CYBERSECURITY

CASE LIBRARY

VOLUME I

CAL POLY California Cybersecurity Institute

ABOUT THE CCI

The California Cybersecurity Institute (CCI) is a robust, multi-agency effort to protect California through enhanced cybercrime forensics and state-wide tactical response training. Our partners are the California National Guard and Cal Poly, San Luis Obispo.

As an extension of Cal Poly's Cybersecurity Center, the CCI aims to educate the next generation cyber workforce and provide faculty and students with a new, hands-on research and learning environment.

The novel partnership between academia, industry and government offers immediate opportunities to local law enforcement, military personnel and even Cal Poly students to help California better protect its citizens.

The CCI serves as an extended Learn by Doing space for Cal Poly students, where they can explore new cyber technologies and train and test tactics side by side with law enforcement professionals and cyber forensics experts. The program helps shape California's cyber standards and practices by offering an environment for cyber defense innovation through advanced study and basic and applied research on emerging issues and technical challenges.

Learn more at cci.calpoly.edu

CAL POLY California Cybersecurity Institute

Cybersecurity Case Library

Volume 1, Issue 1

Mission Statement

The Cybersecurity Case Library is an undergraduate research journal that explores contemporary topics in cybersecurity. We have assembled students from multiple disciplines to co-write case studies regarding specific cyber events with the purpose of educating students campus-wide about cybersecurity. As the cybersecurity field continues to change and expand, our goal is to provide students and faculty with a class resource that will foster a deeper understanding of issues in cybersecurity and their real-world consequences.

Much of the information available concerning cybersecurity is highly technical and inaccessible to groups that do not have specialist knowledge. Through our case studies, we will convey the same information in a way that is instead approachable to anyone who picks up this journal. We value diversity of thought and perspective, and we are committed to presenting accurate and up-to-date information. The case studies in this journal were put through a review process that included academics both in and outside of each case study's discipline in order to ensure accuracy and readability. The Cybersecurity Case Library is associated with the Cybersecurity Center at California Polytechnic State University, San Luis Obispo but is intended to reach audiences in and beyond our campus.

CAL POLY California Cybersecurity Institute

Cybersecurity Case Library

Volume 1, Issue 1

Production Team

Advising & Editorial Board

Note: As an interdisciplinary, collaborative journal *CSCL* has standardized the organization of authors and editorial board members by sorting names in alphabetical order.

CAL POLY California Cybersecurity Institute

Table of Contents

Cybersecurity Case Library
Volume 1, Issue 1

An Examination of Posthumous Harm and Privacy

Nathan Lemay, Tristan Noack

Abstract

This case study examines Apple's security policy and the various authentication methods for iPhones. The two primary ways of unlocking an iPhone are via a remembered secret, such as a PIN or password, and Apple's fingerprint scanner technology, Touch ID. The current authentication methods in place have rendered Apple unable to unlock iPhones, and in this case study we examine a situation in which Leonardo Fabbretti was unable to access the iPhone of his late son. Fabbretti petitioned Apple for assistance, but because their policy is not to unlock devices, as a result of implementing technology that makes it impossible for them to do so, the phone was never unlocked. We examine how the authentication technologies work and the way that these authentication technologies led to certain policies. We consider the moral implications of the effect that these technologies and policies had in this case, as well as in a general context. The moral considerations we focus on include user privacy and the possibility of posthumous harm.

Introduction

Computer technologies have become more portable and widely distributed in recent years. More than three-quarters of Americans own a smartphone and many people carry a considerable amount of personal data on these small, easy-to-lose devices. With this in mind, cybersecurity has become an increasingly relevant consideration for everyone. Cybersecurity is the protection of computer systems from intrusion. This intrusion can manifest in many ways, but in our case study we will be focusing on the protection of one's data from unwanted access. In this paper, we consider whether children or deceased individuals have a right to privacy of their data. Our specific case deals with a father who asked Apple to unlock his late son's iPhone; he was denied because giving him access would have violated Apple's security policy. We examine Apple's encryption and authentication technologies, which are the technologies through which a user may access a device. We demonstrate that Apple's encryption and authentication technologies were made with the intention that not even Apple would be able to unlock a user's locked device. Therefore, the only security policy that could emerge from the implementation of these technologies is one that disallowed intervention in cases such as ours. We also consider the effects that the security policy has on Apple's users. These effects will be understood in terms of how they interact with user privacy and whether they cause harm to users.

