Signalling Cyber Deterrence Through D3FEND

Kimmo Halunen[1,2] and Maria Keinonen[1]
[1]University of Oulu, Finland
[2]National Defence University, Helsinki, Finland

kimmo.halunen@oulu.fi
maria.keinonen@mil.fi

Abstract: States employ cyber deterrence strategies to safeguard their sovereignty in cyberspace. Cyber deterrence encompasses various means to prevent serious cyberattacks. This multifaceted approach incorporates various instruments of state power, including diplomatic, informational, military, economic and legal mechanisms. While all these instruments contribute to a state's overarching deterrence strategy in cyberspace, cyber-specific means offer the most rapid deployment options for countering cyberattacks. The challenge lies in credibly signalling cyber capabilities while preserving their secrecy and effectiveness. This challenge can be countered by carefully curating disclosed information, thereby maintaining the state´s strategic advantages and operational integrity. This research examines the technical implementation of deterrence signalling through a concrete example. By analysing the MITRE D3FEND framework, we aim to demonstrate practical application of cyber deterrence signalling and bridge theoretical concepts with operational cybersecurity practices. The MITRE D3FEND framework is a tool designed to describe cybersecurity countermeasure components and capabilities, and relationships between these elements. The research question posed is whether this framework can be used to signal cyber deterrence. This study evaluates the D3FEND framework's categories to determine which features can be signalled without compromising their effectiveness. Through qualitative content analysis, we develop evaluation criteria based on academic cyber deterrence literature. Each category of the D3FEND framework is methodically assessed against the evaluation criteria, to identify the signalling potential of the framework. The main findings of the study show that, of the seven categories of the D3FEND framework, the "Harden" category contains the most elements that can be used in cyber deterrence signalling, while the "Model" and "Deceive" categories have the fewest. The evaluation helps discern not only the elements to be signalled, but also those aspects of the defence, the exposure of which must be avoided. This research contributes to the academic discourse on cyber deterrence by elucidating the technical aspects of deterrence signalling, thereby offering a novel approach to bridging theoretical frameworks with practical cybersecurity implementations.

Keywords: Cyber deterrence, Cyber defence, MITRE D3FEND framework, Deterrence signalling

1. Introduction

The state signals cyber deterrence to secure its sovereignty in cyberspace. Cyber deterrence can be defined as all the means by which the state protects cyberspace to prevent the most serious cyberattacks. These means include all instruments of state power, including diplomatic, information, military, economic and legal means (Burton, 2018). Although all of these are used as part of the state's comprehensive deterrence strategy to protect cyberspace alongside physical domains, cyber means are often the fastest available to be implemented against attacks in cyberspace.

Signalling cyber deterrence is a difficult issue given the nature of digital assets and cyber capabilities, as well as the state´s aim to keep their cyber capabilities secret (Taddeo, 2018). However, this does not eliminate the need for cyber deterrence signalling. It is necessary to be able to signal cyber capabilities credibly without jeopardising their existence and use. One way to do this is to select the details that can be revealed about the state´s cyber capabilities and actions (Navicky and Tkach, 2023).

The MITRE D3FEND framework is designed to describe cybersecurity countermeasure components and capabilities and relationships between these elements. It is a tool that empowers professionals to comprehensively assess, analyse and architect integrated security solutions across organisational infrastructure (The MITRE Corporation, 2025). For cyber deterrence research, D3FEND provides a structured framework for evaluating signalling from a technical perspective. This article focuses on the technical implementations of the defensive measures presented in the D3FEND framework and explores the potential of this framework in signalling cyber deterrence. As the focus of D3FEND is on defence, the main focus of this paper is on deterrence by denial, which can be defined as convincing the aggressor that the attack will not achieve the desired effects and goals (Mazarr, 2021). Specifically, we will answer the following question: "*Can the MITRE D3FEND framework be used to signal cyber deterrence?*" To answer the research question, we investigate the D3FEND framework by evaluating its elements against criteria tailored for the purposes of this study. The aim is to define which elements of the framework can be signalled without compromising the defence or effectiveness of the technique.

