Abstract—Threat modeling is an analysis activity aimed at eliciting viable and realistic security and privacy threats in the design of a software-intensive system. Threat modeling allows for a by-design approach, mitigating problems before they arise and avoiding later costly development efforts.

However, it mainly pays off in software construction approaches that rely on planned architectures, in which sources of threats can be anticipated beforehand. These axiomatic assumptions are, however, increasingly untrue in contemporary software development practices in which software systems evolve drastically in later stages. In addition, software-intensive systems are increasingly faced with uncertainty in their operational contexts, and these are nearly impossible to enumerate in early development stages.

In this article, we first present the idea of reflective threat modeling, which involves the automated derivation of architectural system models from run-time and operational system artifacts, providing the threat modeler with an accurate and workable run-time inspection view of the system. We then outline and motivate the potential of adopting threat analysis models as a basis for holistic and adaptive threat management through integration of adaptive security and privacy technologies. This will enable systems to autonomously respond to emerging threats by dynamically activating dedicated controls or via run-time reconfiguration.

Index Terms—Threat modeling, threat analysis, threat management, security, privacy, run-time reflection, architecture-centric adaptation

I. INTRODUCTION

Threat modeling [1], [2] involves the systematic analysis of potential security and privacy threats in the context of a specific system under design. It is considered an important cornerstone in the secure software development life-cycle (SDLC) [1], and it contributes to attaining the widely-advocated principles of security and privacy by design: instead of addressing problems and issues whenever they arise, a proactive and thorough assessment of security and privacy requirements in the early development stages will avoid costly, invasive engineering activities in later development stages.

Current threat modeling approaches [3], [4] are heavily shaped by their role in the development life-cycle: (i) they are to be performed manually, by analysts/requirements engineers in a workshop/brainstorm setting, (ii) they act upon an abstraction model of the system under analysis (such as Data Flow Diagrams (DFDs)) that is created in isolation and thus decoupled from implementation or run-time artifacts, (iii) they involve manual risk assessment which involves estimation of risk factors that both cannot be assessed precisely and quantitatively beforehand, and which rely on implicit information and background system knowledge

While existing threat modeling approaches are suited for determining the most relevant security and privacy threats in an a-priori, constructive context, their disconnect from further software development activities and the operational context of software-intensive systems has a number of significant disadvantages. In practice, when threat modeling is conducted, it is done as a one-shot operation, and the analysis is rarely revised in later development stages as it is perceived too time-consuming and costly.

In this article, we provide an in-depth discussion about what is required to integrate threat modeling and threat analysis into the operational context of a system, in two incremental steps.

In a first step, we discuss the opportunities and challenges related to the automated construction of system models through reflection and inspection, an approach that we call reflective threat modeling. We envision the (semi-)automated derivation of integrated and rich input models in which different threat sources such as design flaws, vulnerabilities, and actual incidents can be distinguished and potential threats enumerated and assessed. These inputs models are architectural in the sense that they distinguish explicitly between system design structure, deployment, and run-time events and interactions.

In a second step, we discuss the possibilities and challenges related to implementing adaptive threat management, in which the system itself performs automated threat analysis to identify emerging threats (e.g., when the system or its environmental context has drastically changed), and makes decisions

1For example, to assess the likelihood of a certain threat, the threat modeler must consider aspects such as attacker incentives and capabilities, the trust relations to third parties, the perceived strength of existing countermeasures, etc. In threat modeling practice, these elements and such rationale are kept implicit [5].
to proactively or reactively address threats, for example by employing adaptive security and privacy technologies [10]–[12] or enacting run-time system reconfigurations.

While approaches and technologies to adaptive security and risk assessment enjoy practical adoption [10], [13], they address specific threat types (e.g., focusing exclusively on fraud, anomaly or intrusion), and lack the holistic end-to-end system perspective that is inherent to threat modeling. As such, they are not suited for monitoring threat and risk evolutions across wider threat landscapes.

This article is structured as follows: Section II presents the necessary background, whereas Section III motivates this work. Next, Section IV and Section V respectively propose and discuss the research challenges of reflective threat modeling and adaptive threat modeling. Finally, Section VI concludes.

II. BACKGROUND

This section provides background information on threat modeling, on the integration of security in a development and operations context (SecDevOps), and on adaptive security approaches and enabling technologies.

A. Threat modeling

Threat modeling approaches have emerged in the early 2000s [1] as a means to systematically identify and evaluate security requirements. The most prominent elicitation approach is Microsoft’s STRIDE [2], yet different survey studies [3], [4], [14] show the range of available threat modeling approaches. Recently, the Threat Modeling Manifesto [15] was established to create a common understanding among practitioners and to articulate the core values and principles of threat modeling.

