Securtle: The Security Turtle

People in IoT-enabled homes can not reasonably be expected to be aware of the state of the network; specifically when their own local area networks are targeted. While risk communication can be seen as a solution, often solutions require the installation of scanners, reading textual outputs, and these often assume high levels of security expertise. These tools are static defenses: technically inadequate and failing to align with homeowners' security and computer expertise. In this paper we describe the underlying system which detects scans, password guessing attacks on WiFi, attacks on Zigbee, and other indicators to create a single holistic risk indicator, as well as the implementation of a turtle shaped device which uses a Raspberry Pi as a demonstration of the ambient risk communication prototype. Named the Securtle, it uses simple colors, changing hues, intensity, and blinking LED's to communicate the changing risk levels of a home network in operation. Normally, as attacks flow through the network a home owner, they have no way of reasonably knowing that their own network is under attack. The Securtle detects, in real time, risks that a home owner should immediately address as well as identifies vulnerabilities, ensures that they are patched, and disables unnecessary connections. The LED embedded shell shifts color, providing users of all ages and knowledge levels a simple real time risk exposure indicator of their digital environment. Consequently the homeowner can take immediate self-mitigation actions: power off, disconnect, restart, or perform system updates. These functions are already built in devices on the network, but their use is often neglected.

CCS Concepts: • **Human-centered computing** → *Interaction paradigms*; *Heuristic evaluations*; • **Security and privacy** → *Usability in security and privacy.*

Additional Key Words and Phrases: HCI, AI, IoT, end-user protection

ACM Reference Format:
. 2021. Securtle: The Security Turtle. In *IUI '21: ACM Intelligent User Interface, April 13–17, 2021, College Station, TX*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

People using technology at home can not reasonably be expected to be aware of the security state of the network; specifically when their own local area networks are targeted. Conversely, home technology may include access to our most sensitive data beyond work authentication and financial information. While system monitoring and risk communication can be seen as a solution this requires installation of scanners, which often assumes high levels of security expertise and are inaccessible for general users. Even experts find these unusable [9]. These tools are often static defenses. Such static defenses face critiques that they are technically inadequate, and they must be run continuously or at the correct time. [10] The interactions are designed for experts and fail to align with homeowners' security and computer expertise [1]. Finally, the majority of those technologies are focused on Internet Protocol based end to end communications. Conversely, with technological advancement, IoT (Internet of Things) devices have increased in both ubiquity and market penetration, making user reliance on BLE (Bluetooth Low Energy) and Zigbee technologies making the old rules of safety inapplicable.

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

As a result not only are users unprotected from the risks, they may have no idea their devices pose any risk at all. Securtle gives the users the most vital safety tool they can have: awareness. Securtle uses ambient risk communication so that users can consistently know the level of risk at any moment. This enables them to use whatever mitigation strategies which align with their knowledge and capabilities.

Securtle uses simple colors, changing hues, intensity, and adding a blinking feature, to communicate the risk level of a home network. As attacks flow through the network a home owner has no way of reasonably knowing that their own network is under attack. Securtle identifies the time that a home owner should disconnect and restart assuming color shifting shell to give users of all ages and knowledge levels a simple real time risk exposure reading of their digital environment. An appealing and portable turtle, Securtle uses both passive and active BL/BLE and WiFi tools to rate the real-time risks of a specific local installations of IoT devices. As a result the homeowner can do the one thing that can do –disconnect, restart, and utilize the built-in updating [3]. In the following sections, we discuss the design and implementation of Securtle. We also perform a short heuristic evaluation using Nielsen's Ten Heuristics for User Interface Design [7].

2 DESIGN

Furnell et all [4] showed that users will often make poor choices when given too much information due to 'security fatigue' from making numerous security specific decisions in their everyday life. This necessitates that a device which makes users aware of security attacks must be operational with minimum user intervention that would reduce the cognitive load on users. In this section, we present the Securtle design which is centered on conveying the active information in visual and auditory queues.

