, by doing a time-series analysis they detected and investigated some of the outliers. These investigations lead to the discovery of a malfunctioning router that was causing some of the anomalies. To summarise, working with the data in new ways can reveal interesting anomalies and insights that bring value to the organisation.

On the other hand, some experts shared some of the negative experience they have had with detecting anomalies. Firstly, it is difficult to define a narrowly scoped problem that will generate relevant anomalies. Secondly, the organisation deployed a system that once an attack is detected, it uses algorithms to generate rules to mitigate future attacks. As it turned out, the rules it generated were too specific to detect anything remotely different to the original attack. Thirdly, the have had to abandon models based on promising research; for instance, after struggling to tune them for over a year.

Based on the findings of the case studies and literature review it can be argued that the problem of defining and modelling normal and anomalous is difficult to overcome. Unlabelled data used to build anomaly detectors may well contain the attacks the goal is to find. Promising ideas and models do not work as intended, and promising efforts get abandoned. In fact, unless with a very narrow scope, we have yet to see an operational anomaly detector that is producing satisfactory results.

The main promise of anomaly detection is to detect unknown or undefined attacks. Having to define a very narrow scope seems to beat that purpose of using the technique in the first place. Organisations want to deploy models that can detect all kinds of anomalies, including attack patterns no one has thought of. In that sense, anomaly detection does not seem to live up to expectations.

C. Open source vs commercial tools

The decision on whether to develop the anomaly detector (system) yourself using open source technologies or purchase commercial solutions on the market (or deciding on a balance between the two) was prominent in the case studies. This was possibly the most pressing issue not included in the theoretical methodology.

Whether to build the system internally or purchase commercial solutions is a critical decision for organisations doing anomaly projects.

Both options are expensive, have their advantages and their disadvantages. Building a system requires strong development capabilities, and talent that is difficult to find and retain. Furthermore, the project is more unpredictable, e.g. since many functionalities required by the organisational environment are not included in open source solutions. Buying the solution on the market often comes with a hefty price tag, promises that fail to live up to expectations, and less flexibility to adapt the solution to specific scenarios.

Weighing the advantages and disadvantages the discussion we find ourselves in favour of commercial solutions.

We assume that most organisations are after the capabilities that the techniques promise, as opposed to wanting to have the ability to build such a system. Building a good anomaly detection platform that includes data ingestion, data storage, machine learning capabilities, visualisation, and alerting requires substantial development effort. In addition, there are many functionalities required by organisational rules (privacy, authorisation, authentication, data segregation) that have to built around the open source tools. From what we see in the case studies, going the 'open source route' is unpredictable, and requires talent that is hard to find and retain in today's labour market.

In brief, defining the models are research projects and building the infrastructure is an immense development effort. Unless an organisation is a research driven software development company, they are likely to find themselves way out of their comfort zone.

With commercial solutions organisations saves them from the potential problems that come with building such a platform, although they are sacrificing the flexibility of that open source solution provide.

VII. Future research

In this section we propose recommendations for researchers. A work like this can not thoroughly explore all directions and topics that present themselves. These topics remain as research interest for future research projects. Considering these topics, limitations of this work, and main findings we propose the following directions to be explored further in future research.

a) A practical anomaly detection methodology.: In this work the theoretical methodology's main use was to guide the research and analysis of empirical evidence. Furthermore, data from the case studies was used to test the theoretical methodology, and give new insights into anomaly detection in practice. However, the methodology was not put to the test in practice, e.g. by following it when addressing a cyber security problem. In addition, two case studies are not sufficient to generalise for anomaly detection in general. Future research topic is to use the results of this work to iteratively refine and test the methodology in practice and in different anomaly detection scenarios with the goal of constructing a practical anomaly detection methodology.

b) Working with alerts and false positives.: Further research on usability is needed. The problem of maintaining an anomaly detector and dealing with different types false positives presents a future research challenge. In this work we have found that false positive rates are and will remain a challenge. Addressing the problem only by decreasing the rate by reducing scope may not fit with project where the goal is to detect unknown attacks, and it is challenging to produce interpretable alerts. From the literature these two were identified as the ways to address usability, but the case studies revealed more ways of achieving usability.

First, research is needs to focus on ways to manage and handle false positives, e.g. how to incorporate white-listing of events that are anomalies but should not generate alerts. Second, research should identify good practices of monitoring and diagnosing performance of anomaly detectors so that practitioners can better respond when performance is poor. Third, research must clearly explain any assumptions made regarding the updating of the anomaly detector.

In summary, there is need for research specifically on the usability of anomaly detectors and methods of handling alerts and false alarms.

c) Gathering data and support.: Research is needed on the best practices with gathering and working within this type of data across business organisations, especially large multinational ones that face even greater challenge of working with the different rules and regulations of different countries.