Figure 1 provides a generic overview of the different activities involved in threat modeling, starting from (activity 1) the establishment of an abstraction model of the system. Based on this, (activity 2) systematic threat elicitation approaches involve enumerating all theoretically viable threats, based upon reusable threat taxonomies, trees, vulnerability databases, or threat categories. As this step leads to large bodies of threats to be investigated, the subsequent step (activity 3) involves prioritizing and filtering, by determining which threats pose a significant risk to the system. For this, risk assessment models (e.g., FAIR [16]) can be applied [17] that take into account a wide array of risk components. Finally, for the most critical threats, countermeasures can be selected to mitigate the identified threats (activity 4).

Threat modeling has been successfully adopted in a wide variety of application domains, in cyber-physical systems (CPS) [18], automotive and mobility systems [19]–[22], large-scale enterprise environments [23]–[25], e-health systems [26] and even to validate the complex privacy-preserving architectures that are used in COVID-19 exposure tracking and contact tracing systems [27].

While DFDs [1], [2] are the most common system representations, there is in effect a wide range of system abstractions used in the context of threat modeling (e.g., Process Flow Diagrams [28], Information Flow Diagrams [27], DFD extensions [29]–[31]), or tool-specific representations [32]. For a more elaborate coverage of different secure design notations, we refer the reader to the literature [33].

Threat modeling is predominantly a manual task [15], [22] and in practice is used in the context of (i) educating and raising awareness about threats and common threat sources, (ii) increasing confidence in system concepts and early architectures, and (iii) performing fine-grained analysis at the level of DFD elements or interactions. Due to its heavy reliance on manual effort and the required degree of expertise, continuous threat modeling or frequent iteration of threat analysis activities are considered cost-inefficient [34].

Recently, advances have been made towards automation of threat modeling and threat analysis activities [3]. These efforts are focused on automated threat elicitation (activity 2) and automated threat prioritization and risk assessment (activity 3), e.g., as in SPARTA [30] and the work of Tuma et al. [29]. In addition, complementary attempts to bring threat modeling closer to the implementation phase have led to approaches such as pytm [35], with a code-based model, and threatspec [36], which allows developers to include annotations in code, from which system models can be derived for threat analysis.

B. SecDevOps

DevOps refers to the tight integration of development activities with the operational management of systems [37] and is associated with practices of agile development and continuous integration and deployment (CI/CD). These techniques involve extensive process automation, for example by embedding testing activities in production environments, quality assurance, automated and fine-grained deployment of individual updates or product features. By informing development activities with metrics and results obtained in a production or operational environment, problems can be identified and prioritized more accurately, and the time windows of iterative development can be shortened substantially.

Secured DevOps or SecDevOps [38], [39] involves dealing with novel cyber security risk that emerge in the DevOps context. In its strictest interpretation, it refers to addressing the security implications caused by automated and frequent deployment [40], for example by properly configuring the
target environment such as securing containers or virtual machines [41]. In other work, it is also used to refer to the integration of automated security development and operational practices into the actual CI/CD cycles, e.g., integration of automated security testing or scanning tools [42], [43].

C. Adaptive security and privacy

Approaches that are adaptive in nature are capable of identifying security- or privacy-related issues and reacting appropriately [10], [11], [11]–[13], [44]. Without aiming for exhaustiveness, the term refers to technology classes such as: adaptive firewalls [45], [46], adaptive incident or intrusion detection systems (IDS) [47], [48], adaptive fraud detection systems [49], [50], adaptive cryptography, and adaptive access control systems [51], [52].

For example, when an access request comes from outside of the corporate network or from a new device, an adaptive access control system can decide to require two-factor authentication, or when there is uncertainty about the credentials of a user, the adaptive system can decide to ask for extra credentials, such as biometric identifiers.

Other technology classes of relevance are not uniquely dedicated to security or privacy but represent key enablers for implementing and enacting adaptations in a more generic sense. These include technologies such as virtualization, dynamic container orchestration [53], [54], software-defined networking (SDN), and policy enactment frameworks [55].

III. Motivating case: Maritime system

Figure 2 presents a (strongly simplified) DFD of the system architecture of a maritime system, more specifically with emphasis on the communication subsystems deployed on a vessel used for navigation, management, and interaction with external entities (e.g., tracking of shipping containers).

