Future challenges for smart cities: Cyber-security and digital forensics

Zubair A. Baig*, Patryk Szewczyk, Craig Valli, Priya Rabadia, Peter Hannay, Maxim Chernyshev, Mike Johnstone, Paresh Kerai, Ahmed Ibrahim, Krishnun Sansurooah, Naeem Syed, Matthew Peacock

Security Research Institute & School of Science, Edith Cowan University, Perth 6027, Australia

ARTICLE INFO

Article history:
Received 16 February 2017
Received in revised form
21 June 2017
Accepted 27 June 2017
Available online 16 August 2017

ABSTRACT

Smart cities are comprised of diverse and interconnected components constantly exchanging data and facilitating improved living for a nation’s population. Our view of a typical smart city consists of four key components, namely, Smart Grids, Building Automation Systems (BAS), Unmanned Aerial Vehicles (UAVs), Smart Vehicles; with enabling Internet of Things (IoT) sensors and the Cloud platform. The adversarial threats and criminal misuses in a smart city are increasingly heterogenous and significant, with provisioning of resilient and end-to-end security being a daunting task. When a cyber incident involving critical components of the smart city infrastructure occurs, appropriate measures can be taken to identify and enumerate concrete evidence to facilitate the forensic investigation process. Forensic preparedness and lessons learned from past forensic analysis can help protect the smart city against future incidents. This paper presents a holistic view of the security landscape of a smart city, identifying security threats and providing deep insight into digital investigation in the context of the smart city.

© 2017 Elsevier Ltd. All rights reserved.

Introduction

More than 50% of the world’s population today reside in urban areas and this percentage is expected to increase because of population migration to these regions in the quest for better jobs and education (Khatoun and Zeadally, 2016). The concept of the smart city represents the first major impetus for change in metropolis-sized urban planning since Victor Gruen re-envisioned the urban landscape in America in the 1950s. As a consequence, smart cities have recently gained attention; comprising a collection of entities deployed and maintained in a city to facilitate convenient and improved living for a nation’s population. Various initiatives worldwide have facilitated the emergence of smart cities that address the needs of businesses, institutions, and citizens, through targeted and efficient delivery of service. The smart city promise of provisioning a connected environment for all its citizens is realized through intelligent and sustainable enabling technologies and platforms including the Internet of Things (IoT) and the Cloud.

Smart city services can extend into many diverse domains including the environment, transportation, health, tourism, home energy management and safety and security (Byun et al., 2014; Kantarci and Mouftah, 2014; Lopes et al., 2015). The U.S. National Institute of Standards and Technology (NIST) smart city model is one of the most widely adopted reference models (Khatoun and Zeadally, 2016). It comprises six categories, namely, smart environment, smart mobility, smart economy, smart governance, smart people and smart living; with IoT as the enabling technology. We base our study on four components of the above categories:

- Smart Grids (Smart Environments)
- Building Automation Systems (Smart Living)
- Unmanned Aerial Vehicles (Smart Mobility)
- Smart Vehicles (Smart Mobility)

The smart city will include several types of IoT sensors including those required for smart parking, structural health awareness, urban noise mapping in real-time, traffic level monitoring and route optimization and smart street lighting. The enabling technology for the above smart city components is the IoT whilst the enabling platform for centralized data storage and rendering is the Cloud.
Smart cities are exposed to a diverse set of cyber security threats and criminal misuses. In this environment, a single smart city vulnerability, when exploited by an individual or organized group, may put the entire city at risk (Khatoun and Zeadally, 2016). This complex environment also presents a significant challenge for digital forensic investigations, which will invariably rely upon the data generated by the smart city components. To envision a secure smart city cyber security platform with access to reliable forensic evidence, due diligence for data transfer and storage in the Cloud is mandatory. Such forensic preparedness can provide help to develop more effective ways to detect and prevent problems before they cause widespread harm (Sachowski, 2016; Casey, 2009).

In addition, if a cyber-attack transpires against critical components of a connected smart city ICT infrastructure, as illustrated in Fig. 1, a standard scientifically proven method must be applied for acquisition and subsequent analysis of the data, as part of the forensic investigation.

In this paper, we present a comprehensive analysis of the vulnerabilities and the associated threat landscape for each of the four identified components of a smart city, namely, Smart grids, Building Automation Systems (BAS), Unmanned Aerial Vehicles (UAVs), Smart Vehicles; with enabling IoT sensor technology and the Cloud. Following this, we present a detailed analysis of challenges associated with forensic investigations of smart city data.

**Smart city entities**

**Smart grids**

Smart grid technology is changing the way traditional power grids operate (Fig. 2) by reducing energy demands, global warming and consequently, utility costs. Consumers are required to share information about their energy consumption with their utility providers, over communication channels using smart meters. The interconnection of multiple smart meters and computerized infrastructure of the grid makes them vulnerable to several network based attacks (McDaniel and McLaughlin, 2009).