2.1 Purpose and Scope

Securtle is designed to be used either as a portable or stationary gadget, where it is possible to install it either at home or public networks. A patient with diabetes who uses a BLE insuline pump, can use Securtle to make sure whether the environment is safe enough to use the pump or a Starbucks manager can connect Securtle to its network to detect whether there is possibly a user with malicious intent using their network. Securtle is essentially a honeypot (Figure 1, which is a mechanism through which a system is able to detect unauthorized access by allowing malicious attackers to attack the honeypot first in the network [8]. they have been long used in security to detect and report information about the attacker through networking monitoring. Yet, to the best of authors' knowledge a portable honeypot with characteristics of Securtle is not studied. While Securtle might not discover all kinds of attacker information, it allows the users to: first, determine the level of threat by analyzing the frequency and deviations in nature of incoming traffic, and second, it measures how robust the system is and if it is able to protect its users from cyber attacks. As the first line of defense, Securtle ensure that the users and the system defenders are protected and gain knowledge about a possible security compromise before the actual system is compromised. Note that Securtle is not design to prevent the attacks or defend the users from the attacks. A gadget that can defend the users, devices and networks needs to be much more privacy-invasive with extensive privileges. This naturally is not achievable for a gadget that connects to public networks.

2.2 Design Decisions

Data visualization has been widely used to represent complex security information and notice. In a recent paper by Li et al. [6], the authors used visual tools to communicate which data to protect for security analysts to comprehend large amounts of information effectively and easily. In another study, Shakthidhar et al. [5] showed that the users actually adjust their security choices to colors

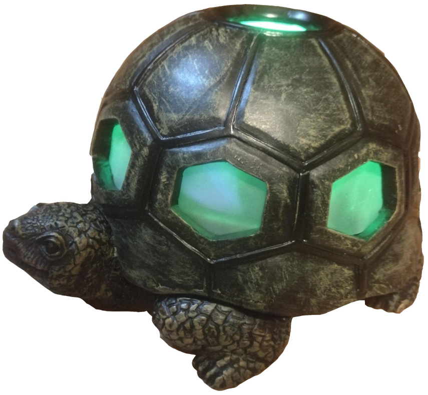

Fig. 1. Securtle prototype in a benign environment glowing green light

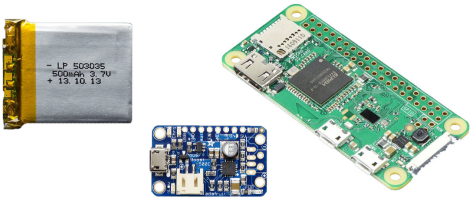

Fig. 2. Main components of Securtle. From left to right: 360 mAh battery, power management module and Raspberry Pi Zero.

and sounds. We follow a similar approach - in order to make cyber attacks more visible to users, we use visual indicators including color type (green, yellow, red) and color intensity (glowing versus solid) to indicate incoming cyber attacks and their severity. This is also a proactive approach to detecting security breaches which allows users to respond to it immediately instead of the reactive approach, where users are informed after a breach occurs.

In order to achieve our goal of building an intelligent sensing device for risk communication that performs network traffic analysis with minimal user intervention, we implemented different color rank modules for risks - Securtle would turn red if there was a high-risk attack and yellow if it was a low-risk attack. For all other times, it would remain a solid green color to indicate that there were no attacks on the system. This was consistent with Berlin and Kay’s work on colors and their associated meanings [2].

Similarly, we used a glowing indicator instead of a solid color indicator to highlight that there was a transition between safe (green) and unsafe (red, yellow) states. The indicator light would remain in a particular state for as long as Securtle was under attack. Aesthetically, we chose a turtle because the shell may provide the user with a sense of ‘protection’, and also an object with the likeness of a turtle would please users across generations in place of a rigid security device that stands out by blending in as an everyday decorative object.

We decided to create a first prototype that supports BL, BLE and WiFi due to the wide usage of these technologies and wide range of cyber threats in them. Threat detection in Securtle is designed in a modular way to make future development easier. New modules can be developed to support new attacks and threats or even newer technologies (more details in section 3). Although the first prototype uses visual queues, Securtle can easily be expanded to use auditory queues as well as Haptic feedback.