Furthermore, it is important to further explore these insights to better understand the organisational side of doing anomaly detection. The organisational issues from the literature review mostly come from research on data mining in general. However, many organisational challenges can affect anomaly detection projects, as we have seen from the case studies. For example, getting data that is distributively generated and governed throughout an organisation, setting up a project team and responsibilities, getting support, and working with different legislation.

d) Open source vs commercial tools.: From the case studies we saw that organisations look to academic research for models to deploy. We observed the difficulties of implementing the tools described in research, e.g. in many papers an algorithm is described or a certain technique used but key information is missing. To help lower to cost of the 'open source route' researchers should increasingly share code used in their work, or the very least provide clear information about the tools used, parameters set, and other criteria that make it easier to replicate the approach.

In summary, a future research direction is to explore, combine and test freely available tools for their potential for deployment, and operational success in business organisations

VIII. Summary

In this article we have explored the topic of anomaly detection, focusing on the discrepancies between practice in academia and industry. When compared to promising research results, anomaly has been relatively unsuccessful when deployed in a business environment. It is important to better understand this topic as these techniques have the potential to detect harmful and sophisticated attacks.

In the section VI we answer the main research question by discussing four key issues that exemplify the discrepancies between doing anomaly detection in practice and in research.

1. We identify an unfortunate combination of organisational challenges that can affect the success of anomaly detection projects.
2. We explain how the technical challenge of defining and modelling normal and anomalous activity affects practice.
3. We identify the main discrepancies between alerts and false positives in research and industry.
4. The decision of choosing between open source and commercial solutions is an important factor for anomaly detection in business organisations.

On our path of answering the main research question we have also gained valuable knowledge.

In a literature study we learnt that:

- Anomaly detection has the promise and potential to detect attacks without the prior knowledge needed for constructing conventional detection rules. The process of building rules is costly and rules constrained by the experience and imagination of the experts that construct them.
- There are significant technical challenges of doing anomaly detection in business organisations, i.e. assumptions often made in research do not hold in reality and there is an unfortunate combination of unique problems with using machine learning to detect cyber attacks.
- There are substantial organisational challenges that anomaly detection faces in practice. With these techniques being new to most organisations it is difficult to define a problem and gather the necessary resources.
- Usability has to be considered from early phases and can be achieved with a combination of two things. First, by having a very low false positive rate. Second, by producing actionable and interpretable alerts.

Using the results of the literature review we then construct a theoretical methodology for deploying operational anomaly detection. The methodology consists of 15 propositions, statements from the academic literature on how organisations should approach the problem. The propositions are evenly divided between the different phases of CRISP-DM and address the most important technical, organisational, and usability issues we have identified (see summary in Figure 1).

To understand what anomaly detection looks like in practice we conducted a series of interviews where we collected information from two different organisations (FI and NET) at different stages of deploying anomaly detection

More specifically, the results of the case study are input for the testing of the theoretical methodology in section IV, and in the discussion on the core discrepancies between theory and practice in section VI.

In brief, we learnt that business organisations approach anomaly detection differently than in academic research, and many important issues from practice have little impact on academic research.

IX. Conclusions

Our observations suggest that organisations are motivated by the promising capabilities of anomaly detection tools, i.e. detect unknown attacks, and its use as a tool to better understand what goes on in the organisation's networks and systems. Consequently, they do not approach this problem in a scientific way with clearly defined problem to solve using statistics or machine learning, it is a capability that they want to have.

Furthermore, people issues play a big role in business organisations. For instance, the choice of tools (e.g. open source or commercial) has implications on what kind of talent is needed within the team. Getting the right talent then depends on the labour market, and as our interviews suggest, it is hard to find and retain people with the right skills (security, data mining, analytics infrastructure).

References

[1] A. Patcha and J.-M. Park, "An overview of anomaly detection techniques: Existing solutions and latest technological trends," *Computer Networks*, vol. 51, pp. 3448–3470, Aug. 2007.

[2] B.-J. Koops, "The internet and its opportunities for cybercrime," *Transnational Criminology Manual, M. Herzog-Evans, ed*, vol. 1, pp. 735–754, 2010.

[3] National Cyber Security Centre - Ministry of Security and Justice, "Cyber Security Assessment Netherlands 2014," tech. rep., 2014.

[4] D. E. Sanger and N. Perlroth, "Bank Hackers Steal Millions via Malware," 2015.

[5] A. Jeffries, "'Celebgate' attack leaks nude photos of celebrities," 2014.

[6] P. Zengerle and M. Cassella, "Millions more Americans hit by government personnel data hack," 2015.

[7] R. Sommer and V. Paxson, "Outside the Closed World: On Using Machine Learning for Network Intrusion Detection," *Security and Privacy (SP), 2010 IEEE Symposium on*, 2010.