Data from smart grid devices can be essential for studying energy consumption patterns and supply/demand management. Traditional data management applications are not designed to handle large scale data generated by the grid. Cloud computing is an appropriate choice that can be leveraged to store and process such large volumes of data (Bera et al., 2015). Data can also be used for detecting anomalous behaviour in smart grids and can assist in forensic investigations. Anomaly detection techniques applied to data from different IoT components operating in a smart grid can detect compromised devices and protect smart grid operations.

Smart grid threats can be categorized into those that affect: network availability, data integrity and information privacy. Devices such as smart meters and IoT devices within a consumer’s household are located in physically insecure locations and can be exploited by an adversary. Since the grid maintains a two-way communication channel with multiple intelligent smart grid devices and the Cloud, these exposed devices create numerous entry points for an adversary to penetrate the smart grid, and also expose smart grid data stored in the Cloud to various security threats.

Consumption patterns could also be utilized by an adversary to extract household information such as the number of individuals living in a house, and the various types of appliances in use (Jokar et al., 2016). Another challenge to privacy of smart grid data is the ownership and accessibility of consumer data stored in the Cloud. Jokar et al. (2016) suggest using anonymization of the data to haze out attribution of any traits to a particular customer.

Smart grids are also vulnerable to attacks that can affect the timely delivery of messages between interconnected systems, which is critical to the successful operations of the grid. Lu et al. (2010) categorize an attack targeting the time constraints in grid communication as a Denial of Service (DoS) attack, exploiting...
vulnerabilities found in the protocol stacks of IPv4, IPv6, 6lowPan, and TCP/IP, commonly used by smart grid components. The smart grid is attacked through generation of legitimate but useless traffic thereby delaying the delivery of legitimate messages and also through launching jamming attacks in wireless power networks. An example of the danger of alert interruption is the North East blackout in 2003, wherein alert messages did not get through to operators.

Electricity theft in a smart grid environment can require digital forensic investigation. Smart meters can be used to analyze the energy consumption patterns in households to detect anomalous behaviour in smart grids. Smart meter usage logs can even help forensic investigators detect in-house marijuana growing operations which has been linked to electricity theft (Depuru et al., 2011). Proper forensic preparation and treatment of alerts and logs in a smart grid can help prevent major problems and can be used in digital investigations.

Building automation systems (BAS) security

Buildings are a core feature of the future smart city. As Lilis et al. (2015) note, there cannot be a smart city without a smart building. Intelligence in buildings is achieved through the use of Building Automation Systems (BAS). BAS centralize the monitoring and control of multiple building services over a shared network medium. Typical building services include heating ventilation and air conditioning (HVAC), elevators, access control, closed-circuit television (CCTV), lighting, water and energy systems. BAS devices, such as sensors and actuators, report and provide physical control through controller devices. By connecting these vastly different building services together, an entire building (in some cases multiple building operations) can be managed automatically, and observed remotely over the Internet.

There are inherent cyber security implications from the interconnection of BAS services (Kastner et al., 2005; Peacock and Johnstone, 2014). The culmination of high trust devices, isolation intended protocol design, extensive lifecycle, and external interconnection has led to a range of security vulnerabilities which can expose a BAS to a range of threats, including physical damage, denial of service, explicit trust in sensors and controllers and associated second order effects.

Due to the cyber-physical properties of BAS, and the limited processing capability of devices, physical damage against building components are possible through normal operation of the protocols (Holmberg, 2003; Johnstone et al., 2015). With the increased connectivity of BAS to shared networks, BAS are exposed to the same threats faced by traditional IT based networks and protocols. Denial of Service against particular building services, for example the access control system, or complete building control takeover is possible (Antonini et al., 2014; Mundt and Wickboldt, 2016) due to the resource constraints of BAS devices.

As a result of their initial internal-only designs, BAS protocols are inherently insecure due to the amount of trust they give to sensors and controllers. Given a controlled, isolated network, it was previously safe to trust messages and devices connected to the network with limited requirements for checking integrity. The consequence of this design is that source authentication is generally non-existent in BAS protocols, as devices are reliant on being truthful about what they do, with limited ability of verification (Holmberg, 2003; Granzer et al., 2010). Second order effects also exist, including damage to physical goods/data inside a building, such as perishables or data servers through temperature manipulation.

In addition, with increased connectivity, BAS can be used as a pivot point into the Cloud to carry out traditional ICT style attacks, such as data theft and exfiltration. Again, in this context, forensic preparation and handling of BAS can help prevent major problems and will facilitate digital investigations.

Unmanned aerial vehicles (UAVs) security

Widespread civilian and commercial Unmanned Aerial Vehicles (UAVs) or drones are a recent addition to the smart city landscape. With the advancement of microprocessors and manufacturing
techniques, small, wireless, plug and play drones are now available to anyone from US$300. These drones host a range of features, also typically containing an onboard camera. Additional components can be added to develop a customizable flying sensor platform; with civilian uses including 3D interactive games, aerial photography and temperature sensing. Commercial applications such as package delivery, coastline patrol and agricultural insecticide distribution have progressed rapidly.

It is expected that future use of drones will impact smart cities dramatically, allowing a platform to provide services and gather data as sensors for an interconnected city. With the applications near limitless, drones do not seem to be going away, with aviation legislation now being defined to encourage the growth of using drones as a service platform, with varying degrees of licensing and regulation.