From a functional point of view, Securtle would indicate both occasional and persistent cyber attacks. One-time attacks over a short period of time would be indicated by an appropriate color with frequency-based blinking or may not be indicated if the threat is not persistent. On the other hand, persistent cyber attacks, which would be the sum of similar occasional attacks would be represented by a blinking yellow or red color depending on the severity. Securtle would also maintain a record of cyber attacks to analyze persistent attacks, and well as to implement machine learning algorithms in the future to understand the scope of these attacks for resource optimization.

3 TECHNICAL IMPLEMENTATION

Securtle has two key components: a physical interaction and an AI algorithm for identification of threats. The description of the implementation begins with the physical interaction then moves to the simple AI for detecting malicious activity from a series of events. The Securtle hardware prototype, shown in Fig 1 contains a small circuit board with BLE and Wifi capabilities, multi-color LEDs (Light Emitting Diodes), and power management circuits for the battery. The threat model identifies ambient ecosystem risks as well as the static and dynamic threats to the system. We have also implemented real world attacks perform proof of concept attacks on our prototype described in Section 3.1. Securtle is implemented with the Raspberry Pi Zero as shown in Fig 2 it gave us the widest range of options for testing, has the correct form factor, and is well documented. This choice resulted in a larger footprint due to the power module as well as a reasonably powerful battery.

The software implementation follows the super loop architecture used in embedded systems where there is an infinite loop in the program and all the tasks of the system are run periodically. In this case, that loop also spawns and handles threads and processes. Additional risk modules can be added to seek other threats, such as identifying attacks on Zigbee, or a SSH module that creates a decoy web server where any future attacks on this decoy server will be indicated as an increased risk. When such profile is activated Securtle emulates a selection of vulnerable version of services that are found on personal laptops to resemble a vulnerable laptop. Similarly, when the BLE profile is activated in a crowded place, Securtle will switch on BLE and will resemble a vulnerable BLE device such as cloudpets toys to detect potential attackers with active BLE scanners.

Currently Securtle is used just for risk-awareness and is not capable of defending the network. Right now it is designed for people to take the action they can take: disconnect, then force reboot to ensure installation of updates. They can also choose to make risk-averse decisions, for example choosing not to log into a bank account on a network when the turtle is yellow. Future implementation of Securtle will move to a custom Printed Circuit Board (PCB), allowing for a minimized, custom footprint, and ultra-low-power design. This approach would also aid in the implementation of several new modules for BLE, BL and Wi-Fi as well as support for other protocols like Zigbee and Z-Wave, if required.

3.1 Demo and Process Flow

Securtle was demoed in a local innovation competition. During the demo, first we performed the Cloudpets BLE attacks using the web bluetooh[1]. Then we switched off the Cloudpet Unicorn and switched on the BLE risk profile of Securtle where it was configured to create a decoy of the Cloudpets toys. The decoy pet was detected by the above-mentioned tool for hacking Cloudpets and attempts were made to penetrate the device. However, since Securtle is not a real Cloudpets toy the attack was unsuccessful. Meanwhile Securtle detected the attack and started glowing red.

The process flow of Securtle is illustrated in Fig 3. Each of the system tasks in Securlte are associated with a security-detection tasks, e.g., checking possible SSH brute force authentications on the system, and are performed by different 'risk modules'. Risk modules can be grouped in a 'risk profile' where every risk profile is mapped with a specific area. For example, a risk profile is associated with public WiFis and that profile is activated when Securtle connects to a public WiFi. After switching on the device, the process flow of Securtle starts by activating the risk profiles. The risk modules that are contained in the risk profile start running and just as they detect a threat based on a sequence of events in the monitoring phase, it will calculate the risk of the threat. This is followed by risk communication where based on the level of the threat, the proper indicator behavior is selected and activated. Afterwards, the user might take action, e.g., disconnect or even switch off critical devices or stop performing risky operations. The potential risk would eventually