In order to analyze this case, we offer several conceptions of privacy and arguments for and against the existence of posthumous harm. After examining these moral dimensions as they relate to this specific case, and other similar,

hypothetical cases, we conclude that while Apple's policy may have yielded unfortunate results, the technologies that led to the creation of this security policy are good because they effectively protect personal data. We hold that the possible negative outcomes of this case and similar cases are not sufficient to call for changing the security policy, which would require changing the relevant technologies to become less secure.

Background

Leonardo Fabbretti's son, Dama, was thirteen years-old when he passed away due to bone cancer in 2013 (Goldman, 2016). Dama was given an iPhone 6 before he passed away and had given his father access to the device through Apple's fingerprint scanner, Touch ID. Touch ID allows a user to access their phone by placing their fingerprint that they have registered with the phone on the Home Button. After Dama's passing, Fabbretti wished to gain access to his son's phone to retrieve any "photos, thoughts, and words which are held hostage" (Goldman, 2016). However, Fabretti could not unlock the phone despite having previously received access with his own fingerprint through Touch ID. In certain cases, such as restarting the phone, Apple requires a user to enter their password to unlock their device rather than using Touch ID. Since Dama's phone had been powered off and then back on, only a valid password could unlock it. Apple's decision to require a password in certain situations is based on their privacy policy and security implementations. Due to Apple's policy and technological decisions, it was impossible for them to access a powered off iPhone without the device's password. This effectively prevented Apple, or Fabretti, from unlocking Dama's phone.

There are a few major technological components that were initially added to the iPhone 5S that have continued into later models, such as Dama's iPhone 6. The first of these is Apple's Touch ID. Fingerprints are one of the oldest biometrics used for personal identification and the most prevalent biometric today (Ashbourn, 2014). Biometrics are defined as any measurable physical characteristics of an individual that can be checked and verified. Fingerprints are believed to be a good biometric because they remain both relatively identical throughout the duration of an individual's life and are unique to that individual (Gold, 2013). These attributes suggest that fingerprints may appear to be a viable substitute for a strong password; fingerprints can provide uniqueness and more complexity than traditional passwords can achieve (Ratha, 2001).

However, fingerprints have an inherent flaw: we leave them on everything we touch. If fingerprints were used as the sole authentication for iPhones, anyone with access to an individual's fingerprint could unlock their device. Unfortunately, it has been shown that lifting fingerprints from a surface and creating a fake fingerprint to unlock a device is relatively trivial (Bhutani & Bhutani, 2013). Despite these pitfalls to fingerprint scanners, recent modifications to Touch ID have made fingerprint replication more difficult; as a result, it is likely to only be exploited by a particularly motivated individual (Apple, 2015). These new developments suggest that Touch ID can still realistically serve as a strong form of authentication, despite concerns surrounding the vulnerabilities of fingerprint biometrics (Byrne, 2014).

While fingerprints may be sufficient to authenticate an individual, Apple did not intend for Touch ID to replace providing a password in the first place. Apple states that the purpose of Touch ID "[overcome] the inconvenience

of a passcode-based lock, not by replacing it but by securely providing access to the device within thoughtful boundaries and time constraints" (Apple, 2015, p.7). In an attempt to preserve the convenience of using Touch ID, as well as minimize the security risks, Apple has placed restrictions on when Touch ID may be used. There are five circumstances in which Apple deems the use of Touch ID alone to be an insufficient form of authentication:

1. The device has just been turned on or restarted.
2. The device has not been unlocked for more than 48 hours.
3. The device has received a remote lock command.
4. After five unsuccessful attempts to match a fingerprint.
5. When setting up or enrolling new fingers with Touch ID (Apple, 2015, p.7).

In this particular case, Dama's phone had been shut off and turned back on, making the device only accessible by entering the device's password. This is a result of Apple's decision to create an overall more secure device, but it has had unforeseen consequences on individuals who are unaware of how or why Touch ID works the way it does.

