A human factor approach to threat modeling

Lauren S. Ferro, Andrea Marrella, and Tiziana Catarci

Sapienza, University of Rome
{lsferro, marrella, catarci}@diag.uniroma1.it

Abstract. Cybersecurity has many challenges to address to ensure the protection of a system from an attacker. Consequently, strategies have been developed to address a system's weakness that an attacker may try to exploit. However, while these approaches may prevent an attacker getting in from the outside, they do not consider the user's actions from the inside and how their behavior may inadvertently allow an attack to take place. This paper presents a human-centered approach to threat modeling titled STRIDE-HF, which extends an existing threat modeling framework (STRIDE).

Keywords: threat modeling · human factors · cybersecurity

1 Introduction

It is human nature to make mistakes. Mistakes can occur for many reasons from feeling stressed, or from a lack of knowledge and understanding about something. One area where human error is becoming increasingly important to focus our attention towards is cybersecurity. With the increasing demand on technology and ubiquitous interaction, there has been a heavy burden to implement cybersecurity policies to protect systems from unwanted access. While there is a focus to prevent unwanted access from the outside (e.g. attackers), there has been a lack of approaches towards addressing vulnerabilities created from the inside due to human error (e.g., sharing passwords, downloading files from unknown senders, etc.). Such human errors may result in a user unknowingly allowing an attacker into a system. The impact of this could be more detrimental if a user is unaware of the consequences of their actions. Thus, making it harder to trace the origin of the breach and consequently causing a delayed response and/or solution towards addressing the breach.

Contemporary research is dedicated to understanding and categorizing human errors, and consequently human factors, across different contexts (e.g., medicine [7], aviation [38]). Unfortunately, these explorations have been mostly specific to the circumstances that they were created for. Therefore, human factor research is limited in scope, consistency, and clarity.

From an outside-in approach, scholars and security practitioners have also examined how to identify weaknesses and errors, but within a system. These studies and approaches have all worked towards anticipating an attack via a concept known as *threat modeling*. Many studies present varied approaches to threat

modeling [2,13,20,21,37,42]. Among the most popular there is the STRIDE approach (**S**poofing, **T**ampering, **R**epudiation, **I**nformation Disclosure, **D**enial of Service, and **E**levation of Privilege), which was introduced by Praerit Garg and Loren Kohnfelder at Microsoft [41] to classify vulnerabilities. However, despite the evident interest in threat modeling approaches, recent work by Xiong and Lagerström [46] found that "threat modeling is a diverse field lacking common ground, and the definitions are numerous and used in many different ways". This is also another issue related to how threat models are represented (e.g., graphical or formal).

Although particular aspects of human and system errors and weaknesses are explored in cybersecurity, a user-centered approach to threat modeling is an under-researched area. If human errors are of high concern in other areas and they are examined to address them, then cybersecurity research should also adopt the same level of scientific rigor to understand how human error can be addressed as a threat to a system like system weaknesses are addressed via threat modeling.

To tackle this challenge, this paper proposes an approach to create a user-centered threat model, which aims to complement traditional threat models to consider how human error could make it easier for an attack to occur. Therefore, the following research questions were answered:

- RQ1: Which specific topics relating to human factors in cybersecurity are discussed within the literature?
- RQ2: What threat modeling techniques exist that work towards to protecting a system from attacks?
- RQ3 : How can we use information to create a user-orientated threat model?

Based on this information, we theoretically developed a user-centered framework based on STRIDE, called STRIDE-HF. Considering that STRIDE has never been studied before in conjunction with human factors, this paper ventures into a new area of inquiry. Thus, a theoretical framework appears to be the most appropriate solution to address this research. The outcomes of this paper present a foundation to extend and iterate upon, which is user focused. Therefore, we provide the following contribution:

- A novel (inside-out) approach towards user-centered threat modeling.
- Insight towards how threat modeling methods and human factors could be considered for developing more secured systems.
- Future research directions for user-centered threat modeling and for future iterations of STRIDE-HF.

The rest of the paper is organized as follows. In Section 2 we present key concepts, relevant definitions, theories, and models and outline the key details of the STRIDE framework. In Section 3 we present our research model and the steps followed towards creating STRIDE-HF and its current implementations. In Section 4 we discuss our observations and suggestions for future research. Finally, in Section 5, we present our conclusion.

2 Background and Related Work

Cybersecurity is a highly relevant area in today's society. In recent times, with the COVID-19 pandemic, our lives have unexpectedly and forcibly become online; resulting in an increase in online data sharing, privacy concerns, and changes to access protocols. With many users having to transition from traditional methods of working and interaction, even with the most simple of tasks (e.g., writing a document in a word processing software), they have inevitably been forced to learn and engage with several new online systems to work remotely. Consequently, there has also been an increase in cyberattacks [3].

Humans possess many flaws that make them vulnerable. Users argue for the privacy of their data while within the same breath they will be posting what they had for lunch, their relationship status, or what they really think about the governments latest decision. All this information may seem trivial at first but it can provide an attacker with enough data to begin developing a plan of attack. This kind of behaviour such as sharing information online or making trade-offs could relate to *Lack of Knowledge* regarding the sensitivity of certain types of information or how that information could be used in an attack. Our (heightened) sense of self also lets us down by allowing a user to be more vulnerable to the influence of attacks because they do not believe they could be the target of an attack or have worthwhile information. In other cases, a users desire to reciprocate the seemingly "alturistic" actions of others or to help those who are seeking assistance allow them to fall victim to a cyber attack. However, the key concept here is that human behaviour can put a user and their community in danger with seemingly little effort. Thus, we need to find ways to protect users from being exploited and effectively from their own bad cybersecurity behaviour.

