Unpacking the Complex Socio-Technical Systems Assemblages in Cybersecurity

Mamello Thinyane
University of South Australia, Adelaide, Australia
mamello.thinyane@unisa.edu.au

Abstract: The ensuing digital transformation means that cybersecurity solutioning increasingly occurs in the context of complex intractable socio-technical systems comprising non-technical elements, including human, social, and societal factors. These evolving cybersecurity ecosystem dynamics, at the confluence of cyber-physical-social spaces, present several challenges to techno-centric cybersecurity solutions including for risk assessment, threat modelling, and incident analysis. This paper unpacks the complexity of the cybersecurity domain and illustrates the associated socio-technical systems assemblages through a case study and situational analysis of a cybersecurity incident. It then reviews socio-technical systems analysis approaches from the safety management domain and discusses the alignment with and relevance for cybersecurity. The utility of these approaches is demonstrated by applying the functional resonance analysis method to the said cybersecurity incident. The situational analysis surfaces the diverse set of factors, including human, non-human, cultural, economic, institutional, and global, that directly played a role in the unfolding of the incident, and which need to be considered in risk assessment and incident analysis. Further, analysing the incident through the functional resonance analysis method shows the functional dependencies and cascade of performance variability between the different elements in this situation, which goes beyond simple, root-cause, linear causality, and purely technical explanations. Overall, the paper explicates the need for cybersecurity risk assessment and incident analysis that is commensurate with the complexity of underlying socio-technical cyber systems.

Keywords: Socio-Technical Systems, Functional Resonance Modelling, Complex Systems, Cybersecurity, Safety Management, Resilience Engineering

1. Introduction and Background

The cybersecurity threat landscape has evolved significantly in tandem with developments in computing; from Charles Babbage’s Difference Engine – the first computing device, Electronic Numerical Integrator and Computer (ENIAC) – the first general purpose computer, Advanced Research Project Agency Network (ARPANET) – the predecessor to the Internet, to the modern day computing that is deeply integrated into cyber-physical-social systems that are supporting critical functions for individuals, organizations, and society at large (Pasandideh et al., 2022). This evolution has heightened the importance of systems perspectives and also concerns regarding not only technocentric threats but also human, organizational and societal threats (Shoemaker et al., 2020; Von Solms & Van Niekerk, 2013; Zoto et al., 2019). As such, systems thinking has increasingly been adopted in cybersecurity to understand computing systems, cyber threats, and cybersecurity solutions holistically as part of a complex ecosystem and as comprising interrelated and interdependent elements that shape the overall cybersecurity posture (Susan M. Tisdale, 2015).

An important element of this evolution has been the integration of computing into increasingly complex and intractable socio-technical systems. Complex systems, which are not just complicated because of the many components and interactions, are dynamic and exhibit emergent behaviour that makes them underspecified and not fully understood. This is in contrast to tractable systems which can be fully described, decomposed into subcomponents, do not change while being described, and whose functions are homogenous and regular (Hollnagel, 2012). Complexity has been considered across many aspects of cybersecurity such as the infrastructure of cyber-physical systems, threat landscape, and organizational architectures (Harbertson et al., 2023; Wen et al., 2017). The complexity of systems has a direct implication on how the functioning of those systems is understood and managed, but more important how dysfunction and malfunctioning are understood and managed. The implication of this, and the motivation for this research, is to give recognition that the approaches, methods, and tools of the trade in cybersecurity need to be commensurate with this complex nature of the underlying socio-technical systems (Perrotin et al., 2022).

The primary line of inquiry in this paper is “How can the cybersecurity domain be understood from a socio-technical systems assemblages perspective?” with the corollary of “How effective are socio-technical systems analytical frameworks (from safety management) for the cybersecurity domain?” To address these questions this paper undertakes a detailed situational analysis of a specific cybersecurity incident to explore the complex socio-technical assemblages implicated in the situation. It then discusses the relevance of socio-technical
Mamello Thinyane

systems approaches from the safety management domain and subsequently demonstrates the utility of the functional resonance analysis method (FRAM) for cybersecurity incident analysis.

