

Identifying Threats, Cybercrime and Digital Forensic Opportunities in Smart City Infrastructure via Threat Modeling

Yee Ching Tok 

*Singapore Univ. of Tech. and Design
Singapore
yeeching_tok@sutd.edu.sg*

Sudipta Chattopadhyay 

*Singapore Univ. of Tech. and Design
Singapore
sudipta_chattopadhyay@sutd.edu.sg*

Abstract—Technological advances have enabled multiple countries to consider implementing Smart City Infrastructure to provide in-depth insights into different data points and enhance the lives of citizens. Unfortunately, these new technological implementations also entice adversaries and cybercriminals to execute cyber-attacks and commit criminal acts on these modern infrastructures. Given the borderless nature of cyber attacks, varying levels of understanding of smart city infrastructure and ongoing investigation workloads, law enforcement agencies and investigators would be hard-pressed to respond to these kinds of cybercrime. Without an investigative capability by investigators, these smart infrastructures could become new targets favored by cybercriminals.

To address the challenges faced by investigators, we propose a common definition of smart city infrastructure. Based on the definition, we utilize the STRIDE threat modeling methodology and the Microsoft Threat Modeling Tool to identify threats present in the infrastructure and create a threat model which can be further customized or extended by interested parties. Next, we map offences, possible evidence sources and types of threats identified to help investigators understand what crimes could have been committed and what evidence would be required in their investigation work. Finally, noting that Smart City Infrastructure investigations would be a global multi-faceted challenge, we discuss technical and legal opportunities in digital forensics on Smart City Infrastructure.

I. INTRODUCTION

In this increasingly interconnected world, humans generate a lot of data in their daily lives through the usage of computing devices and the multitude of increasingly accessible technologies that enable a better quality of life. The technologies include using smart home devices (popularly termed as Internet-of-Things (IoT)) and interaction with novel technical implementations by governments to improve citizens' quality of life. These novel technical implementations range from smart water and electricity meters, smart vehicles, autonomous vehicles and building automation systems [1]. These governments' final goal was to presumably expand the systems into Smart City Infrastructure (SCI) for a better overview of their citizens, environment, safety and resources. Figure 1 illustrates how governments and city planners could gather various data points to obtain a holistic overview of the country, allowing them to provide timely governance intervention where needed.

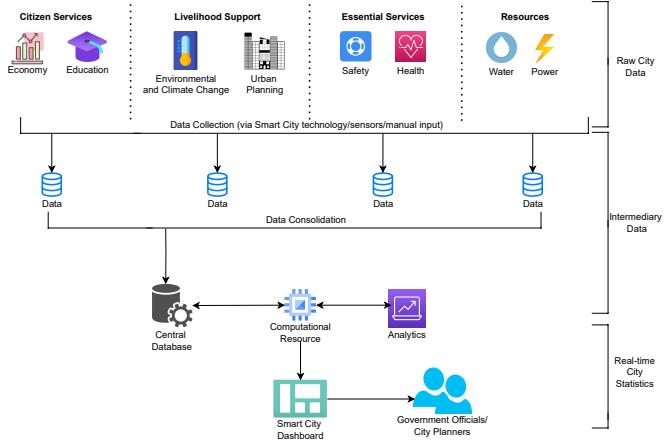


Fig. 1. Overview of Smart City Infrastructure (SCI) Data Collection and Processing

Unfortunately, governments and city planners are not the only ones embracing the adoption of technologies related to SCI. These new systems offer attractive opportunities to state-sponsored adversaries and cybercriminals. If SCI data could be stolen or intercepted, adversaries and cybercriminals could gain access to a plethora of data that could be further exploited. For example, the data could be exploited to cause economic issues to a target country (e.g., sudden restrictions on export of resources or medical supplies) or events intended to destabilize the country (e.g., sudden outbreak of diseases while a country's hospitals are nearing maximum capacity). The adversaries could cause further chaos if they successfully breach SCI systems and trigger anomalous conditions (e.g., opening/closing valves in critical infrastructure or overriding safety mechanisms) to cause emergency shutdowns, destruction of facilities or general chaos in transport systems.

Digital forensic investigators (DFI) and law enforcement agencies (LEA) have been crucial in investigating cyberattacks and cybercrime. However, DFI and LEA will likely require further support if they are called upon to investigate attacks on SCI as described earlier. This is because DFI and LEA personnel are typically not familiar of the unique characteristics of SCI systems as compared to conventional digital systems.

Moreover, DFI and LEA are likely to have other ongoing digital investigations. Their ongoing case commitments may prevent them from being able to perform research on SCI systems to identify the required evidence. Meanwhile, SCI system owners may also be unable to provide the necessary evidence since these requirements were not specified in the first place when SCI was implemented. Caviglione et al. [2] stressed that digital forensic investigation on IoT devices would have to be performed on small-scale implementations (such as homes) to large-scale deployments such as those within a smart city. It was further stated that the technology used to build the infrastructure would be diverse and that their accessibility features must be discerned and appraised independently for forensic investigations [2]. The predicament has not gone unnoticed - there have been calls for further digital forensic research on new digital artifacts and preventing misinterpretations of artifacts [3].

Digital forensics and cybersecurity in smart cities had been previously discussed, where the definition of smart cities was based on National Institute of Standards and Technology (NIST) and only focused on smart environments, living and mobility [1]. While the threats, forensic data, and data sources highlighted did provide some guidance for DFI and LEA [1], the provided information cannot apply to all SCI internationally. For example, different countries have different technical requirements, implementation and data needs for their SCI. This could prove problematic for solutions that suggest a one-size-fits-all approach by specifically listing functional systems, such as the ones suggested by Baig et al. [1].

Due to the complexity of SCI, digital evidence for various components of SCI must be identified before actual cybercrime occurs. This is to reduce the stress faced by DFI and LEA as first responders to cybercrime. DFI and LEA could also better handle SCI cybercrime challenges if a standard definition of SCI is achieved, along with potential threats, offences and evidence sources pre-identified. Although this is a global challenge, adhering to global initiatives and standards allow flexibility of adoption by international DFI and LEA.

The contributions of our research are summarized as follows:

- 1) We highlight current issues in SCI and define a standardized definition of SCI.
- 2) We develop and make publicly available our threat model template to governments, DFI and LEA to identify threats in SCI.
- 3) We map SCI threats to possible offences and corresponding SCI evidence sources and types.
- 4) We discuss future SCI digital forensics opportunities from a technical and legal perspective.

For reproducibility and advancing the research in SCI digital forensics and threat modeling, our threat model is publicly available at: <https://github.com/poppopretn/SmartCityThreatModel>.

The rest of this paper is organized as follows. In Section II, we present the context of the paper and define SCI. In Section III, we highlight the choice of our threat modeling

methodology, showcase our threat model and present the threats identified in SCI. In Section IV, we show the threats, offences, evidence sources and types within SCI that we derived using our threat model. In Section V, we discuss future technical and legal opportunities for SCI digital forensics. In Section VI, we explain the limitations of our research. In Section VII, we summarize current related work in SCI digital forensics. Finally, we conclude the paper in Section VIII.

II. CONTEXTUALIZING SMART CITY INFRASTRUCTURE

Technological innovations and rapid deployment of Internet of Things (IoT) devices have transformed many cities in different geographical regions into smart cities [4]. Arguably, these cities may not have implemented sufficient infrastructure that can deliver futuristic societal outcomes such as accident-free environments or zero-waste scenarios. However, these current implementations have brought about positive changes such as moulding the design of future cities and achieving sustainable use of resources [4].

A. Current Issues in Smart City Infrastructure

The implementation of SCI is an attractive option for governments looking to improve citizens' lives and has increased visibility to critical indicators such as resource utilization and public safety. Nonetheless, there are multiple challenges to implement such capabilities as outlined below:

- 1) **Definition Issues:** It is vital to set the right context and definition when SCI is discussed. From the academic perspective, a commonly agreed definition of a Smart City has yet to be agreed on. A brief literature review of papers regarding SCI was conducted and yielded at least three different definitions of a Smart City [5]–[7]. SCI is not clearly defined from the industry perspective either. Various industry solutions such as Bosch [8], Cisco Kinetic for Cities [9], Microsoft CityNext [10] and Schneider Electric EcoStruxure [11] have offered products touted to allow prospective customers to create smart cities. However, a review of their respective product briefs showed that these solutions appear not to be based on any commonly agreed upon definitions or standards. Many also fail to realize that such forms of definition are constrained by financial budgets and technological maturity of the location SCI are deployed in. This inevitably contributes to the problem of a lack of standard definition in SCI.
- 2) **Interoperability Issues:** This is an extension of the definition issue mentioned previously. There were at least 31 different vendors [12] offering various platforms and technologies to build SCI. It is unlikely that any one vendor could meet all the design requirements (hardware and software) of a city/country. A more realistic outcome would be a myriad of vendors being chosen to implement technologies by their respective strengths. With such a gamut of sensors, protocols, and technologies, interoperability between vendors becomes an issue. Although entities such as FIWARE [13] and the

TALQ Consortium [14] offer Application Programming Interfaces (APIs) to allow interoperability of technologies with different vendors, adoption and implementation of such APIs remain unclear.