Apple's Touch ID is the focus of this case, but other underlying systems are inseparable from this discussion. The Secure Enclave is a processor/coprocessor that is separate from the main processor in an iPhone and provides all cryptographic operations needed by the device (Apple, 2015). Cryptography is defined as any code or cipher system that attempts to hide messages or data from prying eyes. After device data has been fed through a set of cryptographic operations, the resulting data is then encrypted. Ideally, this data should be impossible to access without knowing the secret key that was used as part of the encryption process. The Secure Enclave is responsible for encrypting a user's data until it can be decrypted by a valid form of authentication. Without correct authentication, the device cannot be decrypted and the data is inaccessible. Since Dama was the only individual who knew the password to his device, and Touch ID could not serve as a substitute, Dama's data will remain encrypted.

Apple's decision to restrict Touch ID's utility under certain criteria is based on the strength and convenience of fingerprint biometrics balanced with the fact that we leave fingerprints everywhere. These restrictions were chosen to provide the convenience of biometrics while maintaining the encryption on an individual's device. However, based on the technical limitations of Touch ID, Apple's policy decisions have resulted in users being unaware of potential side effects of these policies. Users are uninformed of policies concerning Touch ID and Secure Enclave technologies and are unexpectedly unable to access their devices. Should Apple design their technologies differently to allow individuals access to their devices given special circumstances? Is the added security more beneficial than the harm individuals face when they lose access to their data?

Process

In order to understand Apple's decision and its security benefits, it is necessary to examine the technical process behind the Secure Enclave and Touch ID. Secure Enclave's sole responsibility is to encrypt and decrypt user data to keep it safe, yet accessible. Secure Enclave achieves this by utilizing a symmetric key cryptography standard known as the Advanced Encryption Standard (AES) (Apple, 2015). AES is widely accepted as the standard for symmetric key cryptography and is used worldwide by both industries and governments. AES works by creating a single key that can both encrypt and decrypt data. Apple uses AES encryption keys that are 256 bits in length, which means that there are possible values for this key (Apple, 2015). This level of key strength makes it extremely difficult to guess or brute force the correct key. If you started guessing keys at the beginning of time with modern computers, you still would not have guessed even half of the possible values. While this level of encryption ensures that no individual with malicious intentions can access your data, it also prevents you from accessing your own data if you lose your encryption key. Since Apple utilizes passwords and fingerprints to generate encryption keys, losing or forgetting your password is synonymous to losing the ability to decrypt your device data.

While one level of encryption may appear sufficient, Apple employs multiple levels of encryption depending on the type of data being encrypted and whether or not Touch ID is enabled. Each individual file on an iPhone has an encryption key associated with it, known intuitively as a "per-file key" (Apple, 2015). Whenever a new file is created, a 256-bit AES key is generated and given to the Secure Enclave so the file can be encrypted. However, varying levels of importance are placed on different types of data and constantly decrypting data slows down the performance of the device. Therefore, it is important to classify data based on how sensitive it is. This need led to the creation of Apple's Data Protection Classes. There are four data classifications that Apple uses to encrypt data: Complete Protection, Protected Unless Open, Protected Until First User Authentication, and No Protection (Apple, 2015). These classifications of data protection provide another level of encryption that protects the per-file key and determines when the Secure Enclave should release the per-file key so device data can be unencrypted.

Data protection classes are important to understand because they directly affect when and how data may be accessed. The data that Fabbretti wished to retrieve from Dama's phone include Dama's photos and his words. This kind of data falls under the "Protected Until First User Authentication" data protection class, which protects "Calendar, Contacts, Reminders, Notes, Messages, and Photos" (Apple, 2015). Data that falls under this classification requires a valid password to be provided at least once to decrypt it; then the data is left unencrypted until the device is powered off. When the device is powered off, the decrypted data protection keys are thrown away by the Secure Enclave, leaving the data encrypted. After the device is powered off, Touch ID is no longer accepted as a valid means of authentication and cannot be used to decrypt the data protection class keys that were protecting the per-file keys.