2.1 Human Factors

Human errors can be the result of negligence, accident, or deliberate action [17]. Human factors has been the topic of study in many areas, namely within the context of aviation, which focuses behaviors leading up to human error. For example, the Dirty Dozen proposed by Dupont [15] describes twelve of the most common human factor related errors, which may lead to aviation related accidents or incidents.

- **Lack of Communication**: people not communicating with each other within a working and/or online environment.
- **Complacency**: a feeling of self-satisfaction that can lead to a lack of awareness of potential dangers.
- **Lack of Knowledge**: not having enough experience and specific knowledge that can lead to poor decisions.
- **Distraction**: when a user's attention has been taken away from the task that they are required to do.

- **Lack of Teamwork**: not providing enough support towards a group of people, co-workers, etc, who rely on your support.
- **Fatigue**: is a physiological reaction resulting from prolonged periods of work and stress.
- **Lack of Resources**: not having enough resources (e.g. time, tools, people, etc.) to complete a task.
- **Pressure**: pressure to meet a deadline interferes with our ability to complete tasks correctly, then it has become too much.
- **Lack of Assertiveness**: not being able or allowed to express concerns or ideas.
- **Stress**: acute and chronic stress from working for long periods of time or other demanding issues such as family or financial problems.
- **Lack of Awareness**: working in isolation and only considering one's own responsibilities, often leading to a disconnect from what others are doing.
- **Norms**: workplace practices that develop over time, which can then influence others behaviors.

While human factors is growing in many areas, one area that can greatly benefit from it is cybersecurity. This is because by understanding human factors we can begin to gain an improved understanding towards addressing human error and improving the security of systems and data.

2.2 Human Factors and Cybersecurity

Human factors in cybersecurity is becoming widely discussed (e.g., [48] [1] [4] [32] [25] [45]), which has led to several issues. The first is that there are many variations for often the same terms due to a lack of consistency or conventions to describe human factors. Furthermore, of the research that does exist, it often has a limited scope [48], ambiguous, or only acknowledges the concept of human factors rather than focuses on it [44].

If we could consider the broad definitions within other areas, we can begin to find commonalities such as the use of the same concepts or similar terms and work towards a more concise list. For example, if we consider *Norms* from Dupont's Dirty Dozen [15], there are similarities with other descriptions. For example, Da Veiga [10] describes pressures from norms that adopt common philosophy for completing tasks in certain ways because that is "the way things are done here" [26] or influential factors such as the personality of the organization [39]. Lastly, Henshel et al. [19] incorporates a user's culture as part of the human factors component within their holistic cybersecurity risk assessment framework. Considering these papers, they all relate to the broader concept of *Norms*. Similar examples also exist for a *Lack of Knowledge* and *Awareness* [23, 49]. Therefore, it is likely that we could begin with one general human factor and continue to develop sub-factors that could relate to more specific circumstances.

Other current trends have also emerged that consider the human factors of users through two lenses: personal/user-centered and organizational/cultural

such as those by Kraemer [25], Al-Darwish et al. [1], Badie and Lashkari [4], and Mortazavi-Alavi [32]. To this end, and like previous studies, human factors could be impacted by several aspects at once depending on a user's previous experiences and how a workplace impacts the user (e.g., both in a social and policy perspective). Therefore, we could consider the user as a node of a larger network that includes part of an (online) team, culture, and ultimately the system that they are interacting with [35]. There is an increase need by organizations to invest time to develop an information security culture [17] that needs to include all the personnel and leadership [18]. By building this culture, organizations can minimize the risk to the exposure of sensitive information [11]. Current research highlights that a positive information security culture can increase security policy compliance, strengthen the overall information security posture, and reduce the financial loss due to security breaches. For example, Chen and Zahedi [9] demonstrated that once users have perceive or experienced a cyber threat, they are more likely to take protective actions. From this study, we could consider this relating to a user's lack of knowledge or competency resulting from a lack of experience in such topics. For instance, Mashiane and Kritzinger [29] identified a large amount of constructs being proposed as the determinants of cybersecurity behavior. It makes it difficult to decide which constructs to focus on when designing cybersecurity behavior interventions. Moreover, it is also important to consider that an employee's attitude and involvement within a company can be influenced by their own experiences. Therefore, it may be key to ensure that employees have had training that provides them with an opportunity to experience first-hand or in real-time threats that they may encounter to allow them to have this experience to internalize. This is also a consideration of Kraemer et al. [25] who identified nine thematic areas where key human and organizational factors were grouped into. Again, highlighting the need for *Training*, thus declaring a fault in a users knowledge for cybersecurity issues. However, Kraemer's study appears to focus more on organizational related issues rather than the user. The study neglects to understand the overall connection between a user and their interaction and behavior within an environment.

2.3 Threat Modeling

To carry out the threat modeling process, we must first understand what are the threats and attacks that we are trying to project ourselves again. Often, a cyber attack can be a highly effective n sophisticated attack, which can bypass even well thought out technological security structures. For a cyber attack to be successful, it typically follows a seven step approach known as the cyber kill chain. In general, a cyber kill chain is a procedural path that an intruder takes to penetrate information systems over time to execute an attack on the target [47].