2. Unpacking the Complexity in Cybersecurity Incidents

A case study of the Equifax data breach that occurred in May 2017 has been undertaken to explore and illustrate the inherent complexity of cybersecurity incidents. This case has been chosen based on the profile of the incident, level of public domain reporting, and detailed documentation available from the subsequent investigations (GAO, 2018; Michael et al., 2017; Miyashiro, 2021). The situational analysis methodological approach is used to explore this incident.

2.1 Methodological Approach: Situational Analysis

Situational analysis, which is a variant of the Grounded Theory research method, is a qualitative critical analysis method developed to explore situations by unpacking the complexity within the situation (Clarke, 2005). The method focuses on the situation as a unit of analysis giving recognition to the fact that it is impossible to disentangle things and actions from the context in which they occur. That all elements in a situation “both co-exist and affect” all other elements in that situation (Clarke, 2021). This is a necessary foundation for explicating complexity because it gives due recognition to the influences and interactions within systems that lead to uniquely emergent behaviours and outcomes.

Situational Analysis employs visually oriented analytical strategy to highlight the key elements present in a situation, the interactions and relationships between those elements, and the context within which situation occurs (Clarke, 2005). The key analytical tools within situational analysis are (Clarke, 2021): situational maps – which identify the main elements present in a situation; relational maps - which map the key interactions between elements; social worlds and arena maps – are collectives around a shared commitment within a situation; and lastly positional maps – which identify key discursive positions on specific issues within the situation of analysis. In this research three tools, messy situational maps, ordered situational maps, and relational maps, are utilized to explore the Equifax incident.

2.2 Equifax Data Breach Case Study

Equifax is a credit reporting agency that suffered a data breach, between May and July 2017, which exploited a vulnerability in Apache Struts, to expose details of over 145 million people (Miyashiro, 2021). The Equifax CEO at the time attributed the data breach to the failure of the employee who was responsible for patching vulnerabilities, however several factors, some of which are discussed hereafter, were prevalent in this situation and contributed to the incident as highlighted in the messy situational map in Figure 1.

Figure 1: Messy situational map of the Equifax data breach incident
In the formal Equifax investigation, the breach was attributed to four main factors: weaknesses in the identification functions, poor detection and monitoring capabilities, lack of segmentation of access to databases, and weaknesses in data governance (GAO, 2018). Further weaknesses were noted in the detection and monitoring functions in Equifax related to the failure to maintain intrusion detection systems; an SSL certificate which was required to inspect encrypted network traffic had expired before May 2017 leading to hackers bypassing the traffic inspection systems (GAO, 2018).

The leadership at Equifax had invested in improving their cybersecurity posture by having a dedicated security operations team, including a Global Threat and Vulnerability Management (GTVM) team. They had also engaged the services of Mandiant, a cybersecurity company, to improve their cybersecurity posture. However, several factors were at play at the organizational level that unfolded in the leadup to the data breach. These include a “talent exodus” from Equifax which saw key figures in cybersecurity functions leaving the company, with some of them having expressed concerns around the data security culture within the company – a former VP of data quality at Equifax echoed this sentiment that “it bothered me how much access just about any employee had to the personally identifiable attributes” (Michael et al., 2017). The reporting on this case details many factors including human elements, technical and non-human elements, international and institutional elements, socio-cultural elements, global issues, and temporal elements that contributed to the overall situation; these are detailed in the ordered situational map in Table 1.

Table 1: Ordered situational map for the Equifax data breach.