[8] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys (CSUR)*, vol. 41, no. 3, p. 15, 2009.

[9] D. Denning, "An Intrusion-Detection Model," *IEEE Transactions on Software Engineering*, vol. SE-13, pp. 222–232, Feb. 1987.

[10] C. Gates and C. Taylor, "Challenging the anomaly detection paradigm: a provocative discussion," in *Proceedings of the 2006 workshop on New security paradigms*, pp. 21–29, ACM, 2006.

[11] Cloud Security Alliance, "Big Data Analytics for Security Intelligence," tech. rep., 2013.

[12] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network Anomaly Detection: Methods, Systems and Tools," *IEEE Communications Surveys & Tutorials*, vol. 16, pp. 303–336, Jan. 2014.

[13] S. X. Wu and W. Banzhaf, "The use of computational intelligence in intrusion detection systems: A review," *Applied Soft Computing*, vol. 10, pp. 1–35, Jan. 2010.

[14] G. M. Weiss, "Data Mining in the Real World: Experiences, Challenges, and Recommendations.," in *DMIN*, pp. 124–130, 2009.

[15] W. Lee, S. Stolfo, and K. Mok, "Mining in a data-flow environment: Experience in network intrusion detection," *... on Knowledge discovery and data mining*, 1999.

[16] C. Kruegel and G. Vigna, "Anomaly detection of web-based attacks," in *Proceedings of the 10th ACM conference on Computer and communications security*, pp. 251–261, ACM, 2003.

[17] K. S. Killourhy and R. A. Maxion, "Toward realistic and artifact-free insider-threat data," in *Computer Security Applications Conference, 2007. ACSAC 2007. Twenty-Third Annual*, pp. 87–96, IEEE, 2007.

[18] D. Hadžiosmanović, L. Simionato, D. Bolzoni, E. Zambon, and S. Etalle, "N-gram against the machine: On the feasibility of the n-gram network analysis for binary protocols," in *Research in Attacks, Intrusions, and Defenses*, pp. 354–373, Springer, 2012.

[19] R. K. Yin, "Case study research design and methods third edition," *Applied social research methods series*, vol. 5, 2003.

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

[21] S. Axelsson, "Intrusion detection systems: A survey and taxonomy," *Technical report*, vol. 99, 2000.

[22] O. Chapelle, B. Schölkopf, and A. Zien, "Semi-supervised learning," 2006.

[23] M. Handley, V. Paxson, and C. Kreibich, "Network Intrusion Detection: Evasion, Traffic Normalization, and End-to-End Protocol Semantics.," in *USENIX Security Symposium*, pp. 115–131, 2001.

[24] K. M. C. Tan, K. S. Killourhy, and R. A. Maxion, "Undermining an anomaly-based intrusion detection system using common exploits," in *Recent Advances in Intrusion Detection*, pp. 54–73, Springer, 2002.

[25] R. Pang, V. Yegneswaran, P. Barford, V. Paxson, and L. Peterson, "Characteristics of internet background radiation," in *Proceedings of the 4th ACM SIGCOMM conference on Internet measurement*, pp. 27–40, ACM, 2004.

[26] M. V. Mahoney and P. K. Chan, "Learning rules for anomaly detection of hostile network traffic," 2003.

[27] R. A. Maxion and F. E. Feather, "A case study of ethernet anomalies in a distributed computing environment," *Reliability, IEEE Transactions on*, vol. 39, no. 4, pp. 433–443, 1990.

[28] P. Chapman, J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer, and R. Wirth, "CRISP-DM 1.0 Step-by-step data mining guide," 2000.

[29] H. R. Nemati and C. D. Barko, "Key factors for achieving organizational data-mining success," *Industrial Management & Data Systems*, vol. 103, no. 4, pp. 282–292, 2003.

[30] A. Hilbert, "Critical Success Factors for Data Mining Projects," in *Data Analysis and Decision Support*, pp. 231–240, Springer, 2005.

[31] A. T. Zhou, J. Blustein, and N. Zincir-Heywood, "Improving intrusion detection systems through heuristic evaluation," in *Electrical and Computer Engineering, 2004. Canadian Conference on*, vol. 3, pp. 1641–1644, IEEE, 2004.

[32] P. Jaferian, K. Hawkey, A. Sotirakopoulos, M. Velez-Rojas, and K. Beznosov, “Heuristics for evaluating IT security management tools,” *Human–Computer Interaction*, vol. 29, no. 4, pp. 311–350, 2014.

[33] T. Ibrahim, S. Furnell, M. Papadaki, and N. Clarke, “Assessing the Usability of Personal Internet Security Tools,” in *Proceedings of the 8th European Conference on Information Warfare and Security*, pp. 102–111, Academic Conferences Limited, 2009.