2. Theoretical Background and Criteria for the Analysis

Signalling forms the core of classical deterrence theory and is based on how a state communicates its military capabilities and resolve to a potential adversary (Schelling and Schelling, 1966). Deterrence is preventive in nature and considers it a failure if an attack occurs (Mazarr and Goodby, 2011). In cyberspace, deterrence is not as absolute concept. Rather, cyber deterrence signalling aims to prevent the most serious cyberattacks and accepts the fact that cyberspace cannot be perfectly defensible due to the sheer number of attack vectors (Lindsay, 2015).

The most important factor in cyber deterrence signalling is to communicate the state's ability to withstand and recover from cyberattacks (Nye, 2017), as well as its willingness to engage in offensive activities to protect its cyberspace (Faesen et al, 2022). In addition to traditional, pre-emptive deterrence signalling, cyber deterrence can be signalled by real-life activities in cyberspace. For deterrence to be credible, it is essential to demonstrate the state's maturity as an actor in cyberspace through real-life actions (Borghard and Lonergan, 2021). In fact, unlike deterrence in general, it is challenging to make cyber deterrence credible without real-life actions in cyberspace (Wanic and Rowe, 2018). Covert cyber operations can be used to send a message to the adversary about the defender's determination to defend its own networks, and the risks and costs potential aggression would cause (Carson and Yarhi-Milo, 2017). However, since states generally strive to keep their cyber capabilities secret, this type of signalling is challenging. It is important to be able to reveal one's ability without jeopardising disposable capabilities or acquired footholds (Schneider, 2019).

The level of secrecy is also related to escalation control. One approach to handling attribution is to selectively disclose information. For instance, a state might choose to reveal details about the attack itself and the preventive measures taken, while deliberately withholding the attacker's identity (Baram and Sommer, 2019). Opting to withhold the attacker's identity, while sharing other details about a cyber incident, can help mitigate the risk of escalation. This approach is particularly effective when the communication emphasises factual evidence and events, avoiding potentially sensitive or controversial aspects of the attack (Lindsay, 2015).

Credibility is also an issue to be considered. For example, if a state blames another state with imperfect attribution, it might appear to be a weak actor in cyberspace (Edwards et al., 2017), but publicly and successfully executed attribution increases the credibility of a state (Egloff and Smeets, 2023). From an escalation perspective, openly promoting or endorsing the use of offensive cyber capabilities could escalate tensions between states, undermine global stability and lead to an increase in cyber hostilities between nations (Klimburg, 2020).

The academic literature mainly focuses on deterrence signalling through offensive tactics and attribution, and purely defensive methods have received less attention. This gap is addressed in this study by investigating the MITRE D3FEND framework through a tailored criterion, based on the theoretical background. We examine, how the MITRE D3FEND framework can be utilised to assess whether signalling a cyber capability might compromise its effectiveness, reveal new attack vectors, or redirect adversaries to alternative targets, thereby informing strategic decisions about cyber deterrence signalling.

The desired setup in cyber defence is such that one's own abilities can be signalled in a way that the defence is not compromised, but the potential aggressor is convinced of the defensive capability. For example, such cyber defence mechanisms can be implemented, which indirectly indicates to the threat actors that they are being closely monitored, and their identities are exposed (Chen, 2023). An unfavourable situation for cyber deterrence signalling is such that nothing can be revealed about one's own defence capability without jeopardising the cyber capabilities. The evaluation criteria used in this study is built between these two extremes. Areas to be evaluated are the visibility of the defensive action to the attacker and the effectiveness of the defence, if a certain action is revealed.

2.1 Visibility

From a visibility perspective, we defined four levels of visibility that occur at different stages of an attack. We simplified the phases of an attack into before, during and after the attack. These phases include all the activities for planning an attack: reconnaissance, gaining a foothold, and the actual attack with installing and using the cyber weapon, and concluding the attack by exiting the target system (see for example, The MITRE Corporation, 2024: ATT&CK Matrix for Enterprise).

“Invisible” means that the attacker cannot be sure if the defence technique is being applied unless the defender reveals this on purpose or inadvertently, or unless the attacker has some intelligence on the use of the

technique. These techniques are generally not openly revealed and in many cases are highly classified. Such invisible methods include models of the defender's systems, as architectural designs, asset inventories and deceptive techniques. For example, with host-based deception, slowing the progress of attackers and instilling a false sense of success could provide a significant defensive advantage (Shade et al., 2020).

