Identification of IoT Devices for Forensic Investigation

Nancy Scheidt
School of Computing
University of Portsmouth
Portsmouth, United Kingdom
nancy.scheidt@port.ac.uk

Mo Adda
School of Computing
University of Portsmouth
Portsmouth, United Kingdom
mo.adda@port.ac.uk

***Abstract*—As the Internet of Things (IoT) environment rapidly grows, so do the distribution and utilisation of IoT devices as well as Cyber-Enabled and Dependent Crime activities. This leads to challenges investigators face when forensically investigating the IoT to present evidence in the court of law. Current tools and frameworks demonstrate a variety of limitations which need to be improved to efficiently deal with IoT forensics. In this article, we introduce a novel approach to uniquely identify IoT devices to improve IoT Forensics while introducing new terms, such as the DNA and Genes of devices. These novel terms for a dynamic and complex environment, such as the IoT, will ease the understanding of unique device identifiers due to its well-known association with traditional crime scene investigation. To elaborate on the importance of unique device identification we establish the context of the IoT, its Forensic Investigation and challenges as well as the approach of Universal Identification Numbers utilising Set Theory. Additionally, we consider the implementation within our proposed IoT Server model environment to demonstrate the necessity in IoT Forensics.**

Keywords*—*Internet of Things; IoT Forensics; Universal Identification Number; Set Theory; Hybrid Systems; DNA of Devices; IoT Server; Hybrid Server

I. Introduction

Every day the number of Internet of Things (IoT) devices per person rises rapidly and will do so continuously. With the fast development and improvement of IoT devices it opens more doors for the variety of usage such devices offer [1]. This provides a multitude of new challenges in digital forensics and increases the complexity of accessing the devices' information if needed. Therefore, worldwide investigators face the issue of collecting evidence of persons of interest if information relating to a case is stored in the IoT environment, which includes devices, networks and the cloud. Solutions to tackle this issue focus mainly on defining specific principles targeting IoT forensics, improving cloud cybersecurity and device-specific approaches due to the differences IoT devices offer when needing to be analysed and investigated [2, 3]. Moreover, [4] consider issues with large data sets and evidence collection on an IoT server which can register devices in their model to provide further solutions to IoT forensics. Therefore, previous research shows that IoT Forensics is still handled similarly to digital forensics in terms of investigation and evidence extraction and is not yet standing by itself.
Considering the research in this field the main disadvantages are linked to the challenges in managing the increasing number of IoT devices worldwide, finding appropriate investigating tools which can be utilised with a variety of operating systems, models and purposes of an investigated device [5, 6]. Furthermore, the storage space for evidence as well as globally working with investigators on criminal cases is still in need of improvement [6].To provide a solution, covering the issues raised in previous research, we aim to answer the following questions which also clarify our contribution in this topic matter:

- How will Universal Identification Numbers help identify suspects within the IoT in the context of Forensics?
- How is the novel way of DNA for devices beneficial in forensic investigations?

The rest of this paper is organised as follows: section II elaborates previous and related work to our proposed research; section III introduces our new terms of DNA of a device as well its implementation and a short case scenario. Finally, this paper is concluded in section IV with recommendations for the following research as well as a short outline of future work.

II. Related Work

This section presents an overview of the IoT, Forensic IoT Investigation and its current issues, additionally, we will touch upon Universal Identification and Management of devices utilising Set Theory when defining such.

A. IoT

The key idea behind the IoT is an infrastructure enabling connection and utilisation for information exchange between

individuals but also devices and it is an environment which is "forever expanding" [2]. The rapid growth of IoT devices is highly apparent in today's society, be it for governmental, corporate or personal use and the billions of devices in use are only speculated to increase in the future [7]. These devices enable the gathering of information and communication with their surroundings while utilising networks and their in-built sensors. Additionally, this leads to homes as well as other frequently used environments such as schools, hospitals and supermarkets being constantly enhanced as well as interconnected on various levels and networks and, therefore, advance daily tasks and communication while providing a variety of services [4, 8, 9]. However, this also results in a more extensive area which is being vehemently attacked and offers a wider variety of utilisation in terms of crime activities [2, 7]. Such activities range from Cyber-Enabled (i.e. fraud, theft, child pornography) to Cyber-Dependant Crimes (i.e. hacking, DDoS attacks, viruses) [10, 11].