Adopting threat elicitation approaches at the basis of the depicted DFD will yield a number of possible and relevant system threats, for example the threat of a malicious agent deliberately steering the vessel off course by tampering with the satellite communications or by spoofing the on-shore management systems to which the vessel remains in contact. However, the different threats and their risk profiles (likelihood and impact) will in reality depend strongly on the operational context of the system. For example, the attack surface that involves access via Port Management Systems (e.g., the Information Disclosure threat that involves a third party that wants to identify specific containers in the cargo in, for example, smuggling scenarios) only becomes relevant when the vessel is docked, whereas such threats are not applicable when the vessel is in open sea. Conversely, the threat of a malignant external entity falsely emitting distress signals via the Search and Rescue Transmission Systems (SART) to deviate the vessel (a spoofing threat) will not be applicable when the vessel is docked.

When performing threat analysis in an a-priori design context, the analyst is required to enumerate and anticipate the different operational contexts of the system, and as these contexts may change significantly over time (as in the motivating example), so will the corresponding threat landscapes. This is especially problematic in systems for which exhaustive enumeration of these operational context is cumbersome or even infeasible, and as a consequence, relevant threats may not be identified, prioritized nor mitigated correctly.

In addition, when the system itself changes or evolves over time, novel threats will not be identified nor mitigated unless the entire threat analysis exercise is repeated. Change scenarios are: new subsystems are activated that were not initially present in the analysis (e.g., the installation of on-deck wireless networks to enable personal communication of crew members), when new external entities emerge (e.g., new means to communicate to base stations, or new types of base stations), or simply, as the capabilities of attackers evolve over type (e.g., new vulnerabilities have been published or new types of attacks have been detected) so that threat types that were previously considered impossible have become actual issues.

IV. Reflective threat modeling

In this section, we motivate and outline a first necessary innovation in threat modeling related to how the system models that are used as the basis for threat analysis are obtained. Our motivation is based on two key problems with the current state of the art: (i) system models used in threat modeling lack expressiveness, (ii) system models created in early stages only convey design information which can become outdated once the system is in operation. We elaborate on both problems below.

a) Lacking expressiveness in system models: Data flow (DFDs) modeling involves a relatively simple and intuitive notation consisting of only five element types (processes, data stores, external entities, data flows and trust boundaries). The simplicity ensures easy adoption by diverse stakeholders and as such facilitates the creation of DFDs in workshop settings. The downside however is that it leads to system models that easily omit relevant information, e.g., DFD models typically model data flows but omit how exchanged information is encoded or which communication protocols are used. In addition, it leads to overloading of element types in terms of the type of system information that is conveyed. For example, a trust boundary may represent an organizational boundary, a communication boundary, a security control boundary, or a
physical boundary, but it may also be used to delineate a unit of computation (a system component), or a logical grouping of elements (e.g., to represent system elements deployed in a virtualized environment such as a distributed cloud).

The correct interpretation of DFD elements\(^2\) as such depends heavily on human disambiguation.

b) \textit{Low accuracy of system models}: When threat modeling is performed in the early stages of requirements engineering, the system model is \textit{prescriptive} in nature, i.e., it is indicative of how the system is expected to be materialized, but in practice, such a holistic view on the system is incomplete in these stages, especially in contemporary development practices that involve drastic and continuous evolution. When relying exclusively on early models, the final product often deviates (or has evolved drastically away) from the planned system.

A. Concept

In the core concept of \textit{reflective threat modeling}, we advocate extensively relying on methods and techniques that involve automated derivation of more expressive system models that leverage this information to create and maintain accurate and rich system models.

These system models are architectural in nature, in the sense that they encode information of distinct architectural viewpoints. We adopt the classification in viewpoints of Bass et al. [8] in terms of (1) a module view, (2) an allocation or deployment view, and (3) a process or client-server view, which we consider a valuable distinction to differentiate between design flaws, physical and infrastructural attacks, and run-time incidents respectively.

- \textbf{Module views} represent the design and development structure of a system. Contemporary practice is increasingly based on expressive models that encode a wealth of information about design structure and remain accessible at run time. Some examples are: data model or schema specifications, workflow descriptions or business process specifications. These can be complemented with techniques of static code analysis (e.g., dependency analysis) and the dedicated annotation-based approaches discussed above [35], [36]. Finally, methods of software architecture reconstruction (SAR) [56], [57] also take into account other development artifacts, such as code comments, information drawn from versioning systems, etc., to retroactively derive the development structure of the system.