- 3) **Cybercrime Issues:** Cyber attacks on smart cities could become the next issue governments face as such projects are implemented. From the legal perspective, respective laws possibly were not updated to include attacks on SCI. Meanwhile, from the law enforcement and incident response perspective, there could be a lack of experience, knowledge and training for professionals called upon to investigate such attacks. This issue is further exacerbated by the definition and interoperability issues. There is no standard definition of a smart city, and multiple technologies are being utilized in a smart city. Legal, remediation and law enforcement actions are hampered due to varying understanding of SCI and technology complexities, leading to a risk of misleading evidence being retrieved and presented to courts of law.

A properly defined and widely accepted definition of SCI could address the issues highlighted above. For example, a properly defined SCI will facilitate and empower Digital Forensic Investigators (DFI) in peer review processes, especially at Levels 3 and 4 of the Peer Review Hierarchy as proposed by [15]. It also facilitates interoperability between vendors and enhances the development of APIs. Finally, it enhances clarity in connections between evidence and criminal hypotheses, reducing the risks of misleading evidence being presented in courts [16].

B. Defining Smart City Infrastructure

The primary issue originates from a lack of a standard definition in SCI as various vendors and entities are vying to be the standard for smart cities. Geographical differences and individual governmental requirements have not helped foster a standard definition of a smart city. A way to transcend such challenges in defining a smart city, along with relevant data indicators is required to facilitate the resolution of issues raised in Section II-A.

As a body that strives to standardize methods to accomplish a final goal, the International Organization for Standardization (ISO) facilitates such an endeavour. After extensive research, we identified multiple ISO standards that provided a suitable framework for a standardized definition of SCI. With reference to Figure 2, the ISO standards that were selected are as follows:

- 1) ISO37101:2016 [17]
- 2) ISO37120:2018 [18]
- 3) ISO37122:2019 [19]
- 4) ISO37123:2019 [20]

As observed from Figure 2, our proposed SCI definition (i.e. the body) is guided by four ISO standards. ISO37101:2016 serves as the guiding vision of a Smart City (i.e., a skeleton) via sustainable development. Concurrently, corresponding core and supporting data indicators from Factor #1

Smart City Infrastructure (SCI) Definition		
<i>Vision:</i> Sustainable development in communities (ISO37101:2016)		
<i>Factor #1:</i> Indicators for City Services and Quality of Life (ISO37120:2018)	<i>Factor #2:</i> Indicators for Smart Cities (ISO37122:2019)	<i>Factor #3:</i> Indicators for Resilient Cities (ISO37123:2019)

Fig. 2. Defining SCI

(ISO37120:2018), Factor #2 (ISO37122:2019) and Factor #3 (ISO37123:2019) provide the required context of the SCI (i.e., the muscles/flesh). Core data indicators are mandatory data indicators that must be captured if the ISO standards are used, whereas supporting data indicators are recommended to be captured (but not mandatory). We further explain our choice as follows:

- 1) **ISO37101:2016** - In contrast to a technical view and definition of SCI, the standard ISO37101:2016 adopts a technology-agnostic approach and uses sustainable development as a common denominator before any form of technical implementation is utilized. Table I lists the underlying fundamental purposes of sustainability in modern society along with sustainability issues raised in ISO37101:2016. Since the issues are systematically listed out before any form of smart city technology is implemented, ISO37101:2016 serves as a suitable component to drive the vision aspect of our proposed SCI definition (see Figure 2).

TABLE I
SUSTAINABILITY PURPOSES AND ISSUES HIGHLIGHTED IN ISO37101:2016

Purposes of Sustainability	Sustainability Issues
<ol style="list-style-type: none"> 1. Attractiveness (e.g., Culture, identity) 2. Preservation and improvement of environment (e.g., Protection of biological diversity and ecosystem) 3. Resilience (e.g., Climate change adaptation, economic shock preparedness) 4. Responsible resource use (e.g., Sustainable production, reusing and recycling of materials) 5. Social cohesion (e.g., Diversity, sense of belonging, social mobility) 6. Well-being (e.g., Happiness, healthy environment, quality of life) 	<ol style="list-style-type: none"> 1. Governance, empowerment and engagement 2. Education and capacity building 3. Innovation, creativity and research 4. Health and care in the community 5. Culture and community identity 6. Living together, interdependence and mutuality 7. Economy and sustainable production and consumption 8. Living and working environment 9. Safety and security 10. Community infrastructures 11. Mobility 12. Biodiversity and ecosystem services

- 2) **ISO37120:2018, ISO 37122:2019 and ISO37123:2019**
 - Building on the technology-agnostic approach of ISO37101:2016, core and supporting data indicators with respect to measuring city services and quality of life (ISO37120:2018), smart cities (ISO37122:2019) and resiliency (ISO37123:2019) are outlined. The data indicators used in the three ISO standards are listed

in Table II. While it is tempting to solely focus on data indicators listed in ISO37122:2019 instead of also using ISO37120:2018 and ISO37123:2019, the standard ISO37120:2018 has stated that ISO37122:2019 and ISO37123:2019 have to be used in conjunction with ISO37120:2018 [18]. Further examination of ISO37122:2019 and ISO37123:2019 also yielded the same considerations that the three ISO standards must be used in tandem to provide a complete overview of a modern city [19], [20]. Finally, another added advantage is that the data indicators in these three ISO standards can also be mapped back to the 17 United Nations Sustainable Development Goals (SDGs) 2015 [18]–[21]. This would allow governments to have better visibility in their progress towards the United Nations SDGs.

TABLE II
DATA INDICATORS LISTED IN ISO37120:2018, ISO37122:2019 AND
ISO37123:2019

Clause Number	Data Indicators
5	Economy
6	Education
7	Energy
8	Environment and Climate Change
9	Finance
10	Governance
11	Health
12	Housing
13	Population and Social Conditions
14	Recreation
15	Safety
16	Solid Waste
17	Sport and Culture
18	Telecommunication
19	Transportation
20	Urban/Local Agriculture and Food Security
21	Urban Planning
22	Wastewater
23	Water

Based on the information provided in the preceding paragraphs, we formally define SCI as follows:

Definition 1: Smart City Infrastructure (SCI) are infrastructure designed to fulfil and address the six purposes of sustainability and twelve sustainability issues (based on ISO37101:2016 and outlined in Table I), with technical implementations that provide visibility to data indicators set out in ISO37120:2018, ISO37122:2019 and ISO37123:2019 (outlined in Table II).

III. THREAT MODELING SMART CITY SYSTEMS

New systems likely have to be designed and built to capture the data indicators based on our proposed definition of SCI (*i.e.*, Definition 1). There could be multiple opportunities where cybersecurity threats creep into the designed systems despite endeavours to reduce such threats. Hence, it is imperative to ensure such systems are appropriately classified and devise a concrete means to identify potential threats. An appropriate Threat Modeling methodology should also

be selected, with a means of clearly documenting identified threats to facilitate threat mitigation efforts.

A. Classification of Smart City Infrastructure Systems

Multiple data indicators (Clauses 5 to 23) shown in Table II appear to be daunting at first [18]–[20]. However, after careful analysis, we were able to further classify the data indicators based on their respective outcomes into 4 distinct groups. With reference to Figure 3, the 4 groups are **Citizen Services**, **Livelihood Support**, **Essential Services** and **Resources**. These identified groups were used as a foundation for our proposed Smart City infrastructure system design and threat model.

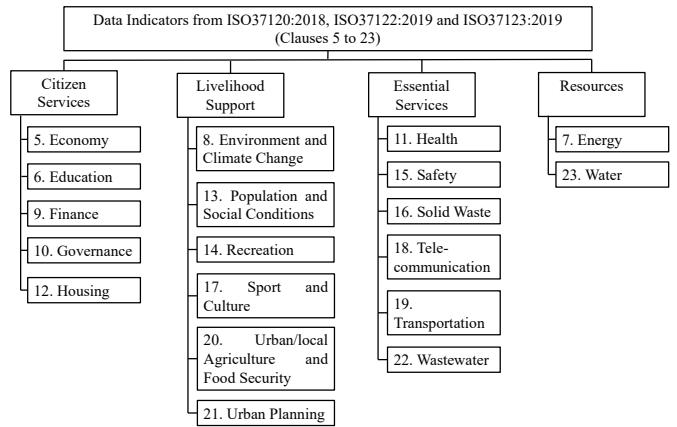


Fig. 3. Smart City Data Indicators Classification

B. Selection of Threat Modeling Methodology

Threat modeling is defined as using abstractions to assist in the identification of risks [22]. It is best executed when systems are still in the design phase since identified risks can be addressed by modifying the design of the systems before they are formally implemented. Ideally, it should be executed after any proposed design change to ensure that potential threats are not introduced due to the change in the system. Having said that, threat modeling is still useful and important for systems that have deployed as well. However, it is likely that additional efforts and resources would have to be utilized to address the risks identified. This is because such systems would have been operational. For example, if the rectification of risks identified in an operational system requires it to be switched off, the system owners would have to divert operational requirements to another asset. This may, in turn, introduce complexities and potential risks. As such, it is highly encouraged that threat modeling should be done before any system is formally deployed.