[1] https://github.com/pdjstone/cloudpets-web-bluetooth

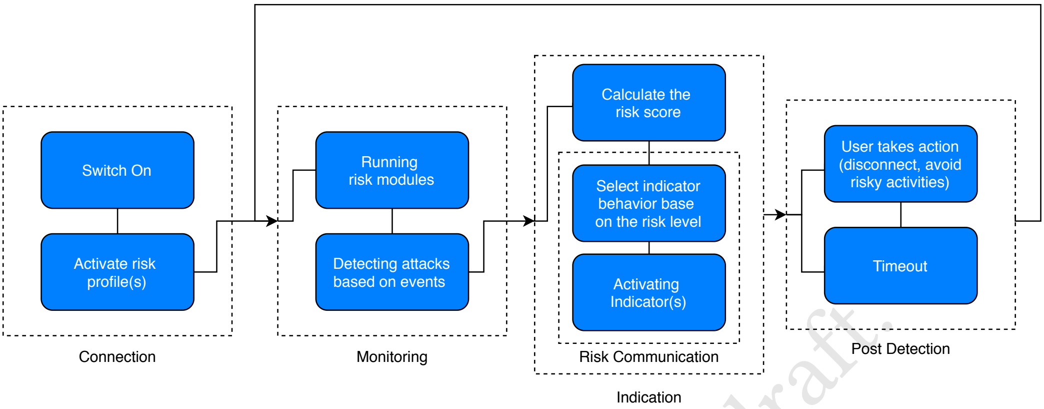

Fig. 3. The process flow from switching on the Securtle to possible user response, assuming that it is pre-configured to the network.

be timed out depending on the severity, e.g. a simple ping will timeout faster than a ssh brute force attack.

4 HEURISTIC EVALUATION

In its current version, Securtle is non-interacting, ambient security risk indicator that tracks and warns users of security incidents. Hence, in order to provide seamless interaction that would require minimum user intervention, we evaluated Securtle using Neilsen's ten heuristics [7] for interface design, discussing the applicability of each heuristic to our design. The ten heuristics are discussed below.

- **Visibility of system status: Present.** The system status is indicated though visual indicators in real-time if the user is present in the immediate vicinity. In order to improve accessibility, Securtle also allows an opportunity for possible upgrade for audio indicators that would let the user know if it is powered on or not.
- **Match between system and the real world: Present.** Securtle uses a standard signalling mechanism, much like traffic lights, in order to alert users. It is essentially a passive intrusion detection system, and this signalling mechanism ensures that it follows the logical order of color-based risk indicators that matches the real world. Again, to improve accessibility, we can always use simple, audio-based alerts to warn the user.
- **User control and freedom: Not applicable.** Users do not necessarily engage in performing operations on the device, and the system does not yet contain an interaction unit on the device that allows users to change settings. In order to respond to a cyber attack detected by Securtle, users have to go through traditional security defense measures without Securtle actively participating.
- **Consistency and standards: Present.** Securtle is consistent both on the front end and the back end. On the front end, it follows standard risk indicator design, and on the back end it follows a system implementation that is consistent with standard honeypot design that security network analysts are familiar with.
- **Error prevention: Absent.** As with all honeypot systems, there is a possibility of detecting false positive and false negatives. For example, Securtle may inform that a possible breach has occurred by red indicators even though it might have been a network anomaly and not a cyber attack. While this remains to be explored in future work, Artificial Intelligence

algorithms can be used to get user feedback on the correctness of the risk indicator and incrementally improve the system.

- **Recognition rather than recall: Not applicable.** Again, the user does not interact with the system except physically, and it is relatively easy to recognize Securtle as a risk indicating device.
- **Flexibility and efficiency of use: Present.** Securtle is designed so that both expert and non-expert users can be informed through the device. These is no learning curve, and the automated process makes the device very efficient without requiring manual configuration. The only requirement is that users are able to connect Securtle to the network that it is supposed to monitor. This is a one-time process and is similar to connecting any device to a Wifi network. However, in order to analyze security incident detected by Securtle, an intervention by a security expert might be needed.
- **Aesthetic and minimalist design: Present.** The minimalist nature of the interaction gives users and network defenders more time to focus on securing the network rather than interacting with the device. Securtle is designed much like a toy and is ambient by blending into a physical space.
- **Help users recognize, diagnose, and recover from errors: Absent.** While Securtle can be upgraded to actively prevent attacks, in its current form, it does not. Users are unable to interact with Securtle to respond to errors, and any hardware failure would require the intervention of an expert without automatic troubleshooting.
- **Help and documentation: Present.** Securtle is designed in a manner that the learning curve of interacting with it, or interpreting its responses, is minimal. This ensures that users do not need extensive documentation to be able to operate Securtle. However, in order for users to learn the basics and the back end, we have included documentation (not yet publicly available) for users who would like further help.