- \textbf{Deployment and allocation views}: Distributed systems commonly rely on middleware platforms that require developers to define deployment abstractions (e.g., applications, (micro-)services), and this is the level of granularity at which distributed applications are deployed. This information is available through reflection and via platform-specific deployment descriptors. In addition, virtualization and orchestration systems [58], [59] employ techniques such as Software-defined networking (SDN) [60], [61] and involve expressive deployment descriptors from which valuable deployment information can be obtained.

- \textbf{Client-server/process views} represent runtime interactions, for example between active objects, threads and processes. These views contribute information about frequency of interactions, provide concrete examples of data elements being processed, run-time user sessions, etc. The construction of these views can leverage upon information provided in audit trails, system logs and session management, but also may rely on dynamic analysis of application execution (e.g., based on call graphs to identify run-time interactions, or taint analysis [62]). In addition, intrusion or fraud detection systems are capable of detecting security incidents in an operational system, and these represent threat instances that may trigger system-wide threat re-evaluation.

B. Research agenda

We define an approach that leverages the aforementioned sources of architectural information to construct or refine system models as a \textit{reflective threat modeling} approach and discuss the required innovations and main research challenges.

- \textbf{Architecture description languages} We argue that dataflow-centric abstractions (DFDs) lack the expressiveness to document and maintain the architectural distinction between (i) design structure, (ii) run-time interactions, and (iii) deployment configurations. As such, we envision adopting or tailoring existing software architecture specification or description languages [63], [64] for the specific purpose of threat analysis, as these allow defining a system from multiple orthogonal and complementary viewpoints [65]. A first problem inherent to adopting more expressive architectural models is related to internal consistency, i.e., ensuring that different views on the system remain consistent [66].

- \textbf{Heterogeneity of architectural information sources}. The potential information sources listed above are heterogeneous in terms of (i) the type of architectural information provided, (ii) their technology context, (iii) how accessible they will be through run-time reflection and inspection. To deal with heterogeneity, we envision a reflection framework that is extensible and capable of integrating these different sources into a comprehensive and complete system model. Also required are facilities to assess the extent to which automatically-generated system models can be considered complete or suited for the purpose of threat analysis.

- \textbf{Derivation logic}. Casting the obtained system information obtained via system inspection and reflection into abstractions suited for threat analysis is not straightforward, not in the least because these target abstractions are not yet well-known or -understood. From a technical perspective, dedicated translation mechanisms and algorithms will be required to raise the level of abstraction or granularity, e.g., by employing clustering algorithms in call or execution graphs obtained through system inspection to identify more coarse grained ‘components’ or ‘processes’. Model-to-model transformation techniques [67] will be required to select and convert the relevant system information encoded

\(^2\)While we focus predominantly on the DFD notation in the above for the sake of simplicity in the argumentation, the same applies to different design notations such as those discussed in Section II-A.
in descriptors and domain-specific languages into the target abstractions. Although these mechanisms will be heterogeneous in nature, they should all contribute to a coherent system model that serves as an appropriate input for system-centric threat analysis.

- **Systemic impact of reflection and inspection.** Continuous inspection and reflection is costly in terms of system performance and the operational disruption of extensive monitoring and inspection systems becomes prohibitive [68]. As such, we argue that an approach that involves gradual and event-based co-evolution of the reflection system model (a digital twin) in parallel with the operational distributed system will be most realistically feasible. However, propagating run-time events and changes between a run-time system and its model-based representation from a variety of sources (run-time state and interactions, versioning systems, software updates, etc.), is a non-trivial problem, for three main reasons: (i) divergence has to be avoided between the run-time state of the system and the model, (ii) the constraints of large-scale distributed systems such as partition tolerance and lack of global time synchronization have to be taken into account, (iii) when the distance between the source abstractions and targets abstractions is substantial [68] translations become complex, difficult to engineer, and verify.

V. A HOLISTIC ADAPTIVE THREAT MANAGEMENT PERSPECTIVE

Traditional threat modeling approaches are mainly used in a constructive, by-design context, and these approaches support design-centric decision-making in the sense that they offer guidance towards addressing identified threats in design or development. For example, Shostack advocates tight integration with security testing and validation approaches and emphasizes overall quality assurance [2], whereas LINDDUN [5] provides taxonomies of architectural patterns and strategies for privacy.

These are less relevant in an operational context, as it is impossible to switch between architectural tactics or design strategies at run time. As such, the key challenge is to mitigate emerging threats by instrumenting and dynamically adapting operational systems.

Security or privacy mitigations always incur a cost and involve essential trade-offs with other system concerns. For example, extensive authentication reduces the usability of a system and an approach to encrypt all communications will incur higher resource consumption (bandwidth, processing power) and increase architectural complexity (e.g., for cryptographic key management). Instead of incurring such costs by default (e.g., enforcing multi-factor authentication by default), adaptive security approaches and techniques have the intrinsic capability to change according to the system’s needs.

