Snap Forensics: A Tradeoff between Ephemeral Intelligence and Persistent Evidence Collection

Yijun Yu
The Open University, UK

Thein Than Tun
The Open University, UK

ABSTRACT

Digital evidence needs to be made persistent so that it can be used later. For citizen forensics, sometimes intelligence cannot or should not be made persistent forever. In this position paper, we propose a form of *snap forensics* by defining an elastic duration of evidence/intelligence validity. Explicitly declaring such a duration could unify the treatment of both ephemeral intelligence and persistent evidence towards more flexible storage to satisfy privacy requirements.

CCS CONCEPTS

• **Security and privacy → Privacy protections**;

KEYWORDS

digital forensics, privacy requirements

ACM Reference format:
Yijun Yu and Thein Than Tun. 2017. Snap Forensics: A Tradeoff between Ephemeral Intelligence and Persistent Evidence Collection. In *Proceedings of 1st International Workshop on Software Engineering and Digital Forensics, Paderborn, Germany, September 4, 2017 (SERF '17),* 2 pages.
https://doi.org/10.1145/3121252.3121255

1 INTRODUCTION

Given wide-spread surveillance such as CCTV, every citizen is empowered to collect forensic intelligence in some digital form, for witness or for self-defence. Generally a forensics evidence require persistence whenever it is needed later, which is not always the case for citizens to maximally protect their privacy, i.e., their rights to be forgotten. Here we propose to model the duration of evidence explicitly so that one can differentiate persistent evidence from ephemeral intelligence. The persistent evidence starts on the date when it was collected, and never ends; the ephemeral intelligence, on the other hand, would disappear after it has served the purpose. There are also two types of persistent evidence, logged and live evidence: the duration between when an evidence was created and when it was collected must be sufficiently short for live forensics.

Notice that for practical reasons the duration of persistent evidence is not forever either. It may be necessary to change how long the evidence is stored. E.g., for regulatory compliance, it may need to be stored for 7 years and can then be deleted, but if it becomes involved in a criminal prosecution, it may need to be stored until the full sentence is served or all appeals have been concluded – which could mean 30 years or more.

SERF '17, September 4, 2017, Paderborn, Germany

ACM ISBN 978-1-4503-5156-0/17/09…$15.00
https://doi.org/10.1145/3121252.3121255

2 MOTIVATION

In Feb 2017, UK National Rails announced a public consultation plan to introducing biometric authentications such as fingerprints or iris scan, as part of digital transformation to replace the current paper ticket-based system at the turnstile gates [8]. For passengers, the proposed changes could make the travel more efficient and enhance user experience; for business stakeholders, the conductor is no longer needed in the new approach to verify tickets on the trains, which could achieve "driver-only trains".

However, fingerprints or iris scans are typical biometrics that our pioneer forensic scientists such as Edmond Locard has proposed the *exchange principle*, according to which the perpetrator of a crime will bring something into the crime scene and leave with something from it, and that both can be used as forensic evidence. The difference is that, in Locard's world, there is always physical contact between the objects. In our digital world, there need not be physical contact - but we choose to record something about a person or object without making contact. His principle does not always hold in the digital world. Given biometric data could carry personal information that could uniquely identify the owner, it is in our view quiet risky for users to give them away to non-authorities.

These motivate us to consider tighter control of ephemeral intelligence, even if they were used for legitimate authentication purposes for the case of biometric scan, or end of lost from unrecoverable persistent storage. Another motivating example is the missing MH370 which suggests the use of a live forensic recorder [9] of ephemeral intelligence of flight data that otherwise may not be recovered as persistent evidence from the physical FDR.

3 LIVE FORENSICS VERSUS SNAP FORENSICS

A more general problem is inspired from the motivating example. According to traditional forensic science, evidence about an incident at the crime scene, according to Locard's Exchange Principle, is trace left found. However, when the incident happens, one may not be able to find the traces any more. For example, in network forensics, the network packets come and go, they may not be kept all the time. Runtime collection of such *ephemeral* evidence requires one to prepare them proactively [3, 4], hence the concept of "forensic readiness" - having organisations and systems prepared to capture and analyse potential evidence.

Previously we have proposed to monitor the environment of the system [7] to diagnose problems and switch to the right solutions [5]. The notion of keeping the log of software systems to diagnosing their failures could be extended to digital forensic evidence collection: to collect and use valid evidence, one needs to

associate each item with a life cycle of validity. We need to keep three time stamps minimal to signify the creation t_0, observation t_1, and the modification t_2 events associated with the evidence, while $t \geq t_0$ is the time of assessing the evidence.

Definition 3.1 (Persistent, Live, and Snap intelligence). Given two finite numbers τ_1 and τ_2, an evidence is *persistent* if $t_2 = \infty$; an evidence is *live* if $t_1 - t_0 < \tau_1$; an evidence is *snap* if $t_2 - t_1 < \tau_2$.

As discussed earlier, the acquisition τ_1 and retention τ_2 durations may be adaptable and changed to meet dynamic needs of satisfying forensic requirements. It is important that forensic evidence is not changed during $[t_1, t_2)$. If $t_2 = \infty$, the evidence will last forever, or *fully persistent*. Of course that is ideal, but as long as $t_2 > t$, at the time of investigation the evidence is valid. It is relatively easy to obtain persistent evidence from persistent storage (disks) with respect to file systems; however, it is much harder to guarantee persistence for network packets because without keeping them on to persistent storage the network packet will get lost.