"Visible during the attack" means that the attacker can see the defence technique when the attack is ongoing without compromising the defence. An example of this type of technique is access control, where the attacker can gain understanding of the different permissions in the system during the attack. However, this does not compromise the utility of this defence in subsequent attacks, although the attacker can of course adapt and choose other attack methods once they learn about the access control. This can be countered with attribute-based access control, which enhances security by proactively collecting unique attributes from entities involved in access requests, then randomly mutating original policies with new rules based on these attributes, creating an additional layer of protection against attacks even if the original attributes are compromised (Rubio-Medrano, 2017).

"Visible before the attack" means that the attacker can see the defence technique before launching the attack without compromising the defence. An example of this type of technique is encryption of network traffic outside the internal network. This can be seen from passive network traffic surveillance even before any active attack measures. However, this does not compromise the defence unless the actual implementation of the method is poor, e.g., weak encryption in the previous example. Real-world examples of this can be found in the many implementation-level issues of the TLS network encryption protocol's earlier versions (Böck et al., 2018; Adrian et al., 2015).

"Visible after the attack" means that the attacker can see the defence technique after a successful attack without compromising the current or future defence. A typical example of this type of technique is recovery from an attack through the use of backups. The adversary can usually see that they have been successful and see the recovery effort, but this will not impact any future efforts of using backups. The issue of backups is old, but there are recent results that improve the potential use of backups, especially against modern criminal threats (Oujezsky et al., 2023).

2.2 Effectiveness

From the effectiveness perspective, we defined three levels of effectiveness, which are tied to the safety of revealing defence technique without compromising its usability.

"Secret" means that even the existence of the technique cannot be revealed to the adversary without compromising its effectiveness. An example of this is honeypots. If an adversary knows that such techniques are employed, they can adapt their attacks in ways that detect such environments (Obaidat et al., 2021). Of course, sometimes adaptations can become weaknesses in the actual attack, as with the WannaCry malware. It contained a hardcoded mechanism to check if a specific domain was active, and if it received a response from the domain, it would terminate its operation (Mohurle and Patil, 2017).

"Existential revelation" in cybersecurity refers to techniques whose mere existence can be disclosed without compromising their effectiveness, but whose specific implementation details must remain confidential. This concept is prevalent in many cyber defence strategies. For instance, acknowledging the presence of a firewall protecting a network typically doesn't undermine its efficacy. Even disclosing the firewall's brand, model or software version may not necessarily compromise the system, unless there are known vulnerabilities associated with those specifics. However, revealing sensitive information, such as firewall rules, administrative passwords or other crucial configuration details, can significantly weaken or completely nullify the defence. The field of firewall bypass research has been a long-standing area of interest in cybersecurity, with recent advancements continuing to emerge in this domain (Emmanuel et al., 2021).

"Detailed revelation" means that the defence is effective even if details of its implementation and positioning in the protected system are revealed. An example of this could be a cryptographic system that is based on open standards. This should offer protection even if the adversary knows which encryption method is used and where, if the required secret key is not revealed (Taran, Rezaeifar and Voloshynovskiy, 2018).

The criteria are presented in **Table 1**.

Table 1: Criteria for analysing the D3FEND framework

Visibility	Rate	Explanation
V1	*Invisible*	The attacker cannot be sure if this technique is being applied unless the defender reveals this on purpose or inadvertently, or unless the attacker has intelligence on the use of the technique
V2	*Visible during the attack*	The attacker can see this technique while attacking, without compromising the defence
V3	*Visible before the attack*	The attacker can see this technique before launching the attack without compromising defence
V4	*Visible after the attack*	The attacker can see this technique after an attack without compromising defence
Effectiveness	Rate	Explanation
E1	*Secret*	The existence of this technique cannot be revealed to the adversary without compromising its effectiveness
E2	*Existential revelation*	Revealing the existence of this technique does not affect its effectiveness, but revealing details about the implementation of it could make it ineffective
E3	*Detailed revelation*	This technique is effective even if details of its implementation and positioning in the protected system is revealed

3. Methodology and Results

We analysed the different techniques presented in the D3FEND framework with the criteria presented in **Table 1**. The criteria consider the technological implementations of the D3FEND techniques and the implications that these have on deterrence signalling. We used the version 1.0.0 of the D3FEND matrix in our analysis. It is important to note that the framework is still under constant development and can change in many ways over time.