For example, when there is uncertainty about the credentials of a user, the system could dynamically decide to ask for extra credentials (adaptive access control). Or, when the system detects that it is near untrusted devices and thus its attack surface increases, it could dynamically reconfigure its security controls. For example, the maritime system could switch to stronger encryption mechanisms for communication when docked in a port but use weaker mechanisms at plain sea (through adaptive cryptography).

A. Concept

We envision an approach in which the threat analysis model becomes the cornerstone of a threat management framework which is capable of installing, activating, and reconfiguring a wide range of adaptive security technologies, such as those described in Section II-C, a concept that we refer to as ‘adaptive threat management’.

As the main advantage, such a holistic analysis framework will maintain an end-to-end system perspective in terms of the different threats that may emerge over time, and it can reason about and decide upon mitigations and reconfigurations in terms of their expected implications on the entire threat landscape. These mitigations may both be preventive and reactive in nature.

![Graphical depiction of the proposed adaptive threat modeling approach](image)

B. Research agenda

The proposed approach of adaptive threat management builds upon the vision of automated reflective threat modeling presented in Section IV, in which the threat model itself acts as the analysis model from which reconfiguration actions can be motivated:

- **Engineering a base platform.** Approaches to determine common framework requirements for the purpose of establishing a system base architecture that integrates a sufficient set of adaptive security and privacy controls to enable mitigation of the most plausible threats that will emerge in the context of the system are required.

- **System model.** An additional challenge is related to finding a representation of these adaptive technologies in the system model, in such a way that the model can be used both for analysis (e.g., for what-if analysis to estimate the impact of potential reconfiguration actions) and to steer and orchestrate actual reconfigurations.

- **Risk events:** Recognition of conditions in which risk significantly changes, at the basis of observed run-time events (e.g., vicinity of new entities, new data types being
processed, ...) requires new threat taxonomies (in which a threat type not only refers to a design flaw or a design weakness but also to vulnerabilities and incidents) and novel risk models and dedicated approaches that act at the level of these run-time events.

- **Technical integration:** Allowing the threat management framework to activate or tune specific security controls in response to identified threat and problems raised by analysis of individual events (reactive threat mitigation instead of proactive) requires access to these enabling technologies (e.g., orchestration or policy systems such as those described in Section II-C).

- **Novel threats:** Establishing an adaptive threat management framework that (semi-)autonomously reconfigures the system at the basis of its internal threat analysis results itself opens up a entirely novel attack surface. Especially when the system is capable of reconfiguration actions (e.g., lowering defenses), malicious entities will be strongly incentivized to gain access and manipulate these models and reconfiguration facilities.

### VI. Conclusion

In a model-based analysis activity, the quality of the inputs greatly impacts the overall effectiveness. This is certainly true for threat modeling: when performing a threat analysis activity at the basis of system model that is outdated, incomplete, ambiguous or inaccurate, relevant security and privacy threats may not be recognized (low recall), or conversely, the reported threats will not be realistic (low precision). In addition, threat modeling approaches currently focus on design-level threats and are less suited to identify threats that emerge at run time, i.e. as the operational context of a system changes.

![Graphical depiction of the proposed threat modeling innovations](image)

Fig. 4. Graphical depiction of the proposed threat modeling innovations (in black), in which (i) the system model is constructed through system reflection, and (ii) the analysis model is used to express and enact system security-oriented adaptations.

To alleviate these shortcomings, we explore and discuss the requirements and research challenges related to adopting more expressive system models, the construction of which is automated by means of system inspection and monitoring, in a novel threat modeling approach called ‘reflective threat modeling’. We complement this approach with a vision on how this can lead to integration of adaptive security and privacy technologies, which are then controlled and informed from this threat-centric perspective, an approach we refer to as ‘adaptive threat management’.

This vision aligns well with the canonical MAPE architecture for constructing self-adaptive systems [69], [70], which involves monitoring the system, performing analysis and planning, and enacting changes at the basis of dedicated and concern-specific analysis models. In that regard, this article discusses and considers the feasibility of adopting traditional threat modeling techniques for the main analysis and planning in support of holistic self-adaptive security and privacy approaches.

### Acknowledgements

This research is partially funded by the Research Fund KU Leuven, and also supported by the European H2020-SU-ICT-03-2018 Project no. 830929 CyberSec4Europe (https://cybersec4europe.eu)

### References