The *liveness* of evidence is guaranteed by a small τ_1. As long as it is small, within short period of the incident, the data describing the incident is already received so that the evidence collection is on the fly, and the forensic analysis is in time to catch problems earlier, and to avoid disasters. In the MH370 incident the missing Boeing 777 is never found, suggesting that evidence needs to be collected proactively. Live blackboxes [9] were proposed to keep track of the aircraft in motion so that their flight data records could be retrieved in real-time (with less than 1 minute delay), to avoid similar incidents like MH370 by reporting evidence of potential incidents to the authorities earlier. The passwords stored by software managers are another form of ephemeral intelligence, where the master password in transient memory may be revealed by unexpected failure of the software system. To reduce the risk of exposing the master passwords, the in-memory plain text passwords should never last long [1]. Live forensic evidence could be combined with persistence storage in order to preserve them in the long run, while typically the storages are remote from the data generators.

Finally, some evidence concerning personal information requires a small τ_2 in order to reduce the risks of leaking personal information of users. One example is SnapChat, whose selling point is to destroy the photos of users within 24 hours. The day-return train ticket, e.g., is another good example whereby the biometric data as evidence of purchase could only last for at most one day.

4 CALL FOR ACTION

Saving evidence as persistent data allows someone re-purposing them without users' consent, violating privacy requirements of users; yet as a trade off, one needs to make the ephemeral intelligence last till being used as evidence for proactive or reactive investigations. Inspired by the success of SnapChat, we define *snap forensics*: instead of making ephemeral evidence fully persistent since t_1, we need to make the privacy sensitive evidence (e.g. biometric) transient again as early as possible, with $\tau_2 = t_2 - t_1$ small. In case one cannot predict how long evidence needs to exist in order for it to be used in an investigation, the incidents which require investigation may not be discovered within the short time periods proposed. Therefore, for investigative purposes, t_2 should be as large as possible. Storing the features of fingerprints in blockchains, a prototype implementation FingerBlox [1] seems to be the right direction to go. It highlights how a fingerprint could be scanned using an Android app and stored as featured points (rather than the raw fingerprints) into the distributed ledgers. The fact that blockchains are tamper-proof helps achieve both persistence and snap forensic requirements, while keeping users privacy into account as well. Due to the slow proof-of-work of blockchain technology, the live forensic requirements may not be accommodated easily though. Another promising biometric authentication solution is called *continuous authentication* [2], where by the biometric pulses of fingers are augmented to the initial measure to incrementally improve the accuracy of authentication.

5 CONCLUSION

By introducing time-stamps based definitions, we differentiate persistence, live, and snap intelligence as temporal forensic requirements, in relation to various security and privacy requirements for biometric authentication, sensitive data protection (e.g., passwords and social relationships). Using the recent examples we demonstrate that live forensics may be inevitable and useful for tracking Internet of Things whose states change continuously. We also show that snap forensics may be made more privacy-friendly by a careful design of the mechanism. Currently we are working with UK Police to implement part of the solution on *social-lift.com*, a platform to support snap forensics by limiting the verifiable and selective disclosure of users' personal information on social media [6].

Acknowledgement

The authors would like to thank Angus Marshall for his valuable comments as forensic science expert.

REFERENCES

[1] Joshua Gray, Virginia N. L. Franqueira, and Yijun Yu. 2016. Forensically-Sound Analysis of Security Risks of Using Local Password Managers. In *24th IEEE International Requirements Engineering Conference, RE 2016, Beijing, China, September 12-16, 2016*. IEEE, 114–121. https://doi.org/10.1109/REW.2016.034

[2] Ivan Martinovic, Kasper Rasmussen, Marc Roeschlin, and Gene Tsudik. 2017. Authentication Using Pulse-response Biometrics. *Commun. ACM* 60, 2 (Jan. 2017), 108–115. https://doi.org/10.1145/3023359

[3] Liliana Pasquale, Sorren Hanvey, Mark Mcgloin, and Bashar Nuseibeh. 2016. Adaptive evidence collection in the cloud using attack scenarios. *Computers & Security* 59 (2016), 236–254. https://doi.org/10.1016/j.cose.2016.03.001

[4] Liliana Pasquale, Yijun Yu, Mazeiar Salehie, Luca Cavallaro, Thein Than Tun, and Bashar Nuseibeh. 2013. Requirements-driven adaptive digital forensics. In *21st IEEE International Requirements Engineering Conference, RE 2013, Rio de Janeiro-RJ, Brazil, July 15-19, 2013*. IEEE Computer Society, 340–341. https://doi.org/10.1109/RE.2013.6636745

[5] Mohammed Salifu, Yijun Yu, Arosha K. Bandara, and Bashar Nuseibeh. 2012. Analysing monitoring and switching problems for adaptive systems. *Journal of Systems and Software* 85, 12 (2012), 2829–2839. https://doi.org/10.1016/j.jss.2012.07.062

[6] Thein Tun, Blaine Price, Arosha Bandara, Yijun Yu, and Bashar Nuseibeh. 2016. Verifiable Limited Disclosure: Reporting and Handling Digital Evidence in Police Investigations. In *iRENIC: 1st International Workshop on Requirements Engineering for Investigating and Countering Crime*. http://oro.open.ac.uk/46914/

[7] Yiqiao Wang, Sheila A. McIlraith, Yijun Yu, and John Mylopoulos. 2009. Monitoring and diagnosing software requirements. *Autom. Softw. Eng.* 16, 1 (2009), 3–35. https://doi.org/10.1007/s10515-008-0042-8

[8] Andy Weir. [n. d.]. UK rail network considers using iris and fingerprint scans as part of digital transformation. ([n. d.]).

[9] Yijun Yu, Mu Yang, and Bashar Nuseibeh. 2017. *Live Blackboxes: Requirements for Tracking and Verifying Aircraft in Motion*. American Institute of Aeronautics and Astronautics. https://doi.org/doi:10.2514/6.2017-0884

[1] https://github.com/jorenham/fingerblox