Given a wide range of threats faced by various systems, multiple different threat modeling methodologies have been proposed throughout the years to identify threats in different scenarios. A collection of threat modeling methodologies have been identified and documented by Shevchenko et al. [23] and presented as follows:

- 1) **STRIDE:** Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, Elevation of Privilege (STRIDE) was first formulated in 1999 by Loren Kohnfelder and Praerit Garg [24]. STRIDE has been successfully used in modeling threats in Cyber-Physical Systems [22], and a software modeling tool is also available [25].
- 2) **PASTA:** Process for Attack Simulation and Threat Analysis (PASTA) was developed in 2012 and has seven stages of modeling activities. It requires both technical and non-technical stakeholders to participate in the process [26].
- 3) **LINDDUN:** Linkability, Identifiability, Non-Repudiation, Detectability, Disclosure of Information, Unawareness, Non-Compliance (LINDDUN) is a six-step process that is focused on modeling privacy-related threats. A noticeable drawback of this method lies in its complexity and labor-intensive process [23].
- 4) **Attack Trees:** This process was originally developed by Bruce Schneider in 1999 [27] and generally requires highly skilled individuals in cybersecurity to be involved in the process. It is a manual process without any guidelines for attack/risk identification, though some attempts to automate the process has been explored [28].
- 5) **OCTAVE:** Operationally Critical Threat, Asset, and Vulnerability Evaluation (OCTAVE) is a risk-based and strategic approach that consists of three phases [29]. It takes a different approach to threat modeling and identifies organizational risks instead of technical risks.
- 6) **VAST:** Visual, Agile, and Simple Threat (VAST) Modeling is a threat modeling strategy utilized by a commercial tool named ThreatModeler [30]. VAST uses application and operational threat models - Process Flow Diagrams (PFDs) are created for application threat models, and operational threat models are built from analyzing the PFDs. Although VAST is useful, detailed specifications of VAST are not available since it is a commercial product.
- 7) **CVSS:** Common Vulnerability Scoring System (CVSS) is a method that assigns a numerical score to a vulnerability based on its severity [31]. Widely used to determine the severity of vulnerabilities, it is not suitable to be used on its own in SCI as it would only display potential vulnerabilities in a numerical format. CVSS is often used in conjunction with other threat modeling methods [23].
- 8) **Trike:** Introduced in 2006, Trike is a threat modeling method that could generate threats and attack trees automatically [32]. However, the methodology and tools have not been updated for a long time and may not be suitable for modern SCI.
- 9) **Persona non Grata:** Persona non Grata is a threat modeling method that focuses on the human aspects of threats [33]. Although it has low false positives, it only identifies the human aspects of threats.
- 10) **Security Cards:** Developed in 2013, Security Cards is geared towards brainstorming about non standard situations and rarely used in the industry [23].
- 11) **Quantitative Threat Modeling Method:** Developed in 2016 and synergizing STRIDE, Attack Trees and CVSS, the creators aimed to address issues arising from complex interdependence in components [34]. Unfortunately, while this method appears to be comprehensive and structured, the tool mentioned in the paper was not released publicly for usage.
- 12) **hTMM:** Hybrid Threat Modeling Method (hTMM) combines threat modeling methodologies such as STRIDE, Persona non Grata and Security Cards for a focused threat modeling activity [35]. However, the process is still largely manual and lacks a dedicated tool that can execute the entire process of hTMM.

TABLE III
COMPARISON OF VARIOUS THREAT MODELING METHODOLOGIES

Threat Modeling Methodology	Software Tool Support	Ongoing Updates	Suitability for Smart City TM
STRIDE	✓	✓	✓
PASTA	✗	✗	✓
LINDDUN	✗	✗	✓
Attack Trees	✓	✗	✗
OCTAVE	✗	✗	✗
VAST Modeling	✓	✓	✓
CVSS	✗	✗	✓
Trike	✗	✗	✗
Persona non Grata	✗	✗	✗
Security Cards	✗	✗	✗
Quantitative Threat Modeling Method	✗	✗	✗
hTMM	✗	✗	✗

1) *Assessment Criteria for Threat Modeling Methodologies:* We examined all twelve threat modeling methodologies to assess if there were any methodologies that would be a good fit for our research objectives. We believe that all the twelve threat modeling methodologies were worth examining as discussed by Shevchenko et al. [23] accounts for the most popular and state-of-the-art threat modeling techniques. There were three key requirements - support for software tools based on the threat modeling methodology, ongoing updates and suitability for Smart City threat modeling.

These three essential requirements were identified as significant and relevant when determining the suitability of the chosen threat modeling methodologies. Firstly, a threat modeling methodology that is supported by a software tool allows collaborative work, enhances design efficiency and reduces potential errors when designing a threat model (e.g. in-built notations and feedback notifications). Secondly, a constantly updated threat model methodology ensures that it remains relevant to the ever-changing cybersecurity landscape. Finally, not all threat modeling methodologies may apply to the SCI

context. As such, it is vital to identify which methodology is appropriate for SCI.

The details of our assessment are summarized in Table III, and we selected STRIDE as our main Threat Modeling methodology for SCI systems designed based on Definition 1. There were a multitude of reasons why STRIDE was selected. Firstly, a freely available and consistently updated threat modeling tool was available. As research into threats on SCI should be an ongoing process, the tool used for threat modeling should also be actively maintained. Secondly, the STRIDE threat modeling methodology pays close attention to threats in the systems while being less complex than other methodologies such as LINDDUN, PASTA and VAST. Although it was also argued that STRIDE could grow in complexity as the systems grow larger [23], the availability and usage of the Microsoft Threat Modeling Tool (TMT) [25] could reduce the potential complexity. Finally, the Microsoft TMT is freely available for all, while methodologies such as VAST and tools such as IriusRisk [36] remain as commercial products. As costs associated with security can sometimes be an important factor, especially in novel systems, we aim to remove that barrier and foster the creation of secure systems during the design phase.

C. Threat Modeling Smart City Infrastructure

As explained in Section III-B, the STRIDE threat modeling methodology and the Microsoft TMT are used to create the corresponding threat model of the proposed Smart City System based on Definition 1. A threat model designed with the Microsoft TMT can contain up to four layers, and are as follows [37]:

- 1) **Layer 0:** This is known as the system layer (it can also be called the context layer), and is a compulsory layer in a threat model designed in Microsoft TMT as it is deemed as a starting point for any system [38]. Layer 0 contains core parts and processes of the system, and the corresponding data-flow diagrams should fit into a single page [38].
- 2) **Layer 1:** This is known as the process layer, and contains the secondary system parts that originate from Layer 0. Like the system layer, the diagram should also be contained within a single page.
- 3) **Layer 2:** This is known as the sub-process level, and contains details of the various components that the secondary systems are composed of. This layer is utilized when the corresponding systems handle sensitive data or are deemed as high-risk.
- 4) **Layer 3:** This is known as the lower-level layer, and used when a kernel-level system is constructed. Every process and sub-process is described in detail via data-flow diagrams.

Based on our defined Smart City Infrastructure, we have deliberately designed the threat model to be up to Layer 1. Firstly, we envision our threat model to be used and applicable to the international audience and community. Different governments and countries may have additional requirements and