1. **Reconnaissance**: usually happens in anticipation before an actual attack. This is the initial phase where attackers select their targets, monitor a network system to try and develop a more informed understanding of the target.

2. **Weaponization**: uses the information from the reconnaissance stage to carefully develop an attack, which may include sending malware, launching a DoS attack, or hacking a system.
3. **Delivery**: is the transmission phase where the weaponization (e.g. malware, attack, etc.) is undertaken. The delivery of a payload or an attack can occur in many different ways (e.g. phishing email) depending on the objective of the attack.
4. **Exploitation**: is the first phase in the execution of a cyberattack where an attacker takes control of the targets environment by exploiting their weaknesses or taking advantage of their access to the system.
5. **Installation**: is where attackers may want to install malware (if they have entered a system) or deploy a payload if it has not already been done by the user (e.g. downloading and installing software from a phishing email).
6. **Command and control**: is where the attackers take (remote) control of a system or device.
7. **Action on objective**: is where the attackers carry out their goals and objectives that have driven the attack in the first place.

The cyber kill chain highlights the steps involved if an attacker can gain enough useful information during the reconnaissance phase. For example, information about a person and/or the company that they work for can help the attacker to develop an angle to contact that user with to gain more information. From here, this information can be used to persuade and deceive victims because it helps to improve legitimacy of the attackers intentions. Consequently, victims are less likely to question the interaction. In some cases, other factors such as timeliness can be used to persuade and deceive victims because it helps improve legitimacy. For example, if there has been a large data breach, an attacker may utilize the fear and contact potential victims posing as a technician to improve security against the threat.

With an understanding of the process that an attacker can follow to access a system, we can begin to analyze how to protect it. One way that this is done is via threat modeling. Threat modeling as defined by Uzunov and Fernandez [43] as "a process that can be used to analyze potential attacks or threats, and can also be supported by threat libraries or attack taxonomies". However, while several other definitions exist that also define threat modeling [5, 6, 12, 14, 28, 31, 40] and as systematically assessed by Xiong and Lagerström [46]. Generally speaking, threat modeling allows security designers to accurately estimate and anticipate an attack and to prevent any unauthorized attacks that gains access to sensitive information, networks, and applications (e.g. Malware, Phishing, Denial of Service (DoS/DDoS)).

Beyond the definition, there are also many different types of threat modeling approaches, frameworks, techniques, models, and theories that all work towards identifying threats and approaches to address them [21, 46]. Each of these have their own context in mind such as preventing attackers from breaching a system, finding weak points within a systems architecture, develop strategies to mitigate potential attacks, among others.

STRIDE The STRIDE method is a mnemonic for six different types of security threats [41]. It supplies the foundation of our theoretical model known as STRIDE-HF (Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of privilege – Human Factor).

- **(S) Spoofing**: using someone else's credentials to gain access to otherwise inaccessible assets.
- **(T) Tampering**: changing data to mount an attack.
- **(R) Repudiation**: occurs when a user denies performing an action, but the target of the action has no way to prove otherwise.
- **(I) Information disclosure**: the disclosure of information to a user who does not have permission to see it.
- **(D) Denial of service**: reducing the ability of valid users to access resources.
- **(E) Elevation of privilege**: occurs when an unprivileged user gains privileged status.

STRIDE has also been used to address many concerns within cybersecurity (e.g., [8, 24, 28, 36]) as well as variations such as STRIDE-per-element and STRIDE-per-interaction [41]. Moreover, Khan et al. [24] differentiate the two variations by describing STRIDE-per-element as a more complex method because it analyzes the behavior and operations of each system component; and STRIDE-per-interaction as a more simpler method to perform because it provides protection strategies sufficient enough to protect a system. However, the general version of STRIDE includes elements that are typical in many cybersecurity related situations.

When a system is developed, it is often driven by requirements that define interaction (i.e., what the user can and cannot do), and how the system is intended to work. Security requirements are driven by what should not occur (i.e., a user gaining access to areas/data that they should not be able to). However, it is extremely difficult to consider every kind of threat and/or behavior that a user can do with a system that can create security issues at a later stage. Yet, of those threats that have been defined after a thorough analysis and risk management, the security analyst must find ways to mitigate them. Risk management consists of risk assessment, risk reduction, and risk acceptance and from here the threats that are identified must be prioritized, often by damage and likelihood. For example, one way to approach managing a risk is presented by Myagmar et al. [33]:

- **Accept the risk**: the risk is so low and so costly to mitigate that it is worth accepting.
- **Transfer the risk**: transfer the risk to somebody else via insurance, warnings etc.
- **Remove the risk**: remove the system component or feature associated with the risk if the feature is not worth the risk.
- **Mitigate the risk**: reduce the risk with countermeasures.

By comparing risk assessment and the assessment of human errors, we can see that they too share similarities. For example, when we look at risks, we consider a potential incident, how it may occur and how we can either mitigate it or reduce the impact of it should a risk occur. For example, Mancuso et al. [27] propose a conceptual framework that aims to maintain interactions between the components of a cyber attack, which is described in terms of three dimensions: adversarial, methodological, and operational. Yet this approach does not consider the behaviour of users like a threat model views the behavior of an attacker. Like human factors, considering the risks of a user sharing a password or downloading a potentially dangerous attachment could also be addressed as part of a user-centered approach with strategies in place.