<table>
<thead>
<tr>
<th>Heuristics</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human elements (individual + collective)</td>
<td>Richard Smith (then-CEO); Attacker; 145 million and 15 million people in US and UK affected; Nike Zheng; “entry crew”; Shell Crew; developer who knew about struts; senior manager; employee meant to apply the patch.</td>
</tr>
<tr>
<td>Non-human elements</td>
<td>Apache Struts; expired SSL certificate; Metasploit; misconfigured Intrusion Detection System; China chopper malware; Moloch; patch management; GTVM team; CVE-2017-5638; input sanitation, 8500 unresolved vulnerabilities; dedicated Security Operations Centre.</td>
</tr>
<tr>
<td>Political economic elements</td>
<td>Data Protection Act 1998; EUR 500,000 fine; Equifax and Mandiant dispute; departure of key staff i.e., “talent exodus”;</td>
</tr>
<tr>
<td>Discursive constructions of actors</td>
<td></td>
</tr>
<tr>
<td>International / institutional elements</td>
<td>US-CERT; Information Commissioners’ Office; Apache; China; United States; United States Federal Trade Commission; nation state actor.</td>
</tr>
<tr>
<td>Major contested issues</td>
<td>Attribution; “blame for the breach”.</td>
</tr>
<tr>
<td>Local to global elements</td>
<td>Tighter regulation; espionage.</td>
</tr>
<tr>
<td>Socio-cultural elements</td>
<td>Concern about data culture; failure to renew SSL certificate.</td>
</tr>
<tr>
<td>Symbolic elements</td>
<td></td>
</tr>
<tr>
<td>Popular and other discourses</td>
<td></td>
</tr>
<tr>
<td>Other empirical elements</td>
<td>War exercises; breach scenarios; “not credit card theft”; “get as much data as you can” play; lack of inventory management; patch management;</td>
</tr>
<tr>
<td>Spatial and temporal elements</td>
<td>10-month certificate expiry; 5 months instead of 48 hours to patch; 76 days’ time to discovery.</td>
</tr>
</tbody>
</table>

All these factors are part of the effects dynamics and complex interplay that culminated in the data breach incident. Invariably different factors had varying levels of influence on the outcome, however, specific interactions can be mapped across many of the elements in the situation through relational mapping. Relational mapping is a technique within the situational analysis suite of tools that shifts the focus from the elements in a situation to the relations between those elements. In relational mapping, each element is taken in turn to consider how it interacts with the other elements, thus explicating the complexity of the effects flows within the situation. For the Equifax case study, the relational mapping is undertaken on two elements, the Apache Struts vulnerability (i.e., CVE-2017-5638) and vulnerability scans undertaken to surface vulnerabilities across the Equifax systems – see Figure 2 and Figure 3 respectively.
Figure 2: Relational mapping for the Equifax data breach centred on the exploited vulnerability (CVE-2017-5638)

Figure 3: Relational mapping for the Equifax data breach centred on “vulnerability scans” element.

Presented in Figure 2, the discovery of the Apache Struts vulnerability, which allowed Object Graph Navigation Language (OGNL) injection attacks by Nike Zheng, started off the chain of events that ultimately led to the data breach. While it could be argued that other vulnerabilities could have been compromised and other vectors used, given that many other vulnerabilities had been noted on previous vulnerability scans on the Equifax systems, this vulnerability that was later assigned CVE-2017-5638 and a CVSS score of 10 was the necessary element in this situation.

The ability of organizations to effectively execute patch management functions was an important element of reducing the risk associated with CVE-2017-5638 vulnerability. In the case of Equifax, the failure of the CVE-2017-5638 alert from the US-CERT, which was relayed inside Equifax through the GTVM mailing list, to reach the employee responsible for patching, meant that their systems remained vulnerable. This is the vulnerability that was exploited by the initial access “entry crew”, which later handed over to a more sophisticated threat actor. The ability to detect and contain the attack was impaired by the ineffective (due to expired SSL certificate)
intrusion detection systems, poor inventory management practices, and internal communication and operational challenges.

The relations between the CVE-2017-5638 vulnerability and the different elements in this situation contributed in varying degrees to how the situation unfolded. Similar effect dynamics can be noted between vulnerability scanning function in Equifax and many other elements in the situation as depicted in Figure 3.