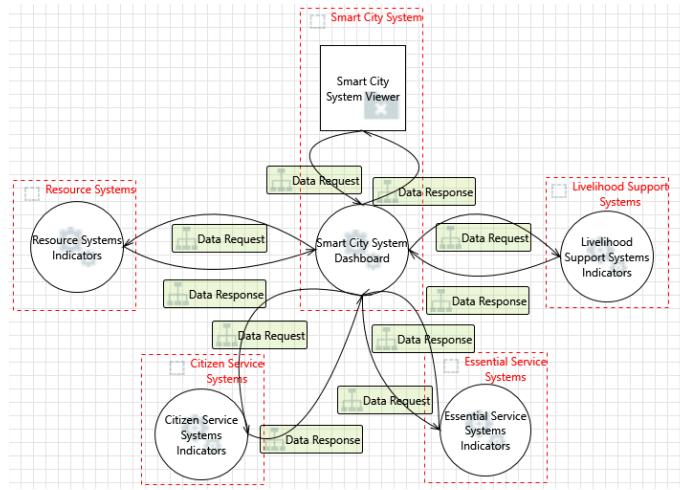


Fig. 4. Smart City System Threat Model - Layer 0

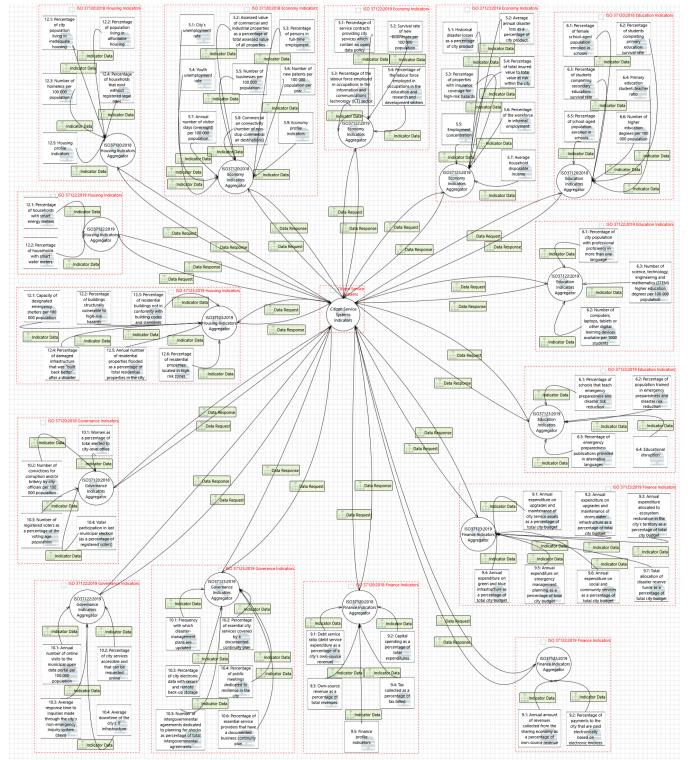


Fig. 5. Smart City Citizen Service System Threat Model - Layer 1

technological preferences for their Smart City Infrastructure. If we had designed our threat model down to Layer 3, that would have been very specific. Being overly detailed and specific could result in our threat model being overlooked instead. This is because it would require further efforts to review and amend the threat model. We have assessed that designing our threat model up to Layer 1 offered the best baseline - It was not overly simplistic and contained the necessary core and secondary systems required to capture the data defined by Definition 1. It also sets the stage for potential users to customize their desired Layers 2 and 3 threat models (if applicable). With reference to Figure 4 and Figure 5, a

screenshot of Layer 0 and Layer 1 of the Citizen Service System (one of the secondary systems) is shown.

In the next section, we discuss the threats analyzed from our SCI threat models (Figure 4 and Figure 5).

D. Threats in Smart City Infrastructure

The threat modeling activity in Section III-C yielded interesting results. There were a total of 1768 potential threats across the six threat categories of STRIDE. The total number and types of each threat category are summarized in Table IV and Figure 6, with threats from the Elevation of Privilege category having the greatest variety of threats.

TABLE IV
TYPES OF THREATS IDENTIFIED DURING THREAT MODELING

Category (Total Number of Threats)	Types of Threats
Spoofing (485)	1. Spoofing in various processes 2. Spoofing of various source data stores
Tampering (114)	1. Potential lack of input validation
Repudiation (114)	1. Potential data repudiation
Information Disclosure (370)	1. Data flow sniffing 2. Potential weak access control for a resource
Denial of Service (228)	1. Potential process crash or stop 2. Data flow potentially interrupted
Elevation of Privilege (457)	1. Systems may be subject to elevation of privilege using remote code execution 2. Elevation using impersonation 3. Cross site request forgery 4. Elevation by changing execution flow in various systems

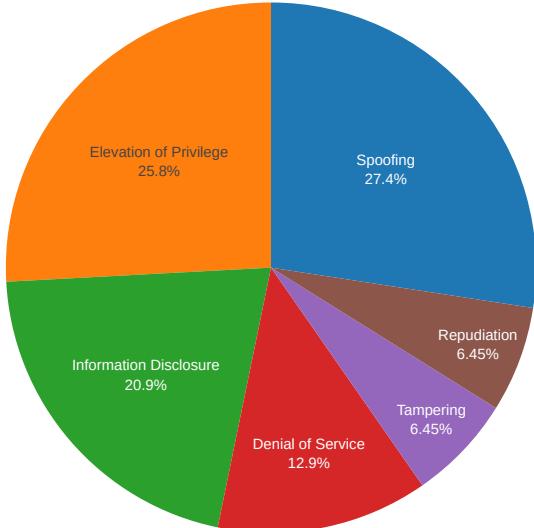


Fig. 6. Overview of Threats Derived from Threat Model Grouped by Category

Figure 7 shows the breakdown of potential threats in our proposed SCI systems, with the Smart City System Dashboard (Level 0) and Resource Systems Indicators (Level 1) having a lower percentage of potential threats (5.09% and 11.3%, respectively). The smaller number of threats was expected as the Smart City System Dashboard only contained major system processes and was less complex since it was at Layer 0. Meanwhile, the Resource Systems Indicators had the smallest number of data indicators (as shown in Figure 3). Thus, they had a smaller number of potential threats than the other Level 1 systems.

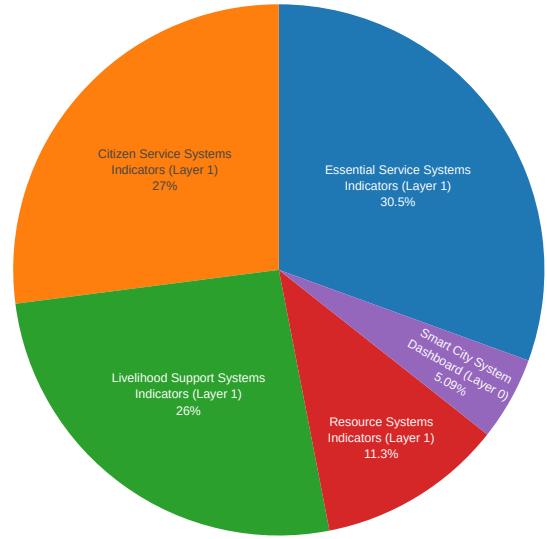


Fig. 7. Overview of Threats Derived from Threat Model

From the threat mitigation perspective, stakeholders can consider a few possible approaches based on the information presented in the preceding (i.e., Table IV, Figure 6 and Figure 7). For a quick-win approach, stakeholders could consider mitigating threats related to the Tampering and Repudiation category and target the Smart City System Dashboard and Resource Systems Indicators first, as they are the lowest in number. If the stakeholders are interested in mitigating a system with the highest number of potential threats and its corresponding category, the Essential Service Systems Indicators and Spoofing category threats could be reviewed. Ultimately, the threat model allows stakeholders to start studying or mitigating any threat they wish. Armed with the knowledge of potential threats that the SCI may face, defenders could also use these threats to anticipate potential cybercrime occurring on these systems.

In the next section, we leverage our SCI threat models to identify the cybercrimes on SCI systems.

IV. CYBERCRIME ON SMART CITY SYSTEMS

Just like SCI, the term "cybercrime" could have differing interpretations. It is vital to iron out a few key terms that may cause confusion and set the scope of discussion. After defining cybercrime for SCI, we map threats to offences that could

occur in SCI and, finally, the associated evidence classification and requirements.

A. Defining Cybercrime in Smart City Systems

Based on our literature review, we observed a few key terms that appeared to be used interchangeably, which could cause confusion and disagreement in our paper if not defined appropriately. The terms in question are "cybercrime", "computers", and "computer crime". The exact definition of cybercrime has been debated at length, especially with the evolution of the threat landscape over the years. Gillespie [39] asserts that it is a crime committed or facilitated by the Internet, with no distinguishing factor as to how access to the Internet was obtained. Meanwhile, with the advancement of technology, computers now come in various form factors and peripherals. Given the broad spectrum of hardware configuration, formally defining what a "computer" entails is difficult. The Budapest Convention on Cybercrime gave a broad definition and considered devices capable of transmitting data (whether in silo or networked) as a computer [40]. Finally, "cybercrime" and "computer crime" have often been used interchangeably to describe illegal activities on systems. Going by how the Budapest Convention on Cybercrime defines computers, "computer crime" should be defined as crimes carried out using computers. Since this paper uses the term "cybercrime", our scope would be limited to the crimes committed or facilitated by the Internet as described earlier.

B. Mapping Threats to Cybercrime and Offences

Given the borderless nature of the Internet, cybercrime is inevitably an international issue that is further complicated by geopolitical boundaries. In an attempt to achieve some form of consensus between nations and various stakeholders, the Budapest Convention on Cybercrime (also known as the Budapest Convention) was formed on 23rd November 2001 by the Council of Europe [40]. As of April 2022, a total of 81 States are now either Parties (66), or have signed it or been invited to accede (15) [41]. This accounts for only about 42% of the represented members in the United Nations (there are currently 193 members in the United Nations [42]). Notable absentees from the signatories of the Budapest Convention include the Russian Federation and China (both permanent members of the United Nations Security Council). The absence of their support for the Budapest Convention could be due to geopolitical tensions and that they were not consulted when the Budapest Convention was first formulated, an observation that was also noted by Gillespie [39].