B. IoT Forensics Investigation and Current Challenges

The increase of Cyber-Enabling and Dependant Crimes calls for more concentrated measures. Therefore, it is of vital importance that the forensic investigation in the IoT environment faces constant improvement, this is especially essential due to the rapid development of technology overall, which will only continue [12]. Moreover, the capacity of data is increasing which is elucidated when portraying IoT forensics as an investigation on the three different levels: Cloud, Network, and Devices [4, 7, 13]. Therefore, digital evidence is of high capacity and needs to be targeted, which [14] aim to do in their designed Open Source Tool, which can screen a variety of platforms to extract evidence by reducing data. Furthermore, the variety of digital forensic investigation tools also reach their limits when competing against the dynamic of the IoT and its devices and as the question of officially defining IoT Forensics arises, we can summarise the examination process as identifying, preserving, analyzing and presenting evidence to the court of law, similarly to digital forensics [6, 15]. Additionally, IoT forensics is a highly complex procedure and rises a variety of security as well as investigation issues and its evidence can be classified into three groups [16]:

- Smart Devices and Sensors
- Traditional Hardware and Software
- Hardware and Software (cloud, social networks, ISPs and mobile network providers, virtual online identities, the internet).

Previous research also demonstrates a variety of challenges which are highly necessary to consider when dealing with IoT Forensics [1, 13, 14, 15]:

- Location of Information
- Type of Device
- Format of Device
- Lifespan Limitation (overwriting of data)
- Preserving the Crime Scene (IoT Environment)
- Lack of Security
- Limitations of current Forensic Tools
- Legal or Jurisdictional Predicament
- High number of Investigating Tools which all cover different tasks.

Considering all these issues, it is important to manage the amount of IoT devices worldwide and tackle the challenges step by step to ensure a successful improvement in the investigation of the IoT environment

C. Universal Identification Number and Set Theory

The increase of IoT devices and the forensic investigation issues arising from the constant development ask for specific measures. Research has dealt with multiple aspects of IoT technology, its application and methods of how devices can be addressed and identified more efficiently [17]. One of such methods, regarding device identification, is demonstrated by [18]. Making use of the formal method and set theory, allowed [18, 19] to simply utilise previous terms as well as develop novel terms and formulae for a Universal Identification Number (UID) for IoT devices which he referred to as the Global Actor Relationship Identifier (GARI) method. For this process, the methods utilised, allowed to approach the issues of IoT device management and identification in the most sensible way. This allows a solution to be employed that was mainly based on a logical and mathematical understanding to then test, secure and develop ways for verification in this area [18, 20]. Considering previous research and the necessity to link such to forensic investigation, we have developed the unique way of IoT device management and identification further to fit the IoT forensic context and created novel terms not only for our research purposes but also to improve the understanding of such, as it will be demonstrated in detail in section III-A.

III. Identification of IoT Devices

A. Model of Devices' DNA and Genes

As briefly touched upon, in traditional crime scene and forensic investigation, the DNA of a living being (further on referred as traditional DNA) can put such at a scene as well as help to create a timeline of events to build a solid base of evidence to then present a case before a court of law [21, 22]. These pieces of evidence are very crucial during an investigation, be it a physical or digital one. Previous research has suggested methods to manage IoT devices globally, however, these did

not consider variables such as the implementation of identification numbers in an IoT Forensics environment, which is of high necessity in today's day and age of SC and the IoT usage increase [19].