By focusing on the whole situation as the unit of analysis, the situational analysis of the Equifax incident highlights multi-dimensional elements and socio-technical assemblages that are at play in this situation. It shows why simplistic attributions of the incident, to technical factors or to the failure of the one employee who was responsible for patching vulnerabilities, are incomplete and provide an unsatisfactory explanation. This highlights the need for solutions that are commensurate with the level of complexity in such incidents and provides an opportunity for exploring relevant solutions from other domains.

3. Socio-Technical Approaches From Safety Management

There are several frameworks for conceptualizing adverse incidents that have been used extensively in domains such as safety management. Most of them are based on three accident philosophies, chain of events models, epidemiological accident models, and systems accident models (Y. Zhang et al., 2022).

The first is the Domino model which is based on the notion of cascade of events or causality effects; that events linearly lead to other events. The original model defined five distinct dominos associated with the worker’s social environment, human factors that lead to faults, an unsafe human act or technical condition, the accident, and the consequences of the accident (Heinrich, 1941). The causality effects in this model understood injuries (or incidents) as being caused by accidents which are caused by unsafe acts or technical conditions which are a result of a human factors (or fault of the person) that are shaped by the social environment.

Figure 4: The cybersecurity domain from the Domino perspective

This model is relevant for the cybersecurity domain because several cybersecurity solutions have similar linear causality mechanisms, for example, using causal modelling to investigate cybersecurity incidents, causal reasoning for malware detection, and employing causality analysis and system provenance graphs to classify ransomware types (Abel et al., 2018, 2020; Mei et al., 2021; H. Zhang et al., 2016). However, beyond accounting just for technical factors, the model identifies human factors which impact on threat conditions which are caused by social environment factors. Applying the model to the Equifax incident, the data breach (incident) could be understood as a result of the unpatched Apache Struts system (unsafe technical condition) which was due to the failure of the employee who was responsible for patching vulnerabilities (fault of the person) which was caused by social environment factors, including the organizational culture at Equifax at the time. However, while this analysis has some merit, the overarching criticism of the Domino model is that it is too simple and therefore leads to simplistic and short-sighted solutions that seek to attribute failure to simple causes.

The second model, the Swiss Cheese model, is based on the understanding of systems as comprising linear interactions, with latent conditions and dependencies (see Figure 5). The complexity of systems from this perspective means that understanding problems entails identifying a combination of conditions that lead to specific systems' outcomes; and addressing problems is a matter of strengthening the layers of barriers that mitigate those problems (J Reason et al., 2006; Reason, 1990). Variations of this model (i.e., Mark II and Mark III) identified different layers or barriers including organizational, workplace environment, and individual elements, as well as the flow of weaknesses from the organizational factors, workplace factors and unsafe
individual acts, but also the long-lasting gaps and weaknesses that are associated with latent conditions within organizations (J Reason et al., 2006).

**Figure 5: Framing cybersecurity domain from the Swiss cheese perspective (adapted from (J Reason et al., 2006))**

The basic elements of hazards, defences, and losses in this model directly map to the cybersecurity language of threats and attacks, controls and countermeasures, and harms and impact respectively. It also captures the notion of having different types of barriers, such as physical controls, administrative controls, and technical controls, as well as having different phases such as detect, deny, disrupt, degrade, and deceive, as different layers of cybersecurity defences (Hutchins et al., 2011). Considering the Equifax incident from the Swiss cheese perspective, it becomes apparent that that the attack exploited weaknesses in several security controls and systems (i.e., holes in the cheese slices). There were weaknesses in inventory management, patch management, and intrusion detection systems. However, there were also human, workplace, and organizational latent factors that contributed to the incident.