If we consider threat modeling and risk management, these approaches are focused on preventing an attacker gaining entry into a system by assessing a system to identify areas of weaknesses that an attacker could exploit. Yet, even with the most well devise plan and risk management, all of these could be for nothing is a user unknowingly opens the proverbial door to an attack. Therefore, this paper aims to consider and theoretically present all the aforementioned concepts and approaches but from a reversed engineered approach - that is to also view the user and their behaviors as risks and threats and to develop strategies in conjunction with traditional approaches.

3 STRIDE-HF

This section describes the theoretical and conceptual process [30] that we took to develop **STRIDE-HF**. This research adopted an inductive approach, that is starting with an observation of contemporary literature surrounding key areas, identified that there is a gap concerning user-centered threat models and looked at how to address this gap by proposing a conceptual/theoretical framework.

The development of this work began by exploring the current literature surrounding human factors, threat modeling techniques, and discussions relating to these within cybersecurity. This was to understand how, if at all, current literature documents risks/security vulnerabilities from a user's perspective and not specifically from the attacker's perspective. Furthermore, this step explored how these vulnerabilities could be classified/related to human factors. We also chose for the time being to exclude aspects of decision making and attitude as they contain several aspects that also need further investigation and studies to determine their impact on human factors and cybersecurity.

After considering the discussion surrounding various human factor models in cybersecurity, we felt that the Dirty Dozen [15] provided an encompassing foundation to start with and to use and iterate upon in the future.

Next, in a similar manner, we looked for a threat model that could provide a general foundation to expand upon, therefore we chose STRIDE (Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of privilege) [41]. The primary rationale behind the use of STRIDE was because it

provided a neutral foundation that could be expanded upon (to consider other elements beyond STRIDE) in future work and empirical validation.

The second step was to consider the STRIDE model in the context of human factors and how it may relate to the STRIDE elements. After examining and discussing the relevant literature, we theorized that human factors could be threat modeled in a similar way to how attackers behavior is. In this way, be began to consider types of behaviors that were discussed within the literature and how they may align with STRIDE elements, as presented in Table 1 taken from Ferro and Sapio [16].

To use the STRIDE-HF model, security analysts and researchers will need to understand how users interact with each other and with the systems inside the workplace. This includes the type of environment (e.g. open-plan, cubicle based, working from home, etc), and culture (e.g. carefree and relaxed or strict and procedural). This may be achieved by qualitative (e.g. observations, questionnaires and interviews with users) or quantitative (e.g. surveys) methods. From here, security analysts can begin to look at their threat models from an attackers perspective and then consider how human factors could impact what has been modeled or managed.

3.1 Implementing STRIDE-HF into an interactive experience: Another Week at the Office

Since STRIDE-HF is still a developing model, it has been implemented within a serious game titled Another Week at the Office (AWATO) [16]. This game has provided a practical way to incorporate the STRIDE-HF framework to educate and assess users behavior in a virtual office space. In AWATO, players take on the role of security analyst who must observe the characters within a typical workplace environment and identify erroneous behavior, such as leaving a computer unattended and unlocked or a post-it note on the ground with login information like in Figures 1 and 2.

From here, the user must decide whether or not this kind of behavior is bad and subsequently report it. Once an incident has been reported, the player must then classify the threat in accordance to STRIDE-HF as presented in Figure 3.

Table 1. STRIDE - HF [16]

Threat	**Human Factor(s)**	**Behavior (*examples*)**	**Response (*examples*)**
Spoofing	Lack of Awareness, Lack of Knowledge, Lack of Resources	Downloading files online or via email attachments.	Educate users about what to look for when accessing links within emails.
Tampering	Distraction, Lack of Awareness, Stress, Pressure, Fatigue	Modifying files to backdate them to avoid punishment. Unblocking blocked ports to get access.	Implement a platform where documents must be uploaded (logs date, time, user, etc.).
Repudiation	Accidentally deleting files.	Not submitting files on time/to the right location.	Change how files are managed and are monitored.
Information Disclosure	Complacency, Distraction, Norms, Stress, Pressure, Lack of Assertiveness	Sharing passwords among colleagues for a time trade-off.	Letting a friend borrow an access card * Enforce stronger punishments for password/access card sharing.
Denial of Service	Distraction, Lack of Awareness, Stress, Pressure	Unplugging hardware for other purposes (e.g. additional charging space).	Clearly label exposed cables to indicate their use.
Elevation of Privilege	Lack of Assertiveness	Giving access to a file because someone with authority asked for it.	Create a more accessible way to report bad behavior of superiors.
	* More likely to be responsible for the STRIDE element over other Human Factors.		