The version 1.0.0 of D3FEND contains 718 artefacts that are divided into seven different categories: *Model, Harden, Detect, Isolate, Deceive, Evict* and *Restore*. We analysed each of the subcategories under these with our taxonomy. The results of this analysis are presented in **Table 2**. The code refers to the description according to **Table 1**. Only the majority assessment for each category is presented in the table. One star (*) after the code means that one subcategory of the main category has been evaluated as different. Two stars (**) after the code means that two or more subcategories have been evaluated as different. Otherwise, the assessment in question is the same for all subcategories.

Table 2: Results of the content analysis

Model	Visibility	Effectiveness	Explanation from the D3FEND framework
Asset Inventory	V1	E2	Asset inventorying identifies and records the organisation's assets and enriches each inventory item with knowledge about their vulnerabilities.
Network Mapping	V3**	E1**	Network mapping encompasses the techniques to identify and model the physical layer, network layer and data exchange layers of the organisation's network and their physical location, and determine allowed pathways through that network.
Operational Activity Mapping	V1	E2	Operational activity mapping identifies activities of the organisation and the organisation's suborganisations, groups, roles and individuals that carry out the activities and then establishes the dependencies of the activities on the systems and people that perform those activities.
System mapping	V1*	E2*	System mapping encompasses the techniques to identify the organisation's systems, how they are configured and decomposed into subsystems and components, how they are dependent on one another, and where they are physically located.
Harden	Visibility	Effectiveness	Explanation from the D3FEND framework

Model	Visibility	Effectiveness	Explanation from the D3FEND framework
Agent Authentication	V3*	E3	Agent authentication is the process of verifying the identities of agents to ensure they are authorised and trustworthy participants within a system.
Application Hardening	V2**	E3**	Application Hardening makes an executable application more resilient to a class of exploits, which either introduce new code or execute unwanted existing code. These techniques may be applied at compile-time or on an application binary.
Credential Hardening	V3*	E3*	Credential Hardening techniques modify system or network properties in order to protect system or network/domain credentials.
Message Hardening	V3	E3	Email or Messaging Hardening includes measures taken to ensure the confidentiality and integrity of user-to-user computer messages.
Platform Hardening	V2*	E2**	Hardening components of a Platform with the intention of making them more difficult to exploit.
Source Code Hardening	V1	E2**	Hardening source code with the intention of making it more difficult to exploit and less error prone.
Detect	Visibility	Effectiveness	Explanation from the D3FEND framework
File Analysis	V2**	E2	File Analysis is an analytic process to determine a file's status: for example, virus, trojan, benign, malicious, trusted, unauthorised, sensitive, etc.
Identifier Analysis	V3*	E2*	Analysing identifier artefacts, such as IP address, domain names or URL(I)s.
Message Analysis	V2	E2	Analysing email or instant message content to detect unauthorised activity.
Network Traffic Analysis	V1**	E2	Analysing intercepted or summarised computer network traffic to detect unauthorised activity.
Platform monitoring	V1**	E2**	Monitoring platform components such as operating systems software, hardware devices or firmware. Platform monitoring consists of the analysis and monitoring of system-level devices and low-level components, including hardware devices, to detect unauthorised modifications or suspicious activity.
Process Analysis	V1*	E2	Process Analysis consists of observing a running application process and analysing it to watch for certain behaviours or conditions which may indicate adversary activity. The analysis can occur inside of the process or through a third-party monitoring application. Examples include monitoring system and privileged calls, monitoring process initiation chains, and memory boundary allocations.
User Behaviour Analysis	V1**	E2**	User behaviour analytics ("UBA"), as defined by Gartner, is a cybersecurity process about detection of insider threats, targeted attacks and financial fraud. UBA solutions look at patterns of human behaviour, and then apply algorithms and statistical analysis to detect meaningful anomalies from those patterns-anomalies that indicate potential threats.' Instead of tracking devices or security events, UBA tracks a system's users. Big data platforms are increasing UBA functionality by allowing them to analyse petabytes worth of data to detect insider threats and advanced persistent threats.
Isolate	Visibility	Effectiveness	Explanation from the D3FEND framework
Access Mediation	V2**	E2*	Access mediation is the process of granting or denying specific requests to 1) obtain and use information and related information processing services; and 2) enter specific physical facilities (e.g., federal buildings, military establishments, border crossing entrances).
Access Policy Administration	V2*	E2*	Access policy administration is the systematic process of defining, implementing and managing access control policies that dictate user permissions to resources.