**Figure 6: Function description in FRAM**

The last framework considered in this paper is the functional resonance analysis method (FRAM) which is an approach for understanding and modelling the functional behaviour of complex socio-technical systems. This method was developed to address the limitation of deterministic, linear, and probabilistic approaches for understanding such systems (Hollnagel, 2012; Patriarca et al., 2020). FRAM describes systems in terms of the interactions between interrelated functions of the system, specifically in terms of how outputs (O) from functions interact with aspects of the downstream functions. These specific aspect include: input (I) – which is processed or transformed by a function and which initiate a function, preconditions (P) – that need to be in place for a function to occur, resources (R) – which are execution conditions or resources needed or consumed to produce the output, time (T) – which set the temporal constraints on the function, and control (C) – which control and monitor the function for acceptable performance (see Figure 6).
The FRAM approach allows for analysis of complex socio-technical systems by focusing on the relations between functions and the resonance that arises from functional coupling. The utility and relevance of FRAM for the cybersecurity domain is illustrated in detail below by applying it to the Equifax data breach incident.


It is outside the scope of this paper to model the full extent of the functional dependencies between all the elements in the situation (see Figure 1). Therefore, the analysis focuses on the one element that the then Equifax CEO attributed the incident to – the failure to patch the vulnerable Apache Struts system. Further discussion is on the functions that are internal to Equifax and those that are the source of significant performance variability.

Table 2: FRAM analysis and representation of the <patch a vulnerability> function

<table>
<thead>
<tr>
<th>Name of function</th>
<th>Patch a vulnerability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>This is the process of applying the software update provided by Apache to mitigate the Apache Struts vulnerability (CVE-2017-5638)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Description of aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>The patch instruction has been received by the responsible employee</td>
</tr>
<tr>
<td>Output</td>
<td>A patched system that can be operated securely</td>
</tr>
</tbody>
</table>
| Precondition | - The availability of the patch from Apache  
               | - An asset inventory that would indicate the presence of affected systems  
               | - The vulnerability being active on their systems.          |
| Resource   | The actual software update to be used to patch the vulnerability |
| Control    | The severity level of the vulnerability (CVSS score) that influences how quickly patches are applied |
| Time       | Not described                                              |

4.1 Analysis of Key Functions

The <Patch a vulnerability> function is undertaken by applying software updates to remediate a risk associated with a known vulnerability and is related to four other key functions, as illustrated in Table 2 and Figure 7. The output of the <Patch a vulnerability> function is a patched Apache Struts system that is not vulnerable to the OGNL injection (i.e., CVE-2017-5638) vulnerability. The effective execution of this function contributes to the secure operation of the Apache Struts system, which, for the scope of this analysis, is a Background and Exit function. Patching a vulnerability is triggered by a patch instruction which is an output of the <Communicate the patch> function. The two key Preconditions for patching a vulnerability are the availability of the patch, which is an output of the <Develop a patch> function, and the vulnerability being active on the organization’s systems, which is an output of the <Undertake vulnerability scan> function.

Another function of relevance in this case is the <Maintain asset inventory> function which involves identifying the digital assets owned by the organization and documenting their cybersecurity status. The Output from the <Maintain asset inventory> function is the actual asset inventory which forms the basis for further asset management and cybersecurity risk management processes.

The other function of interest is the <Communicate the patch> function, which is a condensed function that represents the cascade of external and internal communication that ensued from when the patch was made available to the patch being applied. As will be discussed further below, these series of communications should have triggered the application of the patch; hence why the output of this function is the Input to the <Patch a vulnerability> function.

The next step in analysing this case is to identify performance variability across the different functions and to map how that variability cascaded to the other functions leading to the circumstances of having an unpatched Apache Struts server.
4.2 Performance Variability Analysis

The failure to achieve the desired outcome (i.e., the <Operate systems securely> function) in the FRAM instantiation of the Equifax incident in Figure 7, can be traced to several elements of internal and external performance variability along the technical, human, and organizational dimensions across the key functions.