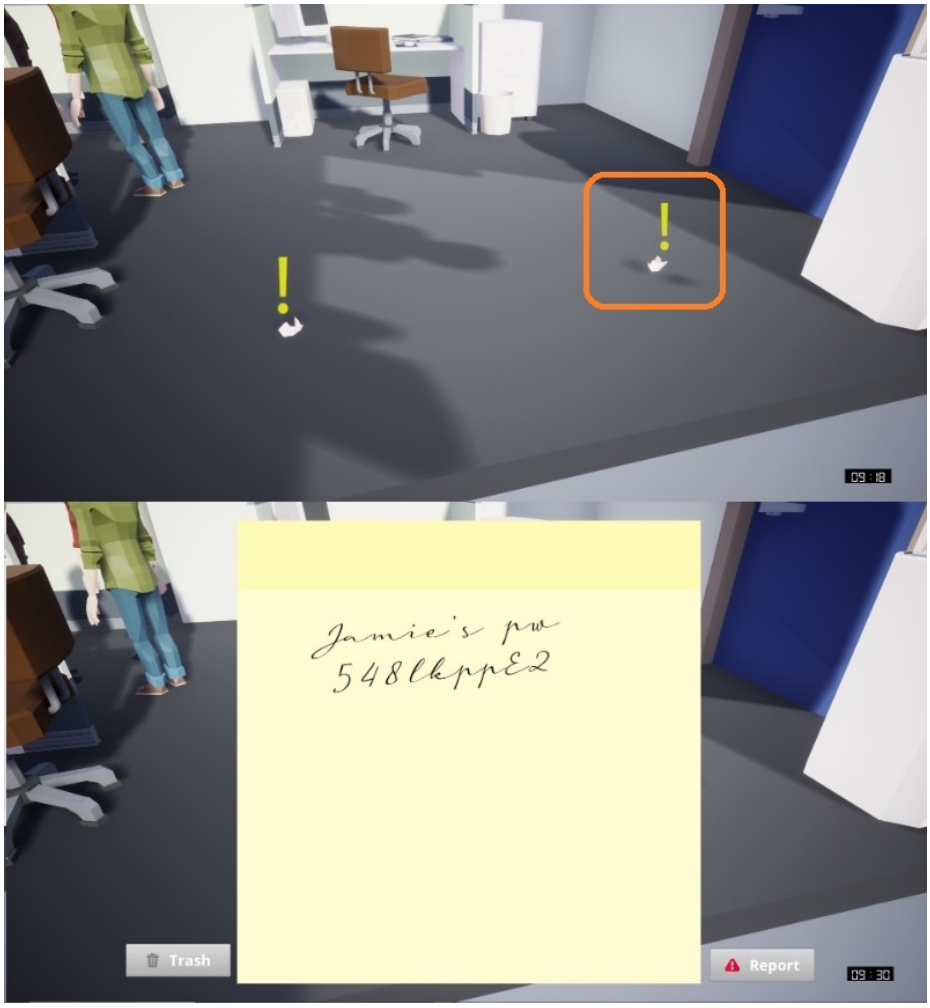

Fig. 1. Example of a post-it note with sensitive information

We analyzed the use of STRIDE in terms of players observing generally bad cybersecurity practices and aligning them to what the most relevant STRIDE-HF element. At this stage, the point was not to validate STRIDE-HF but to understand how a framework such as STRIDE-HF could be used in a practical way. While the game did feature a short text-based primer to make the players familiar with the concepts of threat modeling, human factors, and STRIDE-HF, this study did offer insight on how we could address knowledge gaps observed through game-play in terms of more tailored training. A consideration that will be further elaborated as we begin to empirically validate STRIDE-HF.

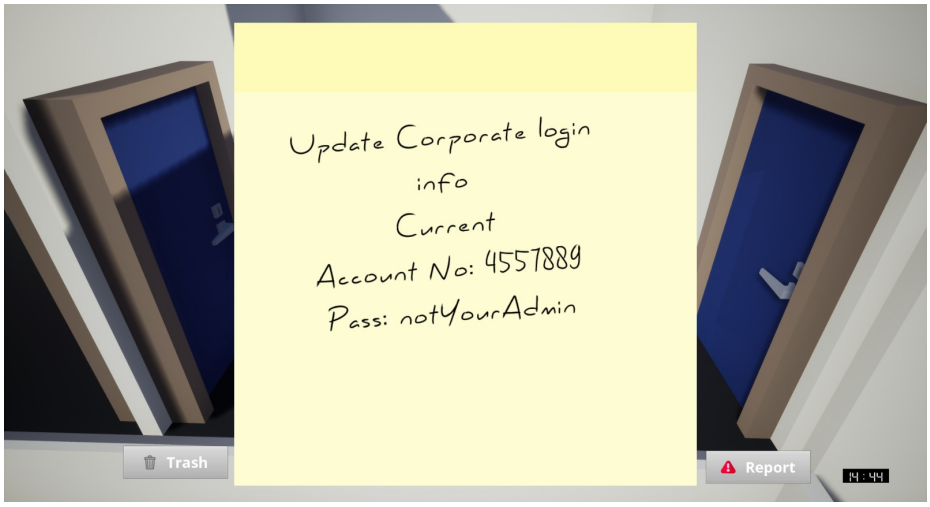

Fig. 2. Example of a post-it note with sensitive information with login and account information

Fig. 3. Example of STRIDE-HF classification inside of AWATO

4 Discussion and Implications

For developing a user-centered threat model, a central issue has been to understand the discussions, definitions, and approaches for addressing human errors and threats to a system. On the one hand, human factor research is gaining momentum in cybersecurity, which has highlighted the role the human factors play in cybersecurity. On the other hand, research and definitions are still very broad and context specific. The purposes of this paper are (a) to explore contemporary research within human factors, including research within the context of cybersecurity, (b) explore contemporary research that describes approaches towards addressing threats to a system, and (c) how can this information be used to develop a more user-centered approach to threats where threats are defined as human errors. STRIDE-HF presents an approach that considers both human factors and threat modeling together to help understand what types of human errors could result in STRIDE elements.