Our proposed idea suggests, an approach similar to DNA analysis regarding IoT forensics, however, instead of using traditional DNA, as we are talking about technology, we propose a method to allocate a unique identification number to IoT devices worldwide. Traditional DNA is a chemical in the structure of a double-helix and provides a living organism with information such as the function of a structure, moreover, traditional Genes refer to very short sections of traditional DNA who provide very specific information on characteristics such as ear shape, hair or eye colour [23, 24].

In our case of IoT Forensics and the allocation of a unique identification number, we will call these numbers the DNA of a device which is made of unique attributes, which we will be calling Genes, Figure 1 visualises how we aim to use the information to construct a device's DNA. An individual (Actor) is purchasing an IoT device and the combination of information, hence Genes, creates a unique string of numbers, the DNA, of the device.

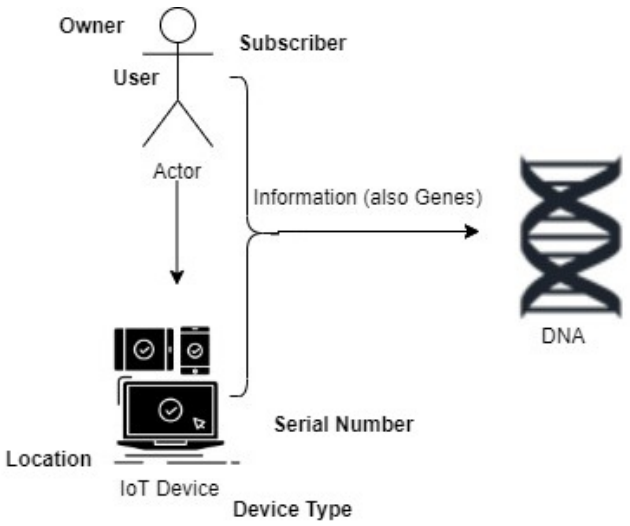

Fig. 1. DNA Development

Table I highlights each of the Genes and their meanings. The owner*(O)* not necessarily has to be the subscriber *(S)* or user *(U)* of a device *(D)* but is the individual having purchased such, whereas the subscriber registers the device for usage on a database. Serial Numbers *(SN)* are always crucial to consider in identifying devices and so are their types *(DT)*. The Location *(L)* of purchase and initial use will support the unique design of the DNA and can provide supplementary information utilised to create such a unique DNA string for a device.

As shown in Figure 2, these six Genes will be allocated specific numbers, unique for each IoT device, which will then be allocated with the DNA_n.

TABLE I
GENES OF A DEVICE

Genes	Attributes
Owner	Individual/Company who purchased a device
Subscriber	Individual/Company registering device, could be owner or receiver (i.e. company phone, present)
User	Individual/Company using device, could be owner or receiver (i.e. company phone, present)
Serial Number	Unique Serial Number of Devices
Location	Place device has initially been registered
Device Type	Brand, Model

Therefore, similar as to UID the DNA would be demonstrated based on a string of a unique series of numbers and this DNA string shows all the necessary information of each section needed for registration, identification and analysis of devices which will support and improve IoT forensics and the investigation process. The following method will show how to calculate the DNA by simply utilising the set method.

$$DNA_n = (O_D, S_D, U_D, L_D, SN_D, DT_D)$$

Thus the DNA is individually created and summarises all the Genes which hold the information needed to uniquely identify an IoT Device. In terms of IoT Forensics, this method and new terminology provides a clear structure and enables a more efficient investigation process.

B. Implementation of DNA

Having a unique identifier for each IoT device is very important and beneficial when forensically investigating the IoT environment. Our approach of new terminology will ease the process to understand and provide the necessary support to successfully investigate, however, this is only the first step in improving IoT forensics. The idea of unique identification is closely linked to our idea of a Hybrid Forensic IoT Server (further referred to as Hybrid Server) which is going to be part of our future research. We will implement the DNA in an IoT Server environment. The idea of our Hybrid Server allows IoT devices to be registered on the server while providing it with all the necessary information the DNA requires. Previously, [4] have suggested a single Forensic IoT Server Model. However, as Figure 3 demonstrates we are aiming to design the Hybrid Server on a hierarchical and distributional level. Hence, a main server will be implemented in each country, sub-servers will be distributed in the regions respectively and the retrieved information and evidence from the devices will be stored on the Cloud for a specific time based on the countries individual legislations. This Hybrid Server is built by the government,