The internal organizational policy at Equifax required critical vulnerabilities to be patched within 48 hours, however, the Struts vulnerability was only patched after 5 months (Miyashiro, 2021). The major sources of performance variability that contributed to this are within the <Communicate the patch> and <Maintain asset inventory> functions. Once communication was received from US-CERT about the Apache Struts vulnerability, the information was relayed to 400 individuals on the GTVM mailing list. The key human function performance variability was the fact that the senior manager who supervised the lead developer (who knew about Apache Struts) and who received the GTVM alert, failed to relay the information to the said lead developer. There is no clear indication of the specific personal circumstances that explain this failure to relay the information. However, there is evidence that organizational climate at the time meant that there was lot of pressure on the cybersecurity functions and that “internally security was viewed as a bottleneck”, which could have contributed to this oversight (Michael et al., 2017; Miyashiro, 2021).

While the analysis has focused on functions related to patching vulnerabilities, there is evidence of variability across many other functions that contributed to the adverse impact of the data breach. For example, the technological performance variability associated with failure of the intrusion detection systems and the human function variability associated with the configuration of the vulnerability scanners.

The functional resonance analysis of the Equifax incident illustrates how internal and external performance variability across the technological, human, and organizational dimensions led to the adverse impacts experienced in this incident. This level of analysis provides a nuanced perspective on the functional dependencies in this situation and accounts for the inherent complexity.

5. Discussion and Conclusion

Complex socio-technical systems, which are typical in the cybersecurity domain, are necessarily intractable and always underspecified. As such, it is impossible to completely prevent adverse incidents in cybersecurity because they are not fully known (e.g., zero-day vulnerabilities), because of the vast and complex attack surface, because of performance variability across functions (e.g., accidental insiders), and because of the functional resonance within STS that could escalate normal performance variability into failure and adverse incidents.
As illustrated in the case of the Equifax incident, the assessment that the data breach was associated with an unpatched Apache Struts system is true and valid, however, the assertion that it could be attributed to the one individual who was responsible for patching the systems, provides an overly simplistic perspective on this situation. The functional resonance analysis undertaken above provides a more nuanced perspective on the dependencies between the different functions both within Equifax and outside the company, that led to the observed outcomes.

Given this complexity, the importance of resilience engineering in cybersecurity becomes apparent. Resilience is the systems’ capability for functional persistence and adaptation during adverse incidents, and it is typically framed in terms of the prepare, withstand, recover, and adapt phases (Kott & Linkov, 2019). However, the means towards resilience differ depending on how systems are conceptualized. For example, from a dominos conceptualization, failures are a cascade of initial failure conditions, and therefore resilience is about strengthening the robustness of the individual elements in the system (i.e., robustness of the individual dominos); from the Swiss cheese perspective, failures are an alignment of weaknesses in barriers, and therefore solutions are about putting in place sufficient barriers to limit propagation of threats; from the FRAM perspective, failures are the result of propagation of variability across functions, and therefore resilience is about dampening performance variability across systems functions. All these different approaches to resilience solutioning are relevant in cybersecurity because, as a complex STS, it compromises both the tractable technical and the intractable human and social dimensions (Ebert et al., 2023).

The analysis of the Equifax incident shows that simple approaches and explanations fail to account for the diverse set of elements and complex effect dynamics at play in such situations. The situational analysis surfaced the myriad of factors, including human, non-human, cultural, economic, institutional, and global, that are implicated in the incident. Further the functional resonance analysis method showed the functional dependencies and cascade of performance variability between the different elements in this situation. Despite some of its weaknesses, including the need for adaptation for the cybersecurity domain, FRAM as a socio-technical systems analysis method holds potential for holistic cybersecurity risk assessment and incident investigation.

If there is any lesson from the history of computer security it is that “hasty and simplistic solutions, while briefly satisfying, [are] not only likely to prove ineffective but also to make problems even worse” (Warner, 2012). While there is merit and utility to some of the simplistic cybersecurity solutions, for example, those that assume that single events cause incidents (e.g., root cause analysis), or that threat events flow linearly through systems, this paper posits that it is necessary to understand cybersecurity threats and solutions holistically along with their complex socio-technical systems assemblages.

References