4.1 Human factors and threat modeling

Humans are prone to making mistakes, especially if our environment facilitates it. One area that has been the topic in the literature is workplace culture. Therefore, it could be an area to start with. There may be opportunities to assess the workplace environment by measuring employees attitudes towards their employer, and workplace culture, to determine if more can be done to address negative issues. For example, if many employees are overworked, it is more likely that they will make trade-offs for time or disregard basic security protocols such as leaving a computer unlocked or sharing passwords. Therefore, by changing work practices or developing more strategic approaches to managing workloads, employees may feel more positive and be more prone to make less errors. In addition, there may be ways to address and change workplace culture by encouraging employees to participate in activities that are orientated towards their commitment to the organizations security goals or to engage with like-minded colleagues [22]. In this way, by addressing environmental and cultural factors it could reduce the likelihood of human error resulting in a breach.

As demonstrated by Chen et al. [9], exposing users to threats in a controlled environment may also offer a way for them to understand the process leading up to a threat, the threat itself, and subsequent consequences of it being successful. However, it is also equally important to consider that more training is not always the solution. In these instances, training provides users with a personalized experience, that is meaningful because of the interaction that it affords. Ultimately, such experiences can also align with the level of access and/or responsibilities that they have within a workplace. This is also a contention that Öğütçü et al. [34] confirmed where users demonstrated more security-focused behavior when they could perceive threats and increase their awareness of the technology. Therefore, it may also come back to addressing a users level of competency by actively exposing them to scenarios that allow them to experience these issues and develop personal and meaningful connections to security issues.

It is clear that there is still a long way to go towards developing a consistent definition of human factors (and what they are) both at a general level and more specifically within a cybersecurity context. Similarly, there is also a lack of consistency when defining threat modeling.

Typically human factors and threat modeling has been considered two separate areas of study. However, threat modeling and risk assessment present similar approaches to human factors in ways to identify, understand, or to anticipate weaknesses or likely errors. Thus, strategies and processes can be implemented to mitigate the effects of these weaknesses and errors. However, with this being said it is reassuring that while the terminology varies, there is some consistency with the errors or human factors that they are addressing. Therefore, it is likely as both fields mature so too will more concise definitions.

Another important consideration is understand how to address human factors throughout a cyber kill chain or where certain factors are likely to be more damaging. For example, in a company that heavily uses social media, a user who is sharing work-related information online or even a photo on social networks, it may be enough to draw the attention of an attacker. However, it is not until the delivery or exploitation stage that human factors are more detrimental. Therefore, in such cases, more focus should be directed towards educating users about email security and social engineering. This may include training to identify persuasive techniques or how to validate the identity of callers or what to look for in emails from unknown senders.

4.2 STRIDE-HF as a user orientated threat modeling approach

STRIDE-HF presents a novel direction to consider threat modeling from a human factor perspective. As discussed within this paper, research within human factors is gaining momentum within the field of cybersecurity. Yet, we are still far from a consistent discourse and terminology. To this end, this paper highlights the importance for working towards a more consistent definition standard and why we should begin to consider human factors as a type of threat towards a systems security in a socio-technological world. In this way, we can begin to work towards the development of user-centered threat models. We aim for STRIDE-HF to become the beginning of a new paradigm that explores human error as a way to further protect the security of systems and the data within them.

The STRIDE-HF Framework functions by offering a way for security analysts to consider human factor related behavior while assessing the types of breaches that could result from them. For example, if a user shares a password (because that is part of the workplace norms) it could result in an elevation of privilege where a user may disable certain settings unknowingly, thus creating a vulnerability. The procedure to use the framework requires that human factor elements are identified within a work environment, which can be done in several ways (e.g. observation or assessment). From here, the framework can help the user to identify the type of STRIDE element that aligns with the human factor that may influence it. Alternatively, the user can use STRIDE-HF in reverse where they identify the likely (STRIDE) issues and then the subsequent human

factor(s). The fundamental difference that STRIDE-HF offers in comparison to traditional threat modeling methods is that it takes a "reversed engineered" approach towards classifying threats that may affect the security of a system from the perspective of users rather than an attacker.

4.3 Future Work

At present, STRIDE-HF is being iterated within the game Another Week at the Office [16]. However, we are striving towards developing this further to include training modules that can be used by security analysts to help them analyze current work practices and identify what human factors could weaken the security practices that are currently in place. Moreover, we would also like to include relevant material (e.g. approaches, activities, information) that could help even a novice security analyst to understand the human factors within their workplace environment and how to address them. This material would also align with what users do within AWATO so that it can become a wholesome training and learning experience.

Future iterations of STRIDE-HF may include additional human factors that are relevant and extend beyond those defined within this paper as well as the incorporation/use of more psychological based principles such as decision making, culture, and attitude. Since it is a requirement in any organization to make decisions on a daily basis. The types of decisions vary from habitual ones such as when to take a coffee break to downloading a document, all of which can lead to negative and positive outcomes. However, these decisions vary greatly depending on the needs of the user and the environmental/psychological factors that may be influencing them. Therefore, there are many approaches that try to predict the way to predict a user or model their behavior. If we are to consider these aspects within the context of cybersecurity, the work environment, and the everyday user, there are several ways that a user could unknowingly/accidentally provide sensitive information to an attacker or leave sensitive information easily available for others to use and thus compromising cybersecurity. Consequently, future development may also explore such decision making behaviors in a quantifiable way (i.e. through questionnaires) to develop a rubric for assessing the risk of a human factor(s) and/or STRIDE element.