Model	Visibility	Effectiveness	Explanation from the D3FEND framework
Execution Isolation	V3**	E2	Execution Isolation techniques prevent application processes from accessing non-essential system resources, such as memory, devices or files.
Network Isolation	V2*	E2	Network Isolation techniques prevent network hosts from accessing non-essential system network resources.
Deceive	Visibility	Effectiveness	Explanation from the D3FEND framework
Decoy Environment	V4*	E1	A Decoy Environment comprises hosts and networks for the purposes of deceiving an attacker.
Decoy Object	V4*	E1	A Decoy Object is created and deployed for the purposes of deceiving attackers.
Evict	Visibility	Effectiveness	Explanation from the D3FEND framework
Credential Eviction	V3	E2	Credential Eviction techniques disable or remove compromised credentials from a computer network.
Object Eviction	V3*	E2*	Terminate or remove an object from a host machine. This is the broadest class for object eviction.
Process Eviction	V3	E2	Process eviction techniques terminate or remove the running process.
Restore	Visibility	Effectiveness	Explanation from the D3FEND framework
Restore Access	V1**	E3	Restoring an entity's access to resources.
Restore Object	V1	E3	Restoring an object for an entity to access. This is the broadest class for object restoral.

In **Table 2**, many subcategories within a given category exhibit very similar characteristics in our taxonomy. Thus, the whole category was given a single evaluation with our taxonomy. There are some categories where the techniques vary more and a single evaluation for the whole category is not necessarily wholly representative of the signalling capabilities of that category. These have been marked with stars.

Optimal cyber defence signalling occurs when defenders can openly reveal their methods, allowing attackers to observe these defences during attacks without the need for defenders to alter their strategies. Thus, techniques with visibility category V3 and effectiveness category E3 would be most beneficial for deterrence signalling. Luckily, many of the techniques in the Harden category seem to fall within these parameters. For example, all the techniques for message hardening are V3 and E3. In addition, credential hardening and agent authentication have this same evaluation with only few exceptions in the subcategories. Also, the Evict category has V3 and E2 in all categories, with only a single exception. Thus, these two categories would be most effective in deterrence signalling according to our analysis.

From the other perspective, any technique that is E1 (secret) and V2 or V3 can be considered fairly ineffective defence for deterrence signalling, as the attacker can gain information on this defence in a phase of the attack where they can easily adapt, and this makes the defence ineffective. In our analysis, there are some techniques in the Model and Harden categories that fall into these categories, such as Access Modelling and Local File Permissions. The most ineffective category for cyber deterrence signalling, based on our analysis, is the Deceive category, where all the techniques need to be kept secret to be effective.

Our analysis shows that there are many categories where it is important to check each subcategory and technique separately. For example, in the Isolate and Detect categories there is a lot of variation within the subcategories, and thus each singular technique needs to be assessed more thoroughly before being able to use these for deterrence signalling. This might require going deeper into the technical details of the technique and its specific implementation to determine whether that technique can be used in signalling or not.

It is important to note that even if a technique is not necessarily very effective in signalling cyber deterrence, it can still be valuable in more traditional organisational cyber risk management and defence building. This means that, although the defence cannot be used to signal deterrence, it can be effective for preventing some of the cyberattacks. Even if the attacker could bypass the defence once it has seen it, many attackers might not want to spend the effort in attacking the system, at least in the criminal setting, where some reports state that speed

is sometimes emphasised (European Union Agency for Cybersecurity, 2024). In the national security setting, the incentives are different, and there the attacker may be more motivated in adapting around defences.

Based on our results, the three most potential categories for signalling cyber deterrence are Harden, Isolate and Evict. Our recommendation is to first investigate these categories for suitable defensive methods to signal cyber deterrence. As the D3FEND framework contains a lot of different methods and each method can be evaluated with our taxonomy, we also recommend looking into other categories and finding the precise methods within that category to signal cyber deterrence.

4. Conclusions

In this paper, we analyse the possibility of signalling cyber deterrence from the perspective of the MITRE D3FEND framework. The framework has recently matured to the version 1.0.0 and can be understood as one key building block in designing and implementing cyber defence methods and strategies in many organisations. This makes it also a good tool for evaluating different defence methods for deterrence signalling.