Currently, civilian-grade drones have varying levels of security features, dependent on their cost point. One study (Peacock, 2014) showed the lack of security in a commonly used civilian drone, the Parrot AR Drone 2. The drone uses an unsecured WiFi connection between a smart device and the onboard Linux-based system to control flight. The onboard system runs with a privileged user account (root), with openly accessible Telnet and FTP services available for interaction. While the study was taken from a defensive perspective (to prevent unauthorized drones entering an area), the identified vulnerabilities have a low barrier to entry, and could be used for malicious purposes.

Additional studies by Pleban, Band and Creutzburg (Pleban et al., 2014) highlight the vulnerabilities present in the AR drone 2, and discuss securing the communications connection with encryption. At a high level, the cyber threats faced by drones can be categorized into two groups, communication threats and device level threats.

From the above studies, it is revealed that the civilian drones investigated do not employ cryptographic techniques to secure communications between controllers and drones. Methods employed to restrict communications to single controllers are also easily defeated. As such, civilian drones can be susceptible to remote hijacking, connection denial, video interception and total control takeover by adversaries (Peacock, 2014).

The specificity of drone hardware varies based on cost point, however a common theme from the aforementioned studies is the trusting nature of the devices. The civilian drones surveyed all operate at root or equivalent permissions on the system by default, allowing access to all files and services on the platform. Given the trust placed in the controller connecting to the drone and the lack of encrypted communications, when an adversary connects to a drone they have full control of the device, rather than requiring to commit further resources to escalate privilege.

A drone controller has the potential to also be vulnerable, dependent on the type of controller, and the software used for control. As the controllers can be mobile devices, or operated through cloud platforms, threats against these controller classes can impact the operability of the drone. Digital data generated by UAVs can be used as a source of evidence in digital investigations, and can result in privacy violations. Therefore, it is important for the smart city to have processes in place to handle UAVs as a source of digital evidence.

**Smart vehicles**

Traditional vehicle networks utilized vehicle-based networking technologies such as bus networks, for Electronic Control Unit (ECU) operation. With the increase in in-car entertainment and GPS, consumer-based devices have been connected to traditional vehicle networks using consumer network protocols. For fully integrated smart vehicles, higher bandwidth network protocols, such as Ethernet are now being used (Lin and Sangiovanni-Vincentelli, 2012).

All modern vehicles use a Unified Diagnostic Services (UDS) protocol for vehicle controller diagnostics, which can be connected to, retrieved, and interacted with through an on-board diagnostics (OBD) port. OBD adapters that typically communicate using Bluetooth facilitate smart phone communication with the underlying vehicle network to retrieve data and also to interact with the network. (Ravenscraft, 2014). Further, devices such as Event data recorders (EDR) record important event data based on predefined circumstances, such as a drastic speed reduction which can correlate with accident events (Canis and Peterman, 2014).

Smart vehicles, like modern vehicles also contain telematic systems such as GPS, and integrated infotainment systems which can link to smart devices and the Cloud, providing a wealth of forensic data. In addition, forensic-capable devices exist for non-invasive extraction of diagnostic data from vehicles (Mansor et al., 2016). Future smart vehicles will offer similar data logging functionality, as many diagnostic services are legally required to be hosted within smart vehicles in countries such as the United States (Canis and Peterman, 2014).

Intelligent vehicles regularly report their status and coordinates to base stations placed on the perimeter of a road network. Vehicle location information has been useful in digital investigations including burglary and homicide. The data is transmitted either periodically or instantaneously, depending on the governing policy and other constraints including device power and communication channel availability (Samie et al., 2016). In addition, contemporary smart vehicles (Baldwin, 2016) are programmed to authenticate to a city’s Cloud infrastructure upon first entry into the smart city zone. As a vehicle is driven around, its GPS coordinates and direction of movement determine the next traffic light that it will encounter. The vehicle’s dashboard is designed to display a countdown clock that operates until the green light appears. Through such seamless integration of smart vehicles, traffic light sensors and the city’s Cloud platform, citizens of the city benefit from real-time information of use whilst driving a smart vehicle.

According to a survey by re-insurer Munich Re, 55% of corporate risk managers surveyed named cyber security as the top concern for self-driving smart cars, while safety, the major focus of the automotive industry was only 6%. (Webb, 2016). Smart vehicles using ethernet protocols require network segregation, which is a different model to the traditional bus topology of vehicle networks. Given vehicle networks were previously isolated local bus networks, secure measures are now required to correctly integrate smart vehicles with the Cloud, with the potential for more traditional IT security technologies, such as firewalls to be deployed in the smart vehicle network (Valasek and Miller, 2014).

The security challenges faced by smart vehicles are typical of the previously isolated networks; albeit now being externally connected. However, unlike some other control networks, smart vehicle network protocol design is taking a vendor agnostic approach, which allows for standardized integration of security features. Security goals of a smart vehicle network comprise: protection of the cyber-physical platform, prevention of physical damage, prevention of remote un-authorized operation and prevention of data theft. A number of threats faced by future smart vehicles exist, which can be classed as physical, interception, abuse and loss of information threats (Lin and Sangiovanni-Vincentelli, 2012; Valasek and Miller, 2014; Anon, 2016; Anon, 2017; Greenberg, 2016).