5 Concluding Remarks

This paper presented an iterated version of the STRIDE threat modeling technique with STRIDE-HF, which is a user-centered threat model that is aligned with Dupont's Dirty Dozen [15].

As summarized in Section 2, the paper has drawn an informed insight to identify the gap that exists when discussing user-centered threat modeling as an additional technique to use in conjunction with traditional threat modeling and as part of the security design of a system. The paper, as discussed in Section 3 presented the development of a novel approach to threat modeling. Lastly, in

Section 4, we discussed the implications of what this research has identified and how it could be continued and applied.

The field of human factors in cybersecurity is still maturing, and more work is needed to quantify the impact of implementing such strategies. This paper contributes by providing initial insight into this developing field and a way to consider an approach to user-centered threat modeling. However, the authors want to stress that while STRIDE-HF does offer a starting point, much like traditional threat modeling, we do not suggest that there is a "one size fits all" approach since the security of systems requires varied approaches. Therefore, further research needs to define human factors in a more concise way so that we can begin to identify those which are more prevalent in specific security situations (e.g. local intranet versus protecting a server). To this end, we have provided a starting point to begin further research, which pushes the considerations of a more inside-approach to threat modeling.

Lastly, this paper also raised several interesting questions for future work. For instance, could more psychological elements be present that impacts human factors and could these be quantitatively assessed to provide more insight towards high-risk human factors and the errors that they could lead to. Such methods could also improve our understanding and highlight the level of impact that psychological aspects could have on human factors to better understand how it could affect the security measures that are currently in place or how new ones should be designed and implemented. Therefore, empirical validation is the logical next step towards not only validating our theoretical model but also these additional considerations. The outcomes of such studies would greatly lead to not only empirical improvements towards the understanding of user specific aspects of threat modeling but also to further define what are human factors within the broader discourse of cybersecurity.

References

1. Al-Darwish, A.I., Choe, P.: A framework of information security integrated with human factors. In: International Conference on Human-Computer Interaction. pp. 217–229. Springer (2019)
2. Alberts, C.J., Behrens, S.G., Pethia, R.D., Wilson, W.R.: Operationally critical threat, asset, and vulnerability evaluation (octave) framework, version 1.0. Tech. rep., CARNEGIE-MELLON UNIV PITTSBURGH PA SOFTWARE ENGINEERING INST (1999)
3. Andrade, R.O., Ortiz-Garcés, I., Cazares, M.: Cybersecurity attacks on smart home during covid-19 pandemic. In: 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4). pp. 398–404. IEEE (2020)
4. Badie, N., Lashkari, A.H.: A new evaluation criteria for effective security awareness in computer risk management based on ahp. Journal of Basic and Applied Scientific Research **2**(9), 9331–9347 (2012)
5. Baquero, A O, K.A.J., Janusz, Z.: Threat modeling for aviation computer security. crosstalk **21** (2015)

6. Bedi, P., Gandotra, V., Singhal, A., Narang, H., Sharma, S.: Threat-oriented security framework in risk management using multiagent system. Software: Practice and Experience **43**(9), 1013–1038 (2013)
7. Bleetman, A., Sanusi, S., Dale, T., Brace, S.: Human factors and error prevention in emergency medicine. Emergency Medicine Journal **29**(5), 389–393 (2012)
8. Chen, X., Liu, Y., Yi, J.: A security evaluation framework based on stride model for software in networks. Int. J. Adv. Comput. Technol (2012)
9. Chen, Y., Zahedi, F.M.: Individuals'internet security perceptions and behaviors: Polycontextual contrasts between the united states and china. Mis Quarterly **40**(1) (2016)
10. Da Veiga, A.: A cybersecurity culture research philosophy and approach to develop a valid and reliable measuring instrument. In: 2016 SAI Computing Conference (SAI). pp. 1006–1015. IEEE (2016)
11. Da Veiga, A., Martins, N.: Information security culture and information protection culture: A validated assessment instrument. Computer Law & Security Review **31**(2), 243–256 (2015)
12. Dahbul, R., Lim, C., Purnama, J.: Enhancing honeypot deception capability through network service fingerprinting. In: Journal of Physics: Conference Series. vol. 801, p. 012057. IOP Publishing (2017)
13. Desolda, G., Di Nocera, F., Ferro, L., Lanzilotti, R., Maggi, P., Marrella, A.: Alerting users about phishing attacks. In: Moallem, A. (ed.) HCI for Cybersecurity, Privacy and Trust. pp. 134–148. Springer International Publishing, Cham (2019)
14. Dhillon, D.: Developer-driven threat modeling: Lessons learned in the trenches. IEEE Security & Privacy **9**(4), 41–47 (2011)
15. Dupont, G.: The dirty dozen errors in maintenance. In: The 11th symposium on human factors in maintenance and inspection: Human error in aviation maintenance (1997)
16. Ferro, L.S., Sapio, F.: Another week at the office (awato)–an interactive serious game for threat modeling human factors. In: International Conference on Human-Computer Interaction. pp. 123–142. Springer (2020)
17. Glaspie, H.W., Karwowski, W.: Human factors in information security culture: A literature review. In: International Conference on Applied Human Factors and Ergonomics. pp. 269–280. Springer (2017)
18. Guo, K.H.: Security-related behavior in using information systems in the workplace: A review and synthesis. Computers & Security **32**, 242–251 (2013)
19. Henshel, D., Sample, C., Cains, M., Hoffman, B.: Integrating cultural factors into human factors framework and ontology for cyber attackers. In: Advances in Human Factors in Cybersecurity, pp. 123–137. Springer (2016)
20. Howard, M., LeBlanc, D.: Writing secure code. Pearson Education (2003)
21. Hussain, S., Kamal, A., Ahmad, S., Rasool, G., Iqbal, S.: Threat modelling methodologies: a survey. Sci. Int.(Lahore) **26**(4), 1607–1609 (2014)
22. Ifinedo, P.: Information systems security policy compliance: An empirical study of the effects of socialisation, influence, and cognition. Information & Management **51**(1), 69–79 (2014)
23. Kemper, G.: Improving employees' cyber security awareness. Computer Fraud & Security **2019**(8), 11–14 (2019)
24. Khan, R., McLaughlin, K., Laverty, D., Sezer, S.: Stride-based threat modeling for cyber-physical systems. In: 2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe). pp. 1–6. IEEE (2017)