5 CONCLUSION AND FUTURE WORK

In this study we presented a design solution for an ambient intrusion detection system and briefly discussed the implementation of Securtle, a gadget that communicates the risk level of the network by creating decoy services not only on WiFi but on BL, and BLE as well. Securtle is modular and supports adding/removing risk modules, allowing the super users and communities to develop their own risk modules and risk profiles. For instance, adding Zigbee risk profile is among the first milestones of our future development. Although the current prototype only uses simple heuristics to determine whether there is a malicious intent behind a series of events, e.g., SSH (Secure Shell Hash) brute forcing after ping, and AI algorithms can be used to improve the efficiency and accuracy of detecting attacks.

As mentioned prior, Securtle currently utilizes the visual queues to communicate the risk. In future, we are going to include auditory queues and haptic feedback as well as we briefly mentioned in the heuristics above to ensure that it is accessible to visually challenged populations. A part of our future work will be dedicated to creating a mobile app for Securtle. That interface can be used to make the interaction smoother by allowing the users to activate or deactivate risk modules or risk profiles manually. It can also be used to give more detailed information about the ambient risk, e.g. details of the events, history of the events, etc. Additionally, a detailed user evaluation in future could help us understand the usability and acceptability of Securtle as people integrate it into their homes.

6 ACKNOWLEDGMENTS

Removed for review.

REFERENCES

[1] Farzaneh Asgharpour, Debin Liu, and L Jean Camp. 2007. Mental models of security risks. In *International Conference on Financial Cryptography and Data Security*. Springer, 367–377.

[2] Brent Berlin and Paul Kay. 1991. *Basic color terms: Their universality and evolution*. Univ of California Press.

[3] Kelly E Caine, Celine Y Zimmerman, Zachary Schall-Zimmerman, William R Hazlewood, Alexander C Sulgrove, L Jean Camp, Katherine H Connelly, Lesa L Huber, and Kalpana Shankar. 2010. DigiSwitch: design and evaluation of a device for older adults to preserve privacy while monitoring health at home. In *Proceedings of the 1st ACM international health informatics symposium*. 153–162.

[4] Steven Furnell and Kerry-Lynn Thomson. 2009. Recognising and addressing ‘security fatigue’. *Computer Fraud & Security* 2009, 11 (2009), 7–11.

[5] Shakthidhar Gopavaram, Omkar Bhide, and Jean L Camp. [n.d.]. Can You Hear Me Now? Audio and Icons in Privacy Permissions. ([n. d.]).

[6] Tianyi Li, Gregorio Convertino, Ranjeet Kumar Tayi, and Shima Kazerooni. 2019. What data should I protect? recommender and planning support for data security analysts. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*. 286–297.

[7] Jakob Nielsen. 1994. 10 Heuristics for User Interface Design. https://www.nngroup.com/articles/ten-usability-heuristics/. (Accessed on 10/09/2020).

[8] Lance Spitzner. 2003. *Honeypots: tracking hackers*. Vol. 1. Addison-Wesley Reading.

[9] Alex Ulmer, David Sessler, and Jörn Kohlhammer. 2019. NetCapVis: Web-based Progressive Visual Analytics for Network Packet Captures. In *2019 IEEE Symposium on Visualization for Cyber Security (VizSec)*. IEEE, 1–10.

[10] Chenxi Wang and John C Knight. 2000. Towards survivable intrusion detection. In *Proceedings of the 3rd Information Survivability Workshop (ISW-2000), Boston, USA*. Citeseer.