From a technical standpoint, we are now able to understand how Secure Enclave encrypts and decrypts data. But how does Touch ID fit in? When Touch ID is enabled, the data protection class keys are wrapped in another AES 256-bit key that the Secure Enclave gives to the Touch ID system (Apple, 2015). If a user-provided fingerprint

matches one stored on the device, Touch ID provides this new encryption key to the Secure Enclave so that the data protection class keys may be decrypted. When the device is powered off or there have been five unsuccessful attempts to unlock the device via Touch ID, these extra keys are discarded and the device can no longer be unlocked by Touch ID (Apple, 2015). Touch ID ultimately fulfills the same duty in decrypting data that a traditional password does but it adds an extra step and an extra encryption key. Despite these subtle differences, the Secure Enclave handles passwords and fingerprints in very similar ways. This suggests that Touch ID is not used as a primary form of authentication due to Apple's policy surrounding Touch ID's use and purpose rather than any technological limitations.

We have now shown that passwords and fingerprints undergo the same process to unlock an iPhone, but the question remains, why is Touch ID not allowed to perform the same authentication as a password? Ultimately, Apple's security policy was not driven by technical limitations of the Secure Enclave but rather Apple's intention to not allow Touch ID to replace traditional passwords. The power of Touch ID is offset by the fact that fingerprints are left everywhere. Although the Secure Enclave processes fingerprints in a very similar manner to passwords, this inherent flaw provides enough ground for Apple to disallow Touch ID to unlock a device under certain circumstances. Taking an in depth look at Apple's security provides justification that Touch ID and passwords are meant to perform their duties cooperatively and that there are tradeoffs for using either as a form of primary authentication. Apple's current security policy does not allow for the reasonable retrieval and decryption of data if a password is lost or forgotten and Touch ID is unable to unlock the device.

Significance

In this case, privacy rights and the possibility of posthumous harm are central issues. A few questions have emerged from the Fabretti case: if Apple overrode the password authentication, would that constitute a violation of Dama's privacy even though he is deceased? Is Apple's security policy too strict and potentially dangerous?

Before investigating the issues of this case, we must first understand the intimate connection between Apple's intentions with Touch ID technology and their security policy. Apple's goal with their authentication technology was to create a system that even the company would be unable to override. The authentication technology at work here then, is overwhelmingly compatible with the policy that states that they will not override the authentication. Because they cannot override this authentication, the technology confirms the policy is in effect. Arguments about how the policy could be changed are merely hypothetical because the technology demands a certain policy—a policy of nonintervention (Cook, 2016). If one were to argue that Apple ought to change their policy to allow for intervention in certain cases, then they must also argue that Apple ought to change the technology to be amenable to that policy. This would make the authentication method less secure. One could also argue that the technology at hand allowed for the creation of a justified policy or that the technology ought to be made less secure to allow for a more flexible policy. Whichever position one chooses to take, they must consider privacy and harm in their argument—both of which are central issues in this case.

Prior to analyzing privacy and harm, the Dama situation could have been avoided had the user understood Apple's security. There was ample time for Dama to give his father access. While that counterfactual situation is plausible,

it is also plausible to make two considerations: (i) a very similar situation could exist where the users understand the security technology at work and their corresponding policies, but still believe that the policies caused them harm and the policies should be changed, which would require a change in technology; and (ii) a situation where the users are unable to prepare for an event like this (i.e., an individual dies in a car crash and would have wanted their data to be accessible to their family members). In this case, it does not matter whether or not the individual understands the technology or policy if they did not have their affairs in order. In response to this, one could argue that a person is always at fault for not understanding technologies and policies relevant to their own affairs; and if they do, they are at fault when their affairs are not in order. Even if we believe it is the user's responsibility to understand the technologies and policies, we can still consider how it is that the policy affects the user.

While privacy may seem like a simple concept, there is significant debate over what privacy actually means. In this paper, we are going to offer a few different approaches to understanding privacy before addressing certain factors specific to the Fabretti case. There are four different views of privacy: **Privacy is Noninterference, Privacy is Control, Privacy is Derivative,** and **Privacy is Convention.** We will see if any of these views capture most individuals' understanding of privacy.

First, the **Privacy is Noninterference** view suggests that privacy is "the right to be let alone" (Warren & Brandeis, 1890, p.195). This conception of privacy holds that any acts of interference constitute an invasion of privacy. This view is rife with problems. Voyeuristic acts are invasions of privacy, but voyeurs do not interfere directly with those they watch. Let us try to add some nuance.