25. Kraemer, S., Carayon, P., Clem, J.: Human and organizational factors in computer and information security: Pathways to vulnerabilities. Computers & security **28**(7), 509–520 (2009)
26. Lundy, O., Cowling, A.: Strategic human resource management. Cengage Learning Emea (1996)
27. Mancuso, V.F., Strang, A.J., Funke, G.J., Finomore, V.S.: Human factors of cyber attacks: a framework for human-centered research. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. vol. 58, pp. 437–441. SAGE Publications Sage CA: Los Angeles, CA (2014)
28. Marback, A., Do, H., He, K., Kondamarri, S., Xu, D.: A threat model-based approach to security testing. Software: Practice and Experience **43**(2), 241–258 (2013)
29. Mashiane, T., Kritzinger, E.: Theoretical domain framework to identify cybersecurity behaviour constructs. In: International Conference on Innovative Technologies and Learning. pp. 320–329. Springer (2019)
30. McGregor, S.L.: Understanding and evaluating research: A critical guide. Sage Publications (2017)
31. Mitnick, K.D., Simon, W.L.: The art of deception: Controlling the human element of security. John Wiley & Sons (2003)
32. Mortazavi-Alavi, R.: A risk-driven investment model for analysing human factors in information security. Ph.D. thesis, University of East London (2016)
33. Myagmar, S., Lee, A.J., Yurcik, W.: Threat modeling as a basis for security requirements. In: Symposium on requirements engineering for information security (SREIS). vol. 2005, pp. 1–8. Citeseer (2005)
34. Öğütçü, G., Testik, Ö.M., Chouseinoglou, O.: Analysis of personal information security behavior and awareness. Computers & Security **56**, 83–93 (2016)
35. Parsons, K., McCormac, A., Butavicius, M., Ferguson, L.: Human factors and information security: individual, culture and security environment. Tech. rep., DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION EDINBURGH (AUSTRALIA) COMMAND ... (2010)
36. Ruffy, F., Hommel, W., von Eye, F.: A stride-based security architecture for software-defined networking. ICN 2016 p. 107 (2016)
37. Saitta, P., Larcom, B., Eddington, M.: Trike v1 methodology document. Draft, work in progress (2005)
38. Salas, E., Maurino, D., Curtis, M.: Human factors in aviation: an overview. Human factors in aviation pp. 3–19 (2010)
39. Saunders, M., Lewis, P., Thornhill, A.: Research methods for business students. Pearson education (2009)
40. Scandariato, R., Wuyts, K., Joosen, W.: A descriptive study of microsoft's threat modeling technique. Requirements Engineering **20**(2), 163–180 (2015)
41. Shostack, A.: Threat modeling: Designing for security. John Wiley & Sons (2014)
42. UcedaVelez, T., Morana, M.M.: Risk centric threat modeling. Wiley Online Library (2015)
43. Uzunov, A.V., Fernandez, E.B.: An extensible pattern-based library and taxonomy of security threats for distributed systems. Computer Standards & Interfaces **36**(4), 734–747 (2014)
44. Vieane, A., Funke, G., Gutzwiller, R., Mancuso, V., Sawyer, B., Wickens, C.: Addressing human factors gaps in cyber defense. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. vol. 60, pp. 770–773. SAGE Publications Sage CA: Los Angeles, CA (2016)

45. Widdowson, A.J., Goodliff, P.B.: Cheat, an approach to incorporating human factors in cyber security assessments. In: 10th IET System Safety and Cyber-Security Conference 2015. pp. 1–5 (2015)
46. Xiong, W., Lagerström, R.: Threat modeling–a systematic literature review. Computers & security **84**, 53–69 (2019)
47. Yadav, T., Rao, A.M.: Technical aspects of cyber kill chain. In: International Symposium on Security in Computing and Communication. pp. 438–452. Springer (2015)
48. Young, H., van Vliet, T., van de Ven, J., Jol, S., Broekman, C.: Understanding human factors in cyber security as a dynamic system. In: International Conference on Applied Human Factors and Ergonomics. pp. 244–254. Springer (2017)
49. Zwilling, M., Klien, G., Lesjak, D., Wiechetek, Ł., Cetin, F., Basim, H.N.: Cyber security awareness, knowledge and behavior: a comparative study. Journal of Computer Information Systems pp. 1–16 (2020)