[34] T. Patil, G. Bhutkar, and N. Tarapore, “Usability evaluation using specialized heuristics with qualitative indicators for intrusion detection system,” in *Advances in Intelligent Systems and Computing*, vol. 176 AISC, pp. 317–328, Springer, 2012.

[35] T. Ibrahim, S. M. Furnell, M. Papadaki, and N. L. Clarke, “Assessing the usability of end-user security software,” in *Trust, Privacy and Security in Digital Business*, pp. 177–189, Springer, 2010.

[36] R. Lippmann, J. W. Haines, D. J. Fried, J. Korba, and K. Das, “The 1999 DARPA off-line intrusion detection evaluation,” *Computer networks*, vol. 34, no. 4, pp. 579–595, 2000.

[37] S. Axelsson, “The base-rate fallacy and its implications for the difficulty of intrusion detection,” in *Proceedings of the 6th ACM Conference on Computer and Communications Security*, pp. 1–7, ACM, 1999.

[38] J. Nielsen, “Usability inspection methods,” in *Conference companion on Human factors in computing systems*, pp. 413–414, ACM, 1994.

[39] S. Rosenbaum, “Usability evaluations versus usability testing: When and why?,” *Professional Communication, IEEE Transactions on*, vol. 32, no. 4, pp. 210–216, 1989.

[40] J. R. C. Nurse, S. Creese, M. Goldsmith, and K. Lamberts, “Guidelines for usable cybersecurity: Past and present,” in *Cyberspace Safety and Security (CSS), 2011 Third International Workshop on*, pp. 21–26, IEEE, 2011.

[41] R. A. Maxion and K. M. C. Tan, “Benchmarking anomaly-based detection systems,” in *Proceedings of the 2002 International Conference on Dependable Systems and Networks*, pp. 623–630, 2000.

Proposition	**FI**	**NET**	**Support**
Business understanding & Data understanding			
P1 Business organisations are motivated to deploy anomaly detection for detecting new/unknown attacks and for automatic refinement of detection rules.	+ / ⋆	+ / ⋆	Partial
P2 It is important that the problem owner defines clear objectives, requirements, constraints, and success criteria for the anomaly detector in early phases.	− / ⋆	− −	No
P3 Anomaly detectors should have a narrow scope, e.g. around a specific attack or malicious activity. A small scope is a success factor for reduced false positive rate and increased likelihood of project success.	+	+ +	Yes
P4 It is important to take inventory of technical and organisational resources in the early stages. E.g. ensure that data is available and of good quality, arrange access to domain experts in advance and verify that the project is sufficiently supported within the organisation.	+ +	⋆	Yes
P5 It is essential to use data that is representative of the traffic of an organisation's networks and systems, e.g. real-life operational data. It is difficult to simulate realistic activity. The results of anomaly detectors on simulated data are not a good indicator of operational results.	+ +	+	Yes
Data preparation & Modelling			
P6 It is important to reduce the variability and dimensionality of the data. For instance, by filtering and aggregating it as much as possible around the scope of the anomaly detector.	−	+ +	Partial
P7 It is important to keep in mind that attacks are not necessarily anomalous and/or rare. Furthermore, anomalies are not necessarily malicious or interesting. The goal is to find specific malicious activity, not statistical anomalies.	+	+ +	Yes
P8 When selecting an algorithm or tool, it is essential to consider the interpretability of its output. It is challenging to transform detected anomalies into actionable alerts.	n.a.	+ +	Yes
P9 Clearly argue why anomaly detection is suitable for a particular problem or for detecting certain activities, and how the features in the data (selected or extracted) are significant for detecting the activity.	+ / −	− / ⋆	No
P10 While supervised machine methods cannot be used on unlabelled data, they can be used to post-process the alerts to reduce false positives and/or provide additional information that accelerates verification of alerts.	+ +	+ / −	Yes
Evaluation & Deployment			
P11 A low false positive rate is essential for operational usability of anomaly detectors. As a result, organisations should prioritise a low false positive rate over a high detection rate.	+ / −	+ / ⋆	Partial
P12 It is important to achieve a very low rate of false positives when working with large amounts of data. Even a small false positives rate ($>1\%$) can result in an unusable anomaly detector.	+ +	+ +	Yes
P13 It is essential that an anomaly detector generates actionable and interpretable alerts. Evaluating alerts is time consuming and is done by expensive and busy experts.	+ +	+ +	Yes
P14 It is important to address the issue of updating operational anomaly detectors as the notion of normal traffic tends to change over time.	n.a.	+ / −	No
P15 Organisations should recognise the value understanding what makes an alert either a true- and false positive. This knowledge can be used to improve detectors or construct detection rules.	+ +	+ ⋆	Yes

Table I
SUMMARY OF CASE STUDY FINDINGS