Traditionally, cyber capabilities are very much kept secret, and this makes the signalling of cyber deterrence somewhat difficult. However, our results show that there are many different techniques in the D3FEND framework that can be used to signal deterrence by denial in a way that does not undermine the effectiveness of the defence. Of course, this does not apply to all techniques in the framework, and in many cases, the details of the implementation and the details of signalling the deterrence can have an impact on this analysis.

Our analysis shows that there is potential for utilising the D3FEND framework in planning and executing cyber deterrence. There are clearly some categories that are more amenable to open signalling (e.g., Harden) and then some categories that are not suitable for signalling (e.g., Deceive). The latter can still be very effective in building the actual defence.

We see that our analysis could have at least two new venues for further research. From the technical viewpoint, it would be interesting to study how the technical details of the implementations of different D3FEND methods affect their usability for deterrence signalling. From a policy perspective, further research is needed in order to understand the effectiveness of these different defensive methods from D3FEND in deterrence building and signalling. The D3FEND framework is not yet mature, but it has the potential to develop into an important tool for organisations to evaluate and utilise defensive measures in the cyberspace.

Ethical clearance was not required for the research.

References

Adrian, D., Bhargavan, K., Durumeric, Z., Gaudry, P., Green, M., Halderman, J. A., ... & Zimmermann, P. (2015) Imperfect forward secrecy: How Diffie-Hellman fails in practice. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security,* pp. 5-17.

Baram, G. and Sommer, U. (2019) Covert or not Covert: National Strategies During Cyber Conflict. *2019 11th International Conference on Cyber Conflict (CyCon),* Tallinn, Estonia, 2019, pp. 1-16, https://doi.org/10.23919/CYCON.2019.8756682

Borghard, E. and Lonergan, S. (2021) Deterrence by denial in cyberspace, Journal of Strategic Studies 46 (2021), pp. 534 - 569. https://doi.org/10.1080/01402390.2021.1944856

Burton, J. (2018) *Cyber Deterrence: A Comprehensive Approach?* Nato Cooperative Cyber Defence Centre of Excellence. https://ccdcoe.org/uploads/2018/10/BURTON_Cyber_Deterrence_paper_April2018.pdf.

Böck, H., Somorovsky, J., and Young, C. (2018) Return Of Bleichenbacher's Oracle Threat (ROBOT). In *27th USENIX Security Symposium (USENIX Security 18)*, pp. 817-849.

Carson, A. and Yarhi-Milo, K. (2017) Covert Communication: The Intelligibility and Credibility of Signaling in Secret, Security Studies, 26:1, pp. 124-156, https://doi.org/10.1080/09636412.2017.1243921

Chen, J. (2023) Vol. 22 No. 1 (2023). A New Interpretation of Integrated Deterrence: Physical and Virtual Strategies, Proceedings of the 22nd European Conference on Cyber Warfare and Security, pp. 116-123. https://doi.org/10.34190/eccws.22.1.1314

Edwards, B., Furnas, A., Forrest, S. and Axelrod, R. (2017) Strategic aspects of cyberattack, attribution, and blame. *Proceedings of the National Academy of Sciences*, vol. 114, no. 11, pp. 2825–2830, https://doi.org/10.1073/pnas.1700442114

Egloff, F. and Smeets, M. (2023) Publicly attributing cyber attacks: a framework, Journal of Strategic Studies, 46:3, pp. 502-533, https://doi.org/10.1080/01402390.2021.1895117

Emmanuel, O.I., Ayodele, A.A., Adebiyi, A.M. and Osang, B.F. (2021) Windows Firewall Bypassing Techniques: An Overview of HTTP Tunneling and Nmap Evasion. In: Gervasi, O., *et al.* Computational Science and Its Applications – ICCSA 2021.

ICCSA 2021. Lecture Notes in Computer Science(), vol 12957. Springer, Cham. https://doi.org/10.1007/978-3-030-87013-3_41

European Union Agency for Cybersecurity. (2024). ENISA Threat Landscape 2024: Cybersecurity Threat Trends., https://www.enisa.europa.eu/publications/enisa-threat-landscape-2024.

Faesen, L., Sweijs, T., Klimburg, A. and Tesauro, G. (2022) *The Promises and Perils of a Minimum Cyber Deterrence Posture: Considerations for Small and Middle Powers*. The Hague Centre for Strategic Studies, April 2022, ISBN/EAN: 9789492102997, pp 69-72.