Unfortunately, there is yet a global consensus on cybercrime. Geopolitical issues, potential lack of awareness and urgency by international governments inadvertently facilitate cybercrime. However, there is a strong need for a collective agreement for cybercrime to be addressed primarily due to its potential global reach. This is especially so if various governments are working towards the implementation of SCI. Academic research, coupled with suitable theoretical models and adaptation of a sensible definition of cybercrime should

help governments be more cognizant of the crimes that could occur.

TABLE V
MAPPING THREATS ON SCI TO CRIMES AND OFFENCES

Type of Crime (based on [39])	Threats on SCI (based on Table IV)	Possible Offences Committed (based on [40])
Crimes against computers	Spoofing in various processes	Illegal Access (Article 2), Illegal Interception (Article 3), System Interference (Article 5), Misuse of Devices (Article 6)
	Spoofing of various source data stores	Illegal Access (Article 2), Illegal Interception (Article 3), Data Interference (Article 4), System Interference (Article 5), Misuse of Devices (Article 6), Computer-related Forgery (Article 7)
	Potential lack of input validation	Data Interference (Article 4), Misuse of Devices (Article 6), Computer-related Forgery (Article 7)
	Potential data repudiation	Data Interference (Article 4), System Interference (Article 5), Misuse of Devices (Article 6), Computer-related Forgery (Article 7)
	Data flow sniffing	Illegal Access (Article 2), Illegal Interception (Article 3), Misuse of Devices (Article 6)
	Potential weak access control for a resource	Illegal Access (Article 2), Data Interference (Article 4), System Interference (Article 5), Misuse of Devices (Article 6)
	Potential process crash or stop	Illegal Access (Article 2), Illegal Interception (Article 3), Data Interference (Article 4), System Interference (Article 5), Misuse of Devices (Article 6)
	Data flow potentially interrupted	Illegal Access (Article 2), Illegal Interception (Article 3), Data Interference (Article 4), System Interference (Article 5), Misuse of Devices (Article 6)
	Systems may be subject to elevation of privilege using remote code execution	Illegal Access (Article 2), Illegal Interception (Article 3), Misuse of Devices (Article 6)
	Elevation using impersonation	Illegal Access (Article 2), Illegal Interception (Article 3), Misuse of Devices (Article 6)
Cross site request forgery	Cross site request forgery	Illegal Access (Article 2), Illegal Interception (Article 3), Misuse of Devices (Article 6)
	Elevation by changing execution flow in various systems	Illegal Access (Article 2), Illegal Interception (Article 3), System Interference (Article 5), Misuse of Devices (Article 6), Computer-related Forgery (Article 7)

After evaluating state-of-the-art works on cybercrime, we assessed that the adaptation of a subset of the taxonomy of cybercrime categories proposed by Gillespie [39] and a subset of the offences outlined by the Budapest Convention would be

suitable to represent cybercrime in SCI. We mapped the types of threats outlined in Table IV to the corresponding cybercrime categories and offences, as shown in Table V below.

There were other categories that Gillespie [39] proposed (e.g., crimes against property, illicit content and individuals). Similarly, the Budapest Convention had also included other types of offences e.g., child pornography, infringement of copyrights and abetting crimes. These are excluded from Table V for being out of scope as we only focus on crimes on SCI.

C. Evidence Requirements for SCI Threats and Offences

Although the issue of pinpointing offences to threats on SCI has been addressed, it is necessary to identify the evidence required to prove such offences have occurred. Based on our earlier efforts in identifying types of threats during threat modeling SCI and the associated offences in Table V, we further identified the possible evidence types that investigators could use. The evidence types were identified by referring to the threats on SCI and examining the corresponding possible evidence sources. After the evidence sources were determined, types of possible evidence (sorted in order of volatility [43]) linked to evidence sources were established. This is presented in Table VI below.

It can be observed that evidence sources and the various types of possible evidence are repeated in some threats on SCI. Although the types of possible evidence may be similar for the various threats on SCI, the context and content within the evidence would be the critical difference for investigators. For example, the threat of spoofing in various processes and potential process crash or stop are different in nature, but the types of possible evidence to distinguish the threat are the same (e.g., network traffic and memory images).

V. DIGITAL FORENSICS IMPLICATIONS AND OPPORTUNITIES

The earlier sections of the paper defined SCI, identified SCI threats, offences committed and the corresponding evidence that could prove such offences were committed. However, it can be argued that SCI is still a work in progress as an international implementation of SCI has yet to be completed. Although SCI has yet to be implemented globally, it presents a chance for law enforcement agencies and investigators to quickly seize the opportunity to develop their digital forensics capabilities. We first discuss how DFI and LEA could benefit from our research and then discuss the possible future opportunities in digital forensics. These are divided into 2 major categories - research and legal opportunities.

A. Implications for Smart City Digital Forensics

We previously identified three issues that would hamper DFI and LEA in their investigation work on SCI attacks (namely definition, interoperability and cybercrime issues). Let us assume a theoretical scenario where a SCI has been fully implemented and consists of various technological implementations by multiple vendors. The successful implementation

of the SCI draws the attention of cybercriminals and sophisticated state-sponsored adversaries, who commence their cyber offensive operations on SCI (e.g. energy monitoring systems and the associated sensors of the nation's households). DFI and LEA are called upon to investigate the multitude of cyber-attacks. Still, they have to contend with current investigative workloads, understand the nuances of SCI, and various underlying technologies and finally obtain the proper evidence to determine the offences that had taken place. In this instance, DFI and LEA could expect to take a long time to investigate the attacks while also trying to resolve their current workload.

However, with our contributions, DFI and LEA can mitigate many challenges. Assuming that the SCI was designed using our threat model, the corresponding threats, cybercrime, and evidence required to identify the offences would have been documented. For example, based on Table V, DFI and LEA would know that there are twelve types of threats within SCI. The adversaries caused a process crash or stop and might have exploited vulnerabilities to change execution flows in the energy monitoring systems and gained elevated privileges. The offences committed would be illegal access, illegal interception, data interference, system interference and misuse of devices for process crash or stop, while the offences committed for gaining elevated privileges via changing execution flow would be illegal access, illegal interception, system interference, misuse of devices and computer-related forgery. After knowing the offences the adversaries committed, DFI and LEA can retrieve the possible evidence sources based on the information presented in Table VI. After consulting Table VI, the investigators know that the evidence sources are the SCI systems, SCI network infrastructure and the systems used by adversaries to commit the crimes. DFI and LEA are further guided to the types of evidence they could obtain, such as network traffic captures, device logs, server logs, system logs and malicious binaries. Moreover, since the various data sources of SCI were already determined and implemented by the chosen vendors, retrieving the evidence would be significantly less challenging for SCI and LEA as methods or capabilities to retrieve these types of evidence should have been supported or documented as these SCI are constructed. In contrast to the instance discussed in the preceding paragraph, DFI and LEA enjoy the benefit of being prepared (knowing potential threats, types of cybercrime related to SCI, and types of evidence that could be collected). They can commence investigations effectively, efficiently, and with confidence.

Figure 8 depicts the potential scenario faced by DFI and LEA and also shows how our proposed work can resolve their future challenges preemptively.

B. Research Opportunities for Digital Forensics

Based on our research performed in Section II to Section IV, we identified a few key technical research opportunities for SCI digital forensics. These are outlined in the following:

1) *Identification and Obtaining Dark Data:* Drawing a parallel connection with dark matter from Physics, Hand

TABLE VI
MAPPING THREATS ON SCI TO POSSIBLE EVIDENCE TYPES AND SOURCES

Threats on SCI (based on Table IV)	Types of Possible Evidence	Possible Evidence Sources
Spoofing in various processes	1. Network traffic (captured from SCI systems), device logs, server logs, system logs, malicious binaries 2. Network traffic (captured from SCI network infrastructure) 3. Memory images, hard disk images, devices used to commit crimes	
Potential process crash or stop		
Data flow potentially interrupted		
Elevation by changing execution flow in various systems		
Spoofing of various source data stores	1. Network traffic (captured from SCI systems), device logs, server logs, system logs, SCI data stores 2. Network traffic (captured from SCI network infrastructure) 3. Memory images, hard disk images, devices used to commit crimes	1. SCI systems 2. SCI network infrastructure 3. Adversarial systems
Potential lack of input validation	1. Network traffic (captured from SCI systems), device logs, server logs, system logs 2. Network traffic (captured from SCI network infrastructure)	
Potential weak access control for a resource	3. Memory images, hard disk images, devices used to commit crimes	
Elevation using impersonation	1. Network traffic (captured from SCI systems), device logs, server logs, system logs, malicious binaries, SCI Executable programs 2. Network traffic (captured from SCI network infrastructure)	
Systems may be subject to elevation of privilege using remote code execution	3. Memory images, hard disk images, devices used to commit crimes	
Potential data repudiation	1. Network traffic (captured from SCI systems), device logs, server logs, system logs 2. Network traffic (captured from SCI network infrastructure)	1. SCI systems 2. SCI network infrastructure
Data flow sniffing	1. Device logs, server logs, system logs, malicious binaries 2. Devices used to commit crimes, hard disk images, memory images	1. SCI systems 2. Adversarial systems
Cross site request forgery	1. Network traffic (captured from SCI systems), web server logs 2. Network traffic (captured from SCI network infrastructure) 3. Browser history, device artifacts (e.g. registry keys for Windows systems) 4. Memory images, hard disk images, devices used to commit crimes	1. SCI web servers 2. SCI network infrastructure 3. User devices 4. Adversarial systems