1. **Physical threats:** can include fault-injection into the ECU to defeat central locking systems, side channel attacks to leak information, or introducing data glitches to gain unauthorized access to debug interfaces.
2. Interception threats: such as man in the middle, reconnaissance, and replay attacks can exist against the data transmitted over the networks internally between ECUs, and between other vehicles and the Cloud.
3. Abuse threats: can include traditional ICT attacks, such as Denial of Service, malicious code execution and unauthorized access to the vehicle, as well as remote execution and operation of the vehicle.
4. Malicious code: Given the increase of integrated infotainment systems which often run embedded versions of Linux, Windows and Android, generic malicious code could be executed against infotainment systems to comprise all connected devices of the smart vehicle network and with potential to leak into the Cloud.
5. Data threats: against the information contained on smart vehicle networks exist, with loss of information from a connected Cloud, and private information leakage if the vehicle is resold.

The broad range of risks and the large amounts of data associated with smart vehicles makes this a challenging area for forensic preparedness and digital investigation. Much work is needed to develop methods and systems for handling incidents involving smart vehicles while preserving personal privacy of individual drivers.

**IoT sensors**

The Internet of Things (IoT) refers to connecting smart devices such as sensors and intelligent vehicles to networks such as the Internet (Nakamura et al., 2016; Razzaque et al., 2016). These devices have become an integral part of a smart city. While smart cities have become an attractive environment for IoT applications, these services must be realized in a scalable and secure manner to support future economic growth and address the existing challenges associated with heterogeneous IoT devices (Jin et al., 2014). These challenges include lack of investment and high cost, high energy consumption, and cyber security.

IoT sensors are deployed and maintained in their respective environments in order to monitor various phenomena and to respond to changes in the smart city environment. These responses are adjusted to enable the smart cities to operate efficiently (Ouerhani et al., 2016). A straightforward application of IoT sensing would be sensors such as those deployed within smart parking meters to monitor the location of available parking bays in a city. Data collected by these sensors are transmitted and stored centrally (Patti and Acquaviva, 2016).

Libelium outline nine categories of sensors which can provide telemetry in a smart city. These categories are listed in Table 1. In addition to the categories listed, there are additional sensor types generally associated with smart cities, concerning the smart grid utility providers and metering.

<table>
<thead>
<tr>
<th>Sensor category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smart parking</td>
<td>Monitoring of parking spaces availability in the city</td>
</tr>
<tr>
<td>Structural health</td>
<td>Monitoring of vibrations and material conditions in buildings, bridges, and historical monuments</td>
</tr>
<tr>
<td>Noise urban maps</td>
<td>Sound monitoring in bar areas and centric zones in real-time</td>
</tr>
<tr>
<td>Smartphone detection</td>
<td>Detect smartphones and in general any device which works with WiFi, Bluetooth, or cellular interfaces</td>
</tr>
<tr>
<td>Electromagnetic field levels</td>
<td>Measurement of the energy radiated by RF capable devices</td>
</tr>
<tr>
<td>Traffic congestion</td>
<td>Monitoring of vehicles and pedestrian levels to optimize driving and walking routes</td>
</tr>
<tr>
<td>Smart Lighting</td>
<td>Intelligent and weather adaptive lighting in street lights</td>
</tr>
<tr>
<td>Waste management</td>
<td>Detection of rubbish levels in containers to optimize the trash collection routes</td>
</tr>
<tr>
<td>Smart roads</td>
<td>Detection of rubbish levels in containers to optimize the trash collection routes</td>
</tr>
</tbody>
</table>

Adapted from (Libelium).

- Eavesdropping: If the communication between the sensor and the centralized server is not secure, the integrity of the data could be compromised. As data are transmitted to the centralized server, communications could be intercepted and manipulated which could cause the sensors to relay incorrect actions and the servers to record incorrect events (Mukundan et al., 2014; Pardeshi and Borade, 2015; Baig, 2014).
- Data loss: Insufficient data management of sensors could impact the operations of a smart city. Data management refers to deployment practices, procedures and policies for utilizing sensors effectively and securely in a smart city environment. If data are not managed adequately, sensors and the sensor data collected, transmitted and stored could be compromised (Mishra et al., 2015).
- Availability compromise: In the event of sensor failure there should be procedures and plans in place to avoid negatively impacting the operations of a smart city. For example, if traffic congestion sensors incorrectly display green lights on all main roads, the smaller side roads would have congestion problems. Without procedures and plans in place a sensor failure could lead to traffic grid lock (Khan et al., 2016).
- Remote exploitation: As sensors connected to a smart city would communicate to a centralized server, insecure communication channels could be used to perform remote exploitations. Remote exploitations could be launched from the main servers, connecting nodes or even an individual sensor and potentially propagate through the network (Cisco, 2015).