Fig. 2. Addressing of the unique DNA String of a Device

authorities such as the police and army, as well as service providers to ensure the public's safety, improve investigation processes but also for profit purposes due to devices only having access to the basic services (calls, messaging) if not being registered on the Server System. This is due to our idea not aiming to eliminate current digital forensic investigation tools but rather to improve the investigation process in cases of a larger and more complex structure, such as international hacking, fraud or trafficking. Generally, the government, such as the city council, will maintain and manage the Hybrid Server system, however, have no access to any information of the databases. These will only be accessible by official court orders to be then analysed by the appointed investigator.

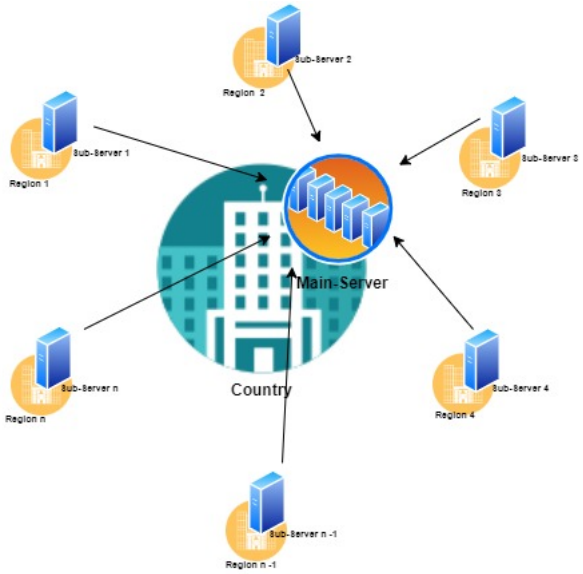

Fig. 3. Structure of future work, Hybrid Forensic IoT Server

C. Hypothetical Case Scenario

To better comprehend the implementation of our approach, we will consider the following scenario as visualised in Figure 4.

Actor A *(A)*, who is residing in Country B *(B)*, has a high number of IoT devices in personal use, from their smartwatch, over multiple smartphones to smart appliances in their kitchen. From their workplace, Actor A received a Laptop as well as a smartphone. All their devices were registered on the sub-server in Region M *(M)* of Country B, which they are currently residing in and where the devices have been purchased initially. Their company laptop *(CL)* and phone *(CS)* were purchased and registered by Company C *(C)* which they are working for, so the owner and subscriber of these devices is the company, whereas Actor A is registered as the user. Other devices in Actor A's possession and use were registered by themselves. Considering our DNA formula presented in section III-A, the DNA of 1) the company laptop, 2) company smartphone and 3) personal device (PD) will be constructed as follows:

1) $DNA_n = (C34, C34, A12, BM65, 976, CL768)$
2) $DNA_n = (C34, C34, A12, BM65, 785, CS928)$
3) $DNA_n = (A12, A12, A12, BM65, 498, PD453)$

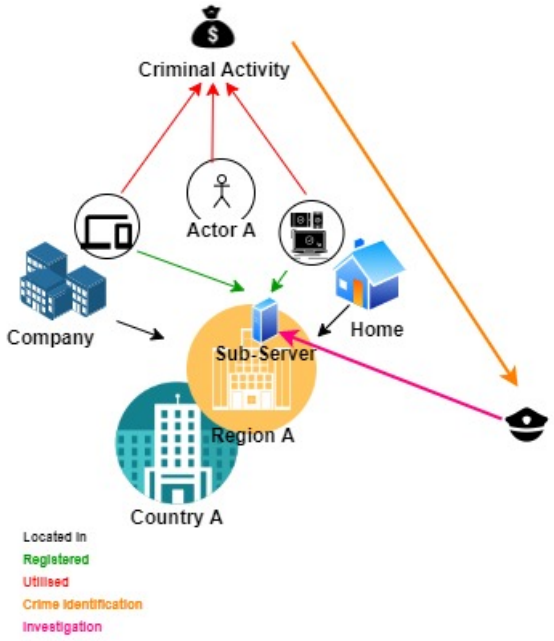

Fig. 4. Case Scenario

Hence, the specific Genes are calculated by Actor and Device information using a combination of letters and numbers and the DNA string of these devices would be demonstrated as follows, in the Hybrid Server environment:

1) C34C34A12BM65976CL768
2) C34C34A12BM65785CS928
3) A12A12A12BM65498PD453

Actor A decides to commit fraudulent activities utilising his company as well as personal devices, to lessen the possibility of leaving a coherent link between the crime and themselves. However, as the devices were registered and all allocated with their unique DNA, Investigator I is able to trace the criminal activities back to Actor A, even though, Investigator I is not able to retrieve evidence from the devices itself, the Hybrid Server is able to show the device's activity and links it directly to its user with help of the unique DNA strings.

D. Discussion

As demonstrated in the previous sections, the novel approach in uniquely identifying IoT devices with DNA strings can provide new opportunities for forensic investigation within the IoT environment. Due to the six specific Genes, devices will be individually identified and widen the opportunity to link them directly to suspects even if the physical devices cannot be accessed by the investigator. It enables and provides a more organised and beneficial way of device management, moreover, our approach allows to deal with challenges efficiently while taking previous research further and focusing more precisely on IoT forensics, which is a necessity due to the rapid growth in numbers and usage of IoT devices and its utilisation in criminal activities. Limitations of this papers' research extend to our future research and proposed model for a Hybrid Server as shortly outlined in section III-B.

IV. Conclusion

In this paper, we have proposed an effective framework in uniquely identifying IoT devices with their Genes, which construct the structure of devices' DNA. Due to the increase of challenges the IoT environment creates, while enabling users to make use of these services on different levels; personal, professional and criminal, it is crucial to improve IoT Forensics just as rapidly as the technological development. Our DNA model demonstrates important forensic factors which were not considered in previous research regarding device management. Moreover, this approach will be of high evidential value, improve device management and examination efficiency during forensic investigations while establishing a case to present in the court of law.
Considering future research, as we briefly touched upon, our future projects will focus heavily on the implementation of the DNA in our proposed model of a Hybrid Forensic IoT Server, which will support the current IoT forensic investigation process greatly. We suggest the hierarchical and distributional Hybrid model to enable main-servers and sub-servers to be in constant communication, also allowing an easier exchange of evidence data worldwide. Implementing the DNA in this Hybrid Server environment will ensure to tackle and deal with the number of issues and challenges regarding IoT Forensics, as highlighted in section II-B. Moreover, we are aiming to test the Hybrid Server by utilising mobile phones to test the real-life location updates and synchronisation on the server. Additionally, future research is aiming to enhance investigators teamwork and data exchange, nationally but also globally, preserve the crime scene in an IoT environment for a longer time period and therefore, oppose the constant overwriting of data and improve trust values regarding jurisdictional predicaments worldwide.

References

[1] B. Safaei, A. M. Monazzah, M. Bafroei, and A. Ejlali, "Reliability side-effects in internet of things application layer protocols," 12 2017.
[2] A. MacDermott, T. Baker, and Q. Shi, "Iot forensics: Challenges for the ioa era," in *2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*. IEEE, 2018, pp. 1–5.
[3] F. Servida and E. Casey, "Iot forensic challenges and

opportunities for digital traces," *Digital Investigation*, vol. 28, pp. S22–S29, 2019.

[4] S. Zawoad and R. Hasan, "Faiot: Towards building a forensics aware eco system for the internet of things," in *2015 IEEE International Conference on Services Computing*. IEEE, 2015, pp. 279–284.

[5] M. Conti, A. Dehghantanha, K. Franke, and S. Watson, "Internet of things security and forensics: Challenges and opportunities," 2018.