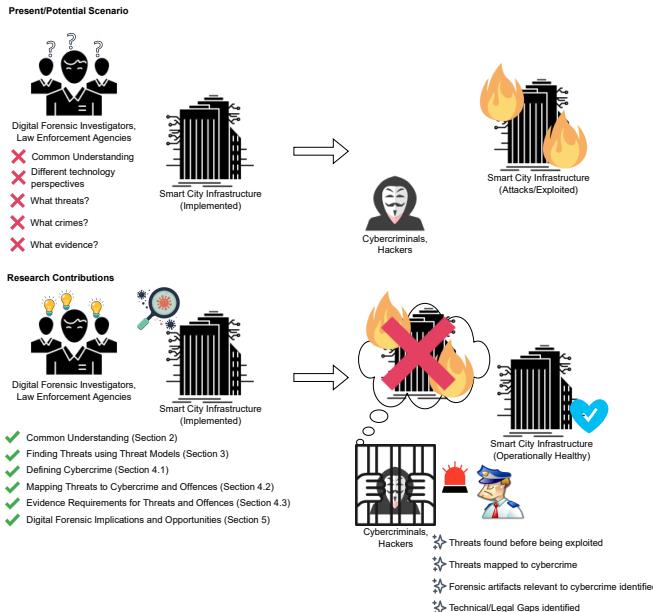


Fig. 8. Overview of our contribution, specific sections are highlighted

contends that missing data is dark data [44]. Hand further postulates that there are 15 types of dark data, and such data is vital, although they have not been noted since the dark data could have a huge impact on corresponding analysis and actions [44]. The current state of digital forensics on SCI exactly presents such a scenario. There is nothing wrong with governments wanting to implement SCI to improve the lives of citizens, but the corresponding identification of usable digital artifacts when cybercrime occurs has not been fully considered. Our research contributions (defining SCI, open-

sourcing our Smart City Threat Model and identification of SCI evidence in Sections II to IV) attempts to address Dark Data-Type (DD-Type) 1 and 2 highlighted by Hand [44]. DD-Type 1 refers to data we know are missing, while DD-Type 2 refers to data we do not know are missing [44].

The premise of our research was mainly driven by DD-Type 2 as we strove to identify the data we did not know that was missing in SCI digital forensics. After determining that there was a lack of a standard SCI definition, we formalize the SCI to facilitate understanding the threats, SCI cybercrime and subsequently, identifying SCI evidence sources. Consequently, certain DD-Type 2 data related to SCI became DD-Type 1, which were presented in our research. Given the nascent nature of SCI digital forensics, it is likely that other forms of dark data suggested by Hand exist in SCI deployments [44]. Additionally, since various governments may customize SCI depending on their use cases, the plethora of dark data could differ, which offers a wealth of technical research opportunities. Having said that, our proposed SCI threat model, identified threats, cybercrime and possible evidence provide a baseline for researchers, DFI and governments to refer to.

2) *Enhancing Evidence Capture in SCI:* Although we have suggested possible evidence sources and types in Table VI, evidence retrieval related to cybercrime will require human intervention. It will be a significant challenge to capture and identify the needed evidence at the right moment, especially when it is hard to predict when and where cybercrime will occur. A naive way would be to capture and store all evidence sources indefinitely, retrieving them when such evidence is required to investigate cybercrime. Unfortunately, such actions would not be economically feasible for governments, given the

associated economical cost required to store the evidence.

A possible solution to this dilemma could be using automated system processes to store evidence related to an ongoing cybercrime. However, this would require the engineering cooperation of the respective SCI vendors since this feature requirement appears to be an optional choice and not related to SCI core functionality. In the absence of vendor support, academic researchers could also possibly create tools or research on techniques that could be used in SCI to facilitate the investigation.

Another possible solution could be the usage of Machine Learning (ML) and Deep Learning (DL) techniques to process the various evidence sources and types to identify potential cybercrime that could have occurred. Such approaches have been explored in general cybersecurity and intrusion detection research [45], [46]. With the support of ML/DL techniques, the strain of sifting through tremendous amounts of potential evidence could be reduced greatly. Investigation support could also be provided, though at this point of time, human supervision would still be a necessary thing due to the probabilistic nature of ML/DL tools.

3) *SCI Evidence Storage and Management:* It is not unusual for DFI to work with large amounts of data. The challenge of managing and working with large volumes of data and evidence has been highlighted in traditional digital forensic investigations [47]. In the case of SCI, the potential amount of varied data gathered for forensic investigation could be tremendous, given the complexity and scale of SCI (with reference to Table VI). Researchers could adopt technical and management perspectives to tackle this challenge.

From a technical perspective, there are a few potential further research directions for forensic evidence storage. Firstly, some efforts could be directed towards a new medium of storage media or format where much more data could be stored resiliently and reliably. Alternatively, a suitable data compression mechanism could be explored to optimize and reduce evidence storage. Finally, the amount of evidence gathered within SCI could be reduced by applying data deduplication techniques. There has been open-sourced data deduplication frameworks such as DeFrame [48] and data deduplication extensions to open-sourced forensic tools such as The Sleuth Kit (TSK) proposed [49].

From a management perspective, researchers could examine various data policies related to SCI forensic evidence, such as storage, usage, access and acquisition vis-à-vis cybercrime investigation requirements. For example, it could be dictated under SCI evidence processing policies that data deduplication has to be carried out to reduce the potential amount of data that has to be examined and stored. Further policies such as access and storage would have to be determined ideally in partnership with LEA and governmental agencies.

C. Legal Opportunities for Digital Forensics

On top of technical opportunities, we also identified some legal opportunities that could facilitate SCI cybercrime investigation based on our work in Section II to Section-IV.

1) *Legal Jurisdiction of Digital Forensic Investigation:* The availability of technology and access to the internet facilitates cybercrime. As such, various SCI systems could potentially be accessed or probed by international users. Adversaries and cybercriminal groups could capitalize on such opportunities to commit crimes and sabotage SCI. Even if SCI systems had their access restricted to a user's geolocation, adversaries located elsewhere could breach into a local victim's computer to access SCI.

Since cybercriminals could be located in an international location, it may be challenging for local legal jurisdiction to prosecute cybercriminals located in another country in the absence of reciprocal legal arrangements. However, LEA should consider the possibility of exercising their rights to extraterritorial jurisdiction. Gillespie suggests that there are four possible grounds in which LEA could apply extraterritorial jurisdiction [39]:

- 1) **Universal jurisdiction:** Reserved for the most grievous offences that was committed. The concept of this jurisdiction was that any nation could charge perpetrators regardless of their nationality and location where the crime was committed [50]. For example, perpetrators who interfere with the usual operations of SCI which leads to massive loss of human lives could be potentially considered under this category.
- 2) **Protective principle:** The protective principle proposes to allow countries seek extraterritorial jurisdiction to defend the victim country's state interests and security [50]. For example, perpetrators who attempt cyberattacks on governmental or military assets owned by international governments could fall under this category.
- 3) **Active personality principle:** This form of extraterritorial jurisdiction is based on the the perpetrator's nationality and/or place of residence, and legal action could commence from there.
- 4) **Passive personality principle:** Similar to the active personality principle, this extraterritorial jurisdiction differs in that legal action commences instead from the victim's nationality and/or place of residence.

2) *International Cooperation in Digital Forensics:* We discussed the possibility of extraterritorial jurisdiction in Section V-C1. However, it would not be possible without a global consensus. The International Criminal Court (ICC) could be a potential platform where serious cybercrimes are judged, but it was observed that the ICC usually handles crimes against humanity [39].

A potential avenue could be the Budapest Convention on Cybercrime where the Second Additional Protocol to the Convention was recently put up for signing with the various parties and signatories [51]. Although the current scope of cooperation was limited to domain name and subscriber information, there was also consideration for extraterritorial cooperation via joint investigations and disclosure of electronic evidence [51]. However, the effectiveness of the Budapest Convention is only as effective as the parties and signatories

who sign the Convention.

VI. LIMITATIONS

Our proposed SCI threat model and the corresponding SCI threats, cybercrime and potential evidence are dependent on Definition 1, the STRIDE threat model and the four ISO standards mentioned in Section II-B. There is a possibility that governments vested in alternative types of SCI (such as proprietary SCI not based on Definition 1) would find a lack of alignment in our work. Nonetheless, the foundation of Definition 1 is based on the goals of an intergovernmental organization (17 United Nations SDGs) and several standards from an international standards body (ISO). Therefore, Definition 1 is technology agnostic and broad. While government agencies are free to select an SCI that best fits their needs, we believe that Definition 1 can assist governments in establishing a solid foundation in cybercrime prevention.