Sensory data are of high forensic value as the ability of these devices to capture events such as: movement patterns of a mobile device or a vehicle over time, can prove to be of significance for forensic investigations. Not only is sensory data of value for investigation, rather the data stored at the centralized base stations is also essential in confirming the sensory readings. Moreover, the acquisition of data from actual sensors placed in their respective environments may not always be practical. Much of the data generated by IoT devices is stored in the Cloud, which presents challenges and opportunities for digital investigations.

**The Cloud**

The data generated in a smart city from each of the four components described above is stored in the Cloud for convenient access by all stakeholders. The threats to the individual components...
of the smart city do indeed affect the security of not only the smart city devices and software, but also the data that are to be stored in the Cloud. Consequently, uncertainty is created over data accuracy and its admissibility as forensic evidence (Quick and Choo, 2014). We now describe the threats posed to smart city data which are transmitted and stored in the Cloud:

- **Data leakage:** When moving the infrastructure or resources to the Cloud, smart city components relinquish control over the data to a third-party Cloud provider. As the data are then hosted on a multitenant environment, data can potentially be accessed by an adversary or even third-party provider personnel (Sabahi, 2011).
- **Insecure APIs:** Most software and application connected to the Cloud infrastructure use APIs to interact with Cloud services. Hence, the APIs used must support secure communication with authentication, access control, encryption and activity logging and monitoring mechanisms (Ashktorab and Taghizadeh, 2012).
- **Malicious Insider threats:** Many Cloud service providers do not reveal the clearance and screening procedures of their personnel and how they grant access to the resources of the client organization (Behl, 2011).
- **Denial of Service Attacks (DoS):** Since an outside party is hosting the services on the Cloud, it is often easy for an adversary to extract information about the smart city infrastructure, as its hosted publicly making data publicly accessible.
- **Malware Injection:** Most Cloud providers host web applications via middleware platforms. If the web applications and servers are not securely configured or patched, an adversary can leverage this opportunity and carry out various malicious scripting-style attacks.
- **System and Application vulnerabilities:** The technology and applications are managed by the Cloud provider and third-party providers, and hence the smart city has no control on management and security. The Cloud provider must be trusted to provide them with robust and secure services.
- **Data Locations and Regulation boundaries:** A smart city may not get a choice on where its data are going to be stored. Hence it makes it difficult to manage the security of the data location.

For instance, if the data are stored in a data center located in another country, a data owner may lose control on how the data are secured, depending on the contractual obligations of the Cloud provider and the local privacy legislation.

### Digital forensic challenges and value for smart cities

In the previous section, we have highlighted the security threats against the smart city components, and the value of forensic preparedness to protect the smart city. In this section, an analysis of the challenges faced when conducting digital forensic investigations in the smart city domain is presented. In addition, the importance of digital forensic investigation in the smart city is discussed.

Due to the interconnected and heterogenous nature of smart cities, situations arise where all the difficulties of each component area are present. In a sense we are presented with the worst of all worlds as the interconnected nature of smart cities means that both volatile and non-volatile, open-source and proprietary systems are involved in transactions and data flows. Therefore, digital forensics for embedded, Cloud and IoT devices is intense and challenging. This is because not all IoT devices have the same network and application architecture. Digital forensics of devices hosted in the Cloud environment is made difficult by the absence of third party agreement with the client, which can allow forensics investigators to access the data stored on the Cloud. Data stored in different countries brings further challenge to digital forensic investigators due to judiciary disparities and information laws. These threats and challenges to the data stored on Cloud infrastructure will also apply to the IoT devices interrogated during forensic investigations (Oriwoh et al., 2013).

An illustration of the forensic value of data stored by smart city devices in a Cloud, is given in Fig. 3.

#### Smart mobile devices

Traditional digital forensics dealt solely with ‘persistent data’, specifically data that is not erased or altered when the artefact of
interest is powered down (Nolan et al., 2005). Many traditional
digital forensic models assume the investigator is in physical
possession of the digital hardware or persistent storage. Subse-
quently, the traditional forensic process assumes that the device in
question can conform to the recommended digital forensic practices
such as; being powered off, the persistent storage is physically
removable, and it is possible to safely duplicate data (through the use
of forensic write blockers). However, smartphones for instance have
changed this traditional landscape and created challenges, which are
yet to be completely recognized or addressed. For instance, open
source and proprietary operating systems may limit the methods by
which an investigator is able to communicate with or locate desirable
evidence on a smartphone (Cusack and Lutui). An abundance of
unique smartphone hardware with insufficient technical documenta-
tion further inhibits the ability for investigators to adequately
understand the intricacies of the underlying architecture.

The freedom for an end-user to install, manage and use
personalized applications creates additional challenges in that inves-
tigators may no longer know if the data of interest are encrypt-
ted, stored locally or in the Cloud (Kechadi et al., 2015). Many of
the aforementioned issues will become mainstream when IoT sensors,
the Cloud and smart cities are increasingly included within digital
forensic investigations. The evolution of mobile computing has also
seen the emergence of devices such as smart watches or wearable
computing devices, which have created both benefits and hin-
drances for traditional digital forensics approaches. Wearable mo-
bile computing devices may operate independently or through a
secondary device such as a smartphone or computer.