Klimburg, A. (2020) Mixed Signals: A Flawed Approach to Cyber Deterrence. *Survival*, 62:1, pp. 107-130, https://doi.org/10.1080/00396338.2020.1715071

Lindsay, J. (2015) Tipping the scales: the attribution problem and the feasibility of deterrence against cyberattack, *Journal of Cybersecurity*, Volume 1, Issue 1, September 2015. https://doi.org/10.1093/cybsec/tyv003.

Mazarr, M. (2021) Understanding Deterrence., in, Osinga, F., Sweijs, T. (eds) *NL ARMS Netherlands Annual Review of Military Studies 2020*. NL ARMS. T.M.C. Asser Press, The Hague. https://doi.org/10.1007/978-94-6265-419-8_8.

Mazarr, M. and Goodby, J. (2011) Redefining the Role of Deterrence, in Hoover Institution on War, Shultz, R., Drell, G. & Goodby, J. *Deterrence: Its past and future: Papers presented at Hoover Institution*, November 2010. Stanford, Calif.: Hoover Institution Press.

The MITRE Corporation (2025). About the D3FEND Knowledge Graph Project. https://d3fend.mitre.org/about/ (Accessed 15th January 2025).

The MITRE Corporation (2024). ATT&CK Matrix for Enterprise. https://attack.mitre.org/ (Accessed 15th January 2025).

Mohurle, S., and Patil, M.M. (2017) A brief study of Wannacry Threat: Ransomware Attack 2017. *International Journal of Advanced Research in Computer Science,* 8, pp. 1938-1940.

Navicky, M. and Tkach, B. (2023) Cross-Domain Cyber Incidents and State Responses, in Billingsley, J. (ed), *Integrated Deterrence and Cyberspace, Selected Essays Exploring the Role of Cyber Operations in the Pursuit of National Interest*. National Defense University Press, Washington, D.C.

Nye, J. (2017) Deterrence and Dissuasion in Cyberspace. International Security 2017; 41 (3), pp. 44–71. https://10.1162/ISEC_a_00266

Obaidat, M.A., Brown, J., and Alnusair, A. (2021) Blind Attack Flaws in Adaptive Honeypot Strategies. *2021 IEEE World AI IoT Congress* (AIIoT), pp. 0491-0496.

Oujezsky, V., Novak, P., Horvath, T., Holik M. and Jurcik, M. (2023) Data Backup System with Integrated Active Protection Against Ransomware, *2023 46th International Conference on Telecommunications and Signal Processing (TSP)*, Prague, Czech Republic, 2023, pp. 65-69, https://doi.org/10.1109/TSP59544.2023.10197687

Rubio-Medrano, C.E., Lamp, J., Doupé, A., Zhao, Z., and Ahn, G. (2017) Mutated Policies: Towards Proactive Attribute-based Defenses for Access Control. *Proceedings of the 2017 Workshop on Moving Target Defense*.

Schelling T. and Schelling T.C. (1966) *Arms and Influence*. Yale University Press, London.

Schneider, J., (2019) Deterrence in and through Cyberspace, in *Cross Domain Deterrence*, Lindsay, J. and Gartzke, E. (eds) https://doi.org/10.1093/oso/9780190908645.003.0005

Shade, T., Rogers, A., Ferguson-Walter, K.J., Elsen, S.B., Fayette, D., and Heckman, K.E. (2020) The Moonraker Study: An Experimental Evaluation of Host-Based Deception. *Hawaii International Conference on System Sciences*.

Taddeo, M., (2018) The Limits of Deterrence Theory in Cyberspace. *Philos. Technol.* 31, pp. 339–355 (2018). https://doi.org/10.1007/s13347-017-0290-2.

Taran, O., Rezaeifar, S. and Voloshynovskiy, S. (2018) Bridging machine learning and cryptography in defence against adversarial attacks*. Computer Vision – ECCV 2018 Workshops: Munich*, Germany, September 8-14, 2018, Proceedings, Part II, pp. 267 – 279. https://doi.org/10.1007/978-3-030-11012-3_

Wanic, E. and Rowe, N. (2018) Assessing Deterrence Options for Cyber Weapons. *2018 International Conference on Computational Science and Computational Intelligence (CSC*I), Las Vegas, NV, USA, 2018, pp. 13-18, https://doi.org/10.1109/CSCI46756.2018.00011 .