We used Microsoft TMT to create our threat model. As such, there is a dependency on a third-party tool. However, we believe the advantages outweigh the disadvantages of relying on a third-party tool. Threat modeling could be a herculean effort for complex systems such as SCI. Compared to manually creating threat models or using visualization tools with limited automation while creating threat models, the Microsoft TMT could track the list of identified threats, generate a summary report of threats and even provide suggested ways to mitigate identified threats.

The possible evidence sources and types in Table VI are dependent on the threat model. In our research, we have limited the depth of our threat model at Level 1 as it was intended to serve as a baseline template for interested adopters. If we went down to more granular levels (e.g. Level 2 or 3) as specified by Microsoft, we would risk being overly specific and may outline redundant threats and evidence. Hence, although Table VI comprehensively maps threats on SCI to possible evidence sources and types, our threat model does not entirely outline all potential threats faced by SCI. Nonetheless, our baseline threat model provides a platform for extending towards different SCI customizations made by a specific organization or vendor.

VII. RELATED WORK

In our preparatory work for the paper, we performed a literature review of state-of-the-art works with the keywords “Smart City”, “Smart Cities”, “Smart City Digital Forensics”, “Cybercrime”, “Threat Modeling” and “Threat Models”. The keywords were used in our searches via Google Scholar, journal databases such as ScienceDirect, IEEE and Association of Computing Machinery (ACM), and search engines such as Google. We manually checked the literature for the relevancy of the topics discussed in this paper and excluded papers unrelated to our research goals.

The idea of cybercrime had been widely discussed on several platforms but was first formally defined and classified by Gordon and Ford [52]. Gordon and Ford described cybercrime as any crime that was aided or executed via electronic devices

such as computers or devices [52]. They also further classified cybercrime into Type I and II crimes - the former is often supported by malicious software such as keyloggers, viruses and vulnerabilities, while the latter uses day-to-day software and protocols. Type I cybercrime also usually happens as a specific event from the victim’s perspective, while Type II cybercrime could repeatedly happen from the user’s point of view. Gordon and Ford further argues that pure Type I or II cases are rarely observed and that cybercrime is usually broad [52]. Our work, particularly Table V, aims to resolve this conundrum by first identifying and mapping the possible threats to related offences. This prevents any confusion faced by DFI and LEA, albeit targeted specifically towards SCI-related cybercrime.

There have also been calls for digital forensics to evolve. MacDermott et al. [53] highlighted that new ways to investigate IoT-related crime have become a growing necessity - threats and criminal abuses in smart cities are increasing while securing the infrastructure was a challenge. Moreover, MacDermott et al. [53] further stated that new threats would have to mitigate by a new generation of digital forensics and best practices to ascertain and identify evidence in an evolving regulatory landscape concurrently. Our research precisely attempts to address this problem by combining digital forensics and industry best practices (e.g. threat modeling systems before implementation) to identify threats preemptively and to highlight cybercrime and needed digital evidence for investigation.

Baig et al. [1] proposed a study of digital forensics and cybersecurity in smart cities based on NIST’s smart city model. However, their study was only based on three categories (smart environments, smart living, smart mobility) out of the six categories mentioned in the NIST smart city model (smart economy, smart governance and smart people were not discussed). As such, this may lead to the misconception by researchers, DFI and LEA that smart cities only comprise smart environments, smart living and smart mobility. Meanwhile, our paper fully follows all clauses and elements of data indicators listed in the ISO standards we referenced. In addition, our paper attempts to define smart cities at an even lower layer of abstraction while embracing conformity to international standards (*i.e.*, Definition 1). Finally, governments will likely implement SCI infrastructure in phases. Our work can factor in technology evolution and facilitates requirement changes by governments (if necessary), yet provides DFI and LEA with the ability to prepare for cybercrime on existing SCI infrastructure.

A digital forensic model to identify interconnectivity between things was proposed by Kim et al. [54], where the authors proposed an improved framework to identify evidence in interconnected devices. While the proposed model is informative, it lacks the context of SCI and assumes that DFI and LEA have an in-depth knowledge of SCI. It also requires users to follow the process iteratively. On the other hand, our threat model lists all possible threats within SCI that are grouped by data indicators via threat modeling immediately. The possible

offences, forensic evidence types and sources based on the threats are then linked together and presented in our research. By using our threat modeling template, governments, DFI, and LEA could obtain personalized guidance and direction in identifying relevant threats and sources of forensic data. If certain data indicators were not required, they could be trimmed away from the template.

Kavallieratos et al. investigated threats in connected smart homes using STRIDE and Microsoft TMT [55]. They acknowledged the dynamic nature of smart, interconnected devices but only outlined six smart home scenarios and solely determined threats. Moreover, Kavallieratos et al. did not adhere to Microsoft's best practices for threat modeling and failed to model their threat models according to appropriate context levels [37]. As such, it is possible that Kavallieratos et al. did not identify all existing threats in their scenarios.

Countries have attempted to specify technical references and standards for smart cities/nations to enrich their citizens' lives. For example, in Technical Reference (TR) 47, Singapore attempted to specify a technology-agnostic reference architecture for IoT and sensor networks [56]. The motivation of the TR was to create an anticipatory government that could offer integrated city services while allowing citizens to enhance their lives with data [56]. Other than technical requirements specified by Singapore, it was observed that TR 47 also mentioned the need for incident response and investigations in the architecture [56]. Our work could provide the investigative context and guidance required from a data indicator perspective.

In summary and to the best of our knowledge, we provide the first systematic SCI threat model that is technology-agnostic and adopts data indicators that have been internationally agreed upon (i.e., ISO standards). Instead of focusing on the technological aspect of SCI, we based our SCI on indicators that reflected the quality of life and used the corresponding data indicators in our SCI.

VIII. DISCUSSION AND CONCLUSION

Determining SCI threats, cybercrime, and retrieving related forensic evidence is not a trivial task for DFI and LEA. Even at a more fundamental level, the definition of SCI could differ due to different needs and perspectives. Our research work led us to the following key observations:

1) SCI Definition. A common definition of SCI supported by international standards helped to reduce ambiguity when we had to prepare for our threat modeling endeavors. We could be technology-agnostic in our threat model by viewing SCI as an outcome-based and data-driven construct compared to a purely technological and role-based construct. With a standard definition and understanding of SCI, DFI and LEA would be better prepared to handle SCI investigations.

2) SCI Threats, Cybercrime and Related Evidence. As Horsman has succinctly pointed out, DFI and LEA have already had their hands full with ongoing cybercrime cases and are hard-pressed to perform further research in

emerging technologies [3]. We believe that adversaries and cybercriminals are very likely to target SCI due to their novelty and the data that flows through these systems. Before SCI is fully implemented and deployed, we preemptively highlight vital details that DFI and LEA are likely to use in their course of investigation and workflows. Based on a standard definition of SCI, we derived potential threats and mapped them to possible offences. We then further mapped the threats to possible evidence types and sources. With these details, DFI and LEA will have some guidance when they are required to conduct investigations related to SCI.

- 3) **SCI Technical and Legal Opportunities for Digital Forensics.** Although our research attempts to get DFI and LEA up to speed about SCI, there are still multiple opportunities for the digital forensics community to explore. We hope that the digital forensics community can work together in looking into the opportunities identified - a two-pronged approach from both a technical and legal perspective would better assist DFI and LEA, who would have to contend with SCI investigations in the future.

We have provided a threat model template freely available for usage by interested parties to model their SCI and discover potential threats. Following that, we also listed possible digital evidence that DFI and LEA could use to investigate cybercrime based on these threats. We believe that our work in this paper will serve as a critical foundation in aiding DFI and LEA in investigating SCI cybercrime, while also factoring in the goal of improving citizens' lives by aligning our threat model with the UN sustainability issues and goals as specified in ISO37101:2016.

Availability: Our threat models are publicly available at <https://github.com/poppopretn/SmartCityThreatModel> for investigation and further research.

IX. ACKNOWLEDGEMENTS

This work is partially supported by the NRF National Satellite of Excellence in Trustworthy Software Systems (Project no. NSOE-TSS2021-01 and NSOETSS2020-03).

REFERENCES

- [1] Zubair A. Baig, Patryk Szewczyk, Craig Valli, Priya Rabadia, Peter Hannay, Maxim Chernyshov, Mike Johnstone, Paresh Kerai, Ahmed Ibrahim, Krishnun Sansurooah, Naeem Syed, and Matthew Peacock. Future challenges for smart cities: Cyber-security and digital forensics. *Digital investigation*, 22:3–13, 2017.
- [2] Luca Caviglione, Steffen Wendzel, and Wojciech Mazurczyk. The future of digital forensics: Challenges and the road ahead. *IEEE Security & Privacy*, 15(6):12–17, 2017.
- [3] Graeme Horsman. Raiders of the lost artefacts: Championing the need for digital forensics research. *Forensic Science International: Reports*, 1:100003, 2019.
- [4] Fadi Al-Turjman. *Intelligence in IoT-enabled Smart Cities*. Routledge, 1st edition, 2019.
- [5] Rudolf Giffinger, Christian Fertner, Hans Kramar, Robert Kalasek, Nataša Miljanović, and Evert Meijers. *Smart cities - Ranking of European medium-sized cities*. 01 2007.
- [6] Fadi Al-Turjman. Mobile couriers' selection for the smart-grid in smart-cities' pervasive sensing. *Future Generation Computer Systems*, 82:327 – 341, 2018.