The types of data collected through wearable mobile computing
devices may either support or refute accusations or claims made
towards a third party (Snyder, 2015). However, investigative digital
forensic models typically assume that the investigator is in com-
plete possession of the physical artefact at the time of acquisition.
Wearable devices as an evidentiary artefact may thus limit or
inhibit an investigator's ability in having control or being in
possession of a device during an investigation. There have been
instances where data from wearable devices have successfully been
used as evidence in a court of law (Olson, 2014). However, due to
a number of inherent privacy-influenced security flaws within the
original design of some devices, data at rest and in transit is now
encrypted by default (Schellevis et al.). The use of encryption
techniques to ensure the confidentiality of data has created
ongoing issues for successful forensic investigations (Casey and
Stellatos, 2008). In an effort to protect the privacy of the end-
users, vendors may place even greater emphasis on implement-
ing cryptographic techniques to protect data in the future. The use
of privacy and security techniques often inhibits or delays suc-
cessful digital forensic investigations.

Dynamic network without infrastructure

The transition from traditional to IoT and smart city forensic
investigations will in itself raise a number of complex issues. User
data are stored in multiple locations; therefore, forensic in-
vestigators must go through various jurisdictions to access the data.
This is a significant challenge for forensic investigators as many IoT
devices use the Cloud to store data. Forensic investigations will be
faced with issues relating to privacy and access to sensitive data.
This is because IoT devices can connect to private, public and
organizational networks. For instance, if the IoT device was con-
ected to a hospital network to access patient records, then inves-
tigators will need to access sensitive patient data for forensic
purposes, which will raise the issue of privacy.

Oriwoh et al. (2013) suggests that prominent future issues for
investigators will include: the expansion and inclusion additional
evidence sources; the inclusion of multiple devices in an inves-
tigation to devise conclusive facts; the inclusion of evidence
which will not be limited to standardized file formats and instead
will be dependent on a vendors proprietary data types; large
quantities of data to analyze; and undefined boundaries between
device data and ownership. Whilst it has been acknowledged that
models and best practices will need to adapt to the evolving
environment, little research has been undertaken with regards to
existing tools and their limitations in being able to analyze data
produced by IoT devices and networks. This limitation puts the
investigator at risk of being unable to derive facts to support
theories and hypothesis during a forensic investigation (Plachkinova et al.).

It would be disingenuous however to dismiss the evidentiary
potential offered by data generated by smart cities. We are pre-
sented with an unprecedented wealth of evidence of the move-
ment, activities, and behaviors of the inhabitants of a such a city.
The sensory apparatus of a smart city has potential to provide input
from CCTV, motion detectors, air quality sensors, smart meters and
RF sensors, setting up a unique avenue for law enforcement to carry
out an investigation. For example, the availability of historical data
as it relates to noise, both acoustic and electromagnetic can be
utilized in order to track the movement of individuals in areas
where CCTV is not present. This data combined with telemetry
from environmental monitoring (such as air quality) can be used to
profile the use of individual vehicles. Such evidence is of value both
in the correlative sense as well as supporting investigative
functions.

One such example is identified by Oriwoh and Conrad (2016),
who examined data from motion detectors in conjunction with
contextual information in order to evaluate the number of occu-
pants present within an area. Their work extended that of Novk,
Bias, and Jakab (Novák et al., 2012), which made use of an ANN to
detect anomalous behaviour after an initial training period in which
a baseline is established. In contrast, Oriwoh and Conrad (2016)
focused on combining motion data with the state of non-living
objects such as smart phones to determine if motion is legitimate
or intrusive. It is possible that similar models could be applied on
a much larger scale via the smart city paradigm. Regrettably, there
exists the potential for criminals to make use of the same data to
more intelligently and efficiently target their existing criminal ac-
tivities. Examples of this have been identified in locationally-aware
social media systems, in which users were identified and targeted
based on household occupancy data (Gambis et al., 2010). In the
smart city scenario there is potential for similar technology-
enabled crime on an expanded scale, as unlike social media,
monitoring in smart cities is not opt-in. For example the availability
of CCTV and motion detectors provide mechanisms to determine
the average activity within an area at a particular time, air quality
sensors allow for the tracking of vehicles and smart meters provide
details of when individuals present at a particular property. In
aggregate, such data enables a new avenue for criminal enterprise.

Forensic investigators expect to focus on various sources of
information.

These could include computer systems, network devices, mobile
phones, USB drives and hard drives. However, for IoT forensics,
objects such as household appliances, smart cars, smart homes,
digital cameras and various other IoT devices poses challenges for
device-level investigations (Oriwoh et al., 2013). The number of
devices, with IoT devices being interconnected with various other
technologies and devices, creates a vast collection of IoT devices.
Any crime, including robbery and homicide, can have vital clues on
IoT devices. Therefore, when investigating a crime in this context,
forensic investigators might need to analyze data from devices that
the offender or victim came in contact with.
Data formats and logs

The size and format of data is expected to vary significantly from one case to another. This is because of the intercommunication and exchange of data and information between various IoT devices across the network. Also, the format of the data is not the same or not common to normal digital forensics processes. Therefore a digital forensic investigator must unveil, understand and reformat to readable and usable format.