[6] S. Watson and A. Dehghantanha, "Digital forensics: the missing piece of the internet of things promise," *Computer Fraud & Security*, vol. 2016, no. 6, pp. 5–8, 2016.

[7] A. Alenezi, H. F. Atlam, R. Alsagri, M. O. Alassafi, and G. B. Wills, "Iot forensics: A state-of-the-art review, challenges and future directions."

[8] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of things for smart cities," *IEEE Internet of Things journal*, vol. 1, no. 1, pp. 22–32, 2014.

[9] H. F. Atlam, A. Alenezi, R. J. Walters, G. B. Wills, and J. Daniel, "Developing an adaptive risk-based access control model for the internet of things," in *2017 IEEE international conference on internet of things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and ieee smart data (SmartData)*. IEEE, 2017, pp. 655–661.

[10] M. McGuire and S. Dowling, "Cyber-crime: A review of the evidence research report 75, chapter 1: Cyber-dependant crimes," *Home Office*, pp. 1–27, 2013.

[11] S. R. Islam, D. Kwak, M. H. Kabir, M. Hossain, and K.-S. Kwak, "The internet of things for health care: a comprehensive survey," *IEEE Access*, vol. 3, pp. 678–708, 2015.

[12] A. Botta, W. De Donato, V. Persico, and A. Pescapé, "On the integration of cloud computing and internet of things," in *2014 International Conference on Future Internet of Things and Cloud*. IEEE, 2014, pp. 23–30.

[13] I. Yaqoob, I. A. T. Hashem, A. Ahmed, S. A. Kazmi, and C. S. Hong, "Internet of things forensics: Recent advances, taxonomy, requirements, and open challenges," *Future Generation Computer Systems*, vol. 92, pp. 265–275, 2019.

[14] D. Quick and K.-K. R. Choo, "Digital forensic intelligence: Data subsets and open source intelligence (dfint+ osint): A timely and cohesive mix," *Future Generation Computer Systems*, vol. 78, pp. 558–567, 2018.

[15] S. Alabdulsalam, K. Schaefer, T. Kechadi, and N.-A. Le-Khac, "Internet of things forensics–challenges and a case study," in *IFIP International Conference on Digital Forensics*. Springer, 2018, pp. 35–48.

[16] P. Dirac, "The lorentz transformation and absolute time," *Physica*, vol. 19, no. 1-–12, pp. 888–896, 1953.

[17] M. Imani, A. Q. Moghadam, N. Zarif, O. Noshiri, K. Faramarzi, H. Arabnia, and M. Joudaki, "A comprehensive survey on addressing methods in the internet of things," *arXiv preprint arXiv:1807.02173*, 2018.

[18] A. A. Majeed, "Global management of the internet of things," Ph.D. dissertation, University of Surrey, 2018.

[19] A. Majeed and A. Al-Yasiri, "Formulating a global identifier based on actor relationship for the internet of things," in *Interoperability, Safety and Security in IoT*. Springer, 2016, pp. 79–91.

[20] D. A. Peled, "Formal methods," in *Handbook of Software Engineering*. Springer, 2019, pp. 193–222.

[21] M. C. K. Michel and M. C. King, "Towards an adaptable system-based classification design for cyber identity," in *2018 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*. IEEE, 2018, pp. 1–2.

[22] E. Alamoudi, R. Mehmood, A. Albeshri, and T. Gojobori, "A survey of methods and tools for large-scale dna mixture profiling," in *Smart Infrastructure and Applications*. Springer, 2020, pp. 217–248.

[23] J. Jenkins, "Dna: What's your story?" *Journal of the American Association of Nurse Practitioners*, vol. 31, no. 10, pp. 555–557, 2019.

[24] B. Bruijns, R. Tiggelaar, and H. Gardeniers, "Massively parallel sequencing techniques for forensics: A review," *Electrophoresis*, vol. 39, no. 21, pp. 2642–2654, 2018.