- [7] Fadi Al-Turjman and Sinem Alturjman. 5g/Iot-enabled Uavs for multi-media delivery in industry-oriented applications. *Multimedia tools and applications*, 79(13-14):8627–8648, 2020.
- [8] Bosch. *Smart Cities | Bosch Global*, 2020. (Accessed 29 July 2020).
- [9] Cisco. *Cisco Kinetic for Cities - Cisco*, 2020. (Accessed 29 July 2020).
- [10] Microsoft. *Microsoft CityNext*, 2020. (Accessed 29 July 2020).
- [11] Schneider. *EcoStruxure: IoT – Internet of Things | Schneider Electric Global*, 2020. (Accessed 29 July 2020).
- [12] ABI Research. *Smart Cities Platforms and Standards*, 2018. (Accessed 17 April 2020).
- [13] FIWARE. *The Open Source platform for our smart digital future - FIWARE*, 2020. (Accessed 29 July 2020).
- [14] TALQ Consortium. *Home - TALQ - Consortium*, 2020. (Accessed 29 July 2020).
- [15] Graeme Horsman and Nina Sunde. Part 1: The need for peer review in digital forensics. *Forensic Science International: Digital Investigation*, 35:301062, 2020.
- [16] Nadine M. Smit, Ruth M. Morgan, and David A. Lagnado. A systematic analysis of misleading evidence in unsafe rulings in england and wales. *Science & Justice*, 58(2):128 – 137, 2018.
- [17] ISO. Sustainable development in communities — Management system for sustainable development — Requirements with guidance for use. ISO 37101:2016, International Organization for Standardization, Geneva, Switzerland, 2016.
- [18] ISO. Sustainable cities and communities — Indicators for city services and quality of life. ISO 37120:2018, International Organization for Standardization, Geneva, Switzerland, 2018.
- [19] ISO. Internet of Things (IoT) - Reference Architecture. ISO 37122:2019, International Organization for Standardization, Geneva, Switzerland, 2019.
- [20] ISO. Internet of Things (IoT) - Reference Architecture. ISO 37123:2019, International Organization for Standardization, Geneva, Switzerland, 2019.
- [21] United Nations. The 17 GOALS | Sustainable Development, 2021. (Accessed 10 January 2021).
- [22] Adam Shostack. *Threat Modeling: Designing for Security*. Wiley Publishing, 1st edition, 2014.
- [23] Nataliya Shevchenko, Timothy A. Chick, Paige O'Riordan, Thomas Patrick Scanlon, and Carol Woody. Threat modeling: A summary of available methods. *Software Engineering Institute | Carnegie Mellon University*, 2018.
- [24] Adam Shostack. *The Threats to Our Products - Microsoft Security Blog*, 2009. (Accessed 29 July 2021).
- [25] Microsoft. *Microsoft Threat Modeling Tool release notes - Azure*, 2020. (Accessed 30 July 2021).
- [26] Tony UcedaVélez and Marco M. Morana. *Risk Centric Threat Modeling: Process for Attack Simulation and Threat Analysis*. Wiley Publishing, 1st edition, 2015. ISBN: 9781118988374.
- [27] Bruce Schneier. *Academic: Attack Trees - Schneier on Security*, 1999. (Accessed 2 August 2021).
- [28] R. Vigo, F. Nielson, and H. Nielson. Automated generation of attack trees. In *2014 IEEE 27th Computer Security Foundations Symposium (CSF)*, pages 337–350, Los Alamitos, CA, USA, jul 2014. IEEE Computer Society.
- [29] Christopher Alberts, Audree Dorofee, James Stevens, and Carol Woody. Introduction to the octave approach. Technical report, Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA, 2003.
- [30] ThreatModeler. *Threat Modeling Methodologies: What is VAST? | ThreatModeler*, 2018. (Accessed 5 August 2021).
- [31] Forum of Incident Response and Security Teams. *CVSS v3.1 Specification Document*, 2019. (Accessed 5 August 2021).
- [32] Paul Saitta, Brenda Larcom, and Michael Eddington. *Trike v.1 Methodology Document*, 2006. (Accessed 5 August 2021).
- [33] Jane Cleland-Huang. How well do you know your personae non gratae? *IEEE Software*, 31(4):28–31, 2014.
- [34] Bradley Pottenger, Goncalo Martins, and Xenofon Koutsoukos. Software and attack centric integrated threat modeling for quantitative risk assessment. In *Proceedings of the Symposium and Bootcamp on the Science of Security*, HotSos '16, page 99–108, New York, NY, USA, 2016. Association for Computing Machinery. ISBN: 9781450342773.
- [35] Nancy Mead, Forrest Shull, Krishnamurthy Vemuru, and Ole Villadsen. A hybrid threat modeling method. Technical Report CMU/SEI-2018-TN-002, Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA, 2018.
- [36] IriusRisk. *Threat Modeling Platform*, 2021. (Accessed 5 December 2021).
- [37] Microsoft Docs. *Data-Flow Diagram Depth Layers - Learn | Microsoft Docs*, 2022. (Accessed 5 January 2022).
- [38] Microsoft Docs. *Layer 0 - The system layer - Learn | Microsoft Docs*, 2022. (Accessed 5 January 2022).
- [39] Alasdair Gillespie. *Cybercrime: Key Issues and Debates*. Routledge, 2nd edition, 2019. ISBN: 9781351010269.
- [40] Council of Europe. *CETS 185 - Convention on Cybercrime*, 2001. (Accessed 5 April 2022).
- [41] Council of Europe. *Ecuador invited to join the Budapest Convention on Cybercrime - News*, 2022. (Accessed 5 April 2022).
- [42] United Nations. *Growth in United Nations membership | United Nations*, 2022. (Accessed 5 April 2022).
- [43] Eoghan Casey. *Digital Evidence and Computer Crime: Forensic Science, Computers, and the Internet*. Academic Press, Inc., Orlando, FL, USA, 3rd edition, 2011.
- [44] David J. Hand. *Dark Data: Why What You Don't Know Matters*. Princeton University Press, 2020.
- [45] Yang Xin, Lingshuang Kong, Zhi Liu, Yuling Chen, Yanmiao Li, Hongliang Zhu, Mingcheng Gao, Haixia Hou, and Chunhua Wang. Machine learning and deep learning methods for cybersecurity. *IEEE Access*, 6:35365–35381, 2018.
- [46] Dilara Gümuşbaş, Tulay Yıldırım, Angelo Genovese, and Fabio Scotti. A comprehensive survey of databases and deep learning methods for cybersecurity and intrusion detection systems. *IEEE Systems Journal*, 15(2):1717–1731, 2021.
- [47] Darren Quick and Kim-Kwang Raymond Choo. Introduction. In *Big Digital Forensic Data: Volume 1: Data Reduction Framework and Selective Imaging*, pages 1–4, Singapore, 2018. Springer Singapore.
- [48] Dan Feng. The framework of data deduplication. In *Data Deduplication for High Performance Storage System*, pages 135–154, Singapore, 2022. Springer Nature Singapore.
- [49] Izabela Savić and Xiaodong Lin. The analysis and implication of data deduplication in digital forensics. In Weizhi Meng and Mauro Conti, editors, *Cyberspace Safety and Security*, pages 198–215, Cham, 2022. Springer International Publishing.
- [50] Anthony Aust. *Handbook of International Law*. Cambridge University Press, 2 edition, 2010.
- [51] Council of Europe. *Opening for signature of the Second additional Protocol to the Cybercrime Convention*, 2022. (Accessed 15 May 2022).
- [52] Sarah Gordon and Richard Ford. On the definition and classification of cybercrime. *Journal in computer virology*, 2(1):13–20, 2006.
- [53] Aine MacDermott, Thar Baker, and Qi Shi. Iot forensics: Challenges for the iot era. In *2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*, pages 1–5, 2018.
- [54] Jieon Kim, Jungheum Park, and Sangjin Lee. An improved iot forensic model to identify interconnectivity between things. *Forensic Science International: Digital Investigation*, 44:301499, 2023.
- [55] Georgios Kavallieratos, Vasileios Gkioulos, and Sokratis K. Katsikas. Threat analysis in dynamic environments: The case of the smart home. In *2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pages 234–240, 2019.
- [56] Enterprise Singapore. *IoT reference architecture for Smart Nation*. Technical reference ; TR 47 : 2016. SPRING Singapore, Singapore, 2016.