All the four components as well as the underlying IoT sensors and the Cloud are prone to communication interception/modification as well as frequency jamming attacks. Encryption of data transmitted over the wireless channels is one approach for attaining end-to-end data security. For instance, the various types of data transmitted from a UAV to a central base station include: flight (including videos), connection, system, and GPS data. Whilst the storage of all data may not be feasible from a network bandwidth as well as storage cost perspective, flight and system logs contain highly desirable data required for forensic analysis. However, the extraction of live state information whilst a UAV is powered down remains an unaddressed challenge. Preservation costs associated with storage of UAV-generated data, is a factor that needs to be considered when designing a forensically-sound data transfer and storage mechanism.

In smart grids, various solutions have been proposed in the literature to tackle the issue of security. However, despite the presence of effective countermeasures for each identified threat, many challenges are still open for research. Some of the challenges that still need to be resolved include; the requirement to analyze large volumes of ingress data of the smart grid at near real-time speeds, so as to identify adversarial attempts to penetrate the network. In addition, the lack of a standard security solution encumbers data verification on the diverse set of smart grid components comprising smart meters, network transmission lines, wireless channels (with varied security protocols) and sensors. Smart grid security solutions (as given in Table 2), are lacking in key security domains including availability of smart grid resources, tamper-proof smart meters, secure network transmission lines (both wired and wireless), and the presence of solutions including key management, user authentication and data encryption mechanisms, to achieve data confidentiality, integrity and availability (Colak et al., 2016). The level of requirement for the forensic data obtained from smart grids and subsequently stored in the Cloud comprises connection information, client usage patterns as well as utility-provider system logs.

For Building Automation System (BAS) networks, the following security threats associated with storage of data generated by individual components of the BAS, including intrusion detection systems and firewalls: the presence of highly trusting devices accompanied with the non-existence of a rapid device/software refreshing cycle, encumbers the data verification process. In addition, the lack of secure networking protocols to interconnect the various components of a BAS as well as the lack of message authentication, also make it hard to attain seamless network security. Several types of logs are generated by BAS devices for storage in the Cloud including trends and connections. Device data logging in a BAS network is typically manual and requires strong policy definitions to ensure that log data transferred for subsequent storage in the Cloud is not tampered with during communication, and is thus submissive for forensic analysis.

Cloud forensics

Data collected at the Cloud from individual smart city components is also vulnerable to security threats posed by an adversary. Several aspects of Cloud security point towards the proneness of the Cloud to data leakage, malware injection, and network intrusions. The transient nature of some Cloud data encumbers the verification of findings of a forensic investigation. Moreover, the geographical boundaries and regulations may make it difficult to retrieve Cloud data required to carry out forensic investigations postincident. However, the corroboration of evidence extracted from Cloud-based data with data stored in the individual devices of the smart city (from the four other smart city components), will increase the reliability of forensic data.

The basic steps of traditional digital forensics include collection of the medium in the scene, preservation of the content in the medium, verification, analysis, interpretation, documentation and presentation of the result in court. However, Cloud forensics is not as simple as traditional digital forensics. The main challenges come from the data acquisition, static, elastic and live forensics, evidence segregation, virtualized environments, internal staffing, multi-jurisdictional laws, external dependency chain of custody and finally service level agreements. The preservation and verification of the potential evidence, and legal issues should be of concern. Hence, Cloud forensics will be more complex, time intensive and expensive. Analysis of Cloud artifacts as those acquired from client devices helps portray the state of the system, though proving to be challenging as Cloud data has no universally accepted standardized representations (Roussev and McCulley, 2016).

The evolution of Cloud computing forensics is in its infancy. Currently there is not a standard method or tool set for conducting Cloud investigations, or even for evaluating and certifying proposed tools. The presentation of evidence derived from a Cloud service will likely be problematic in the near future.

Case study: reckless driver in a smart city

An illicit driver enters a smart city attempting to subvert all policing controls in place for traffic regulation. The driven vehicle enters the city and increases speed past the stipulated limit, threatening other motorists and commuters alike. Two scenarios can be identified here. In the first case, the vehicle may not be a smart vehicle and so would be detected by road-side IoT sensors as violating the speed limits stipulated by the city council. In such a case, an incident response team will be alerted immediately by sensory data which is transmitted over to the centralized Cloud and measured as being anomalous by the relevant data analysis engine. Subsequently, the traffic control authority (the police) will be alerted with the data emerging from the Cloud, and necessary tactical action will be taken to control the incident.

A second scenario may emerge if the violating vehicle is a smart vehicle, in which case the vehicle controls will automatically alert the driver of the impending danger associated with speeding. In addition, the smart vehicle will communicate with the Cloud to update various parameters including the vehicle’s coordinates, speed as well as the numbers of passengers in the vehicle. With additional data to facilitate law enforcement action, the vehicle can either be alerted and guided using a UAV and/or relevant personnel from the law enforcement may be dispatched for taking tactical action to control the vehicle and the traffic.

After having responded to the incident, relevant digital forensic investigators will retrieve data from the smart vehicle itself, UAVs involved in tracking, and the IoT sensors. The data can also be acquired from the Cloud, and can be cross-checked for validity. The threats posed to the acquisition, transmission and subsequent storage of criminal data from the vehicle are:

- Eavesdropping on the vehicle data in-transit to the Cloud (Data confidentiality breach),
## Conclusion

Challenges associated with acquisition and storage of smart city data emerging from its individual components i.e., Smart grids, Building Automation Systems, Unmanned Aerial Vehicles and Smart Vehicles; and enabling IoT Sensors, Cloud platform, remain largely unaddressed. Through this article we have provided a thorough insight into the smart city threat landscape for each of its four components and as well as for the enabling technologies/platforms. A detailed assessment of the type and source of data originating from the smart city components, for secure storage in Cloud was also elaborated upon. A summary of our findings are presented in Table 2. Based on our findings, some of the questions that could be posed for forensic investigation on smart city data include:

- Are log files or memory dumps available for recovery from individual devices?
- Can the data stored on commercial Cloud providers be accessed and/or recovered?
- How principles of forensic investigation can be applied so that evidence is not altered during the course of an investigation?
- And lastly, what would be the optimum level of forensic preparedness/readiness for a given smart city service?

### Table 2
Summary of the smart city security and forensic data landscape.

<table>
<thead>
<tr>
<th>Security threats</th>
<th>Data sources</th>
<th>Forensic data (Level of requirement)</th>
<th>Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smart Grids</td>
<td>Protocol vulnerabilities, Privacy, Eavesdropping, Rogue/infected devices, Attacks on internet connected devices</td>
<td>Communication gateways, Smart meters, ICS systems, Smart Appliances, Sensors/Actuators, Firewall, IDS/IP</td>
<td>Connection logs (syslog, console logs, network packet capture): High, Usage logs: Medium, System logs (authentication, OS, application): High</td>
</tr>
<tr>
<td>BAS</td>
<td>Highly Trusting devices, Long device lifecycle, Lack of source authentication, Insecure protocols</td>
<td>Internal Logger devices, Firewall/IDS (if implemented)</td>
<td>Trend logs (value changes): High, Connection Logs: High, CSV: High, Other DB format: Variable, SQL: High</td>
</tr>
<tr>
<td>Smart vehicles</td>
<td>Physical threat, Communication interception, Communication jamming/DoS, Data security</td>
<td>GPS, Infotainment system, Event data recorder</td>
<td>System logs: High, GPS data: High</td>
</tr>
<tr>
<td>IoT sensors</td>
<td>Maintaining confidentiality of data, Secure communication, Data management, Data storage, Sensor failure, Remote exploitation</td>
<td>Smart parking meters, Structural health measurement devices, Smartphone detection sensors, Electromagnetic field level monitors, Traffic congestion monitors, Smart lighting sensors, Waste management monitors, Smart roads sensors</td>
<td>System logs: High, Connected device logs: High, GPS data: High</td>
</tr>
<tr>
<td>Cloud</td>
<td>Data Leakage, Malicious insider threat, Insecure API, Denial of service (DoS), Malware injection attacks, System and application vulnerabilities, Data locations and Regulation boundaries</td>
<td>Firewall, Database, Application, IDS and IPS, Active Directory, Cloud system</td>
<td>Connection logs, IPS and IDS logs, user logs: High, Database logs: Medium, Application logs: Low, System logs: High</td>
</tr>
</tbody>
</table>

- Tampering of the in-transit data (Data integrity breach),
- IoT sensor failure (Hardware fail — no breach),

A challenge associated with smart city cyber-security is to maintain services that refresh encryption keys regularly to avoid persistent tampering of sensor data in case of key compromise. Similarly, maintaining secure transmission lines, such as those that adopt protocols such as IPSec for network security, are also essential for attaining the goals of cyber-security for the smart city.

Following the retrieval process, the data will be analyzed by the forensic investigators to answer several key questions associated with the investigation:

- The response time taken by the law enforcement to contain the incident,
- The precise vehicle manoeuvres carried out by the perpetrator,
- The effect of the actions of the perpetrator on the city traffic, and
- Actions that could have been taken to better contain the threat posed. Through the coordinated effort of multiple smart city components i.e., data sources, not only is proper incident handling achievable, but also accurate and relevant data is conveniently provided from various distributed sources, to facilitate the digital forensic investigation process.
In summary, smart city data is generated in vulnerable environments by all data sources, for storage in a backend Cloud. Consequently, security of the data transmission and storage facilities is essential in order to preserve forensically valuable evidence, required for conducting investigations for committed cyber-crime in the smart city. Based on the analysis of the threat landscape of the smart city presented in this article, it is essential to have relevant security controls and forensic readiness in place to ensure that data transfer through the ICT infrastructure of the smart city into the Cloud is secure and available for preventing, detecting, and resolving incidents. The security of the data stored in the Cloud is effective only if data is kept confidential i.e., not leaked out to non-owners, is integrity checked, and is available for ready access. During a forensic investigation, the evidence thus derived from the smart city data can then be used to establish guilt of cyber criminals, and to feed into the smart city security improvement process, to consequently improve the overall security of the smart city ICT infrastructure.

References