The **Privacy is Control** view suggests that privacy is "control over when and by whom the various parts of us can be sensed by others" (Parker, 1973, p.281). On the face of it, this view seems very similar and may even just be a reinterpretation of "the right to be let alone." This view incorporates all things you sense in its consideration of the right to privacy. This also has some issues as it is too broad. For example, by allowing a security company to install cameras in your house, you are forgoing your control that they will not spy on you; however, while the cameras are on and no one is looking at the feed, your right to privacy is not violated, regardless that you have relinquished your control over whether or not you are seen.

The view that **Privacy is Derivative** holds that privacy itself is a secondary right, a right derived from others (Thomson, 1975). This means that an individual has the right to privacy by virtue of having other rights; an act violates not only their right to privacy but also another right. Therefore, every violation of an individual's right to privacy depends on the violation of another right. For example, if an individual's sensitive documents are looked at without consent, this constitutes a violation of the rights they have over their property (i.e., the right to control who can look at certain documentation). Simultaneously, it violates their right to privacy. The main concern when addressing privacy violations then becomes the more fundamental right that is violated prior to the secondary violation of their privacy.

Finally, the view that **Privacy is Convention** assumes that any understanding of the right to privacy is conventional, but privacy can still be appealed to as a right. It rejects reducing privacy to a secondary right and instead holds that

there are social norms and legal conventions under which one can understand violations of privacy. Regardless of appeals to ownership, one can argue that if an individual picks up and looks through the phone sitting on the desk in the library that you are currently using, your right to privacy is violated. You may not own the phone, but there are shared social norms suggesting that a violation has occurred. Whether or not your right to privacy has been violated does not depend on whether or not you own the phone.

Here, we provide a distinction between the views of **Privacy as Convention** and **Privacy as Derivative**. The privacy violation does not come from a primary violation of a separate right—the right over things that one owns—but it is still evident. While this view may seem unphilosophical and cumbersome due to its cultural and temporal fragility, it does have some merit in establishing a common understanding. Given these distinctive views on privacy, we will now discuss posthumous harm.

Considering the rights of deceased individuals is tricky. We cannot interact with deceased individuals and any of their wishes made prior to their death are their own. There are two opposing views on deceased individuals: the view that no posthumous harm can be done to individuals versus the view that posthumous harm is conceivable. The **Common Sense** view of death is the one most frequently held. This view says that an individual cannot be harmed post-death because they no longer exists. Even given religious conceptions of death and the afterlife, many still hold that any harm done to a person's body or slander against them is no longer a harm to that person, since they no longer exists. However, this is a harm done to their family or any others who survive them. With this understanding, it is still wrong to desecrate a body because you are harming not only the deceased person but also those who care for them.

The **Posthumous Harm** view holds that an individual can be harmed after death. Similar to all the positions that we have presented, this view can take several forms. One form holds that the inability to fulfill a rational desire constitutes a harm to the person. With this understanding of harm, we can see how a person may have projects that carry on after their death that may or may not be successful. If their project is unsuccessful (even after death), this could constitute a harm. If a person's legacy is tarnished, some may argue that they are harmed. This can be difficult to understand unless we believe an individual has interest in their reputation after death. For example, any of the horrible things that surface about famous people after death, or even things that are read in the journals of the deceased, may still tarnish their reputation. If everyone's conception of that person drastically changed in a negative way then one might argue—using the **Posthumous Harm** view—that the individual has been harmed.

To return to our case, would Dama be harmed if his father had been granted access to the phone? Furthermore, would that harm be due to a privacy violation or something entirely different? It is best to work backwards. From an initial look at the view advocating that there is posthumous harm, whether or not Dama was harmed depends on whether or not his rational desires would be disappointed. If he had a desire to keep certain things secret from his family, even after death, then it would be beneficial to him that his father was not given access to the phone. However, if he would have wanted his family to have access to his photos and messages, then it is arguable that Apple's policy has, in fact, harmed Dama. Given the facts as we know them, it is difficult to argue that granting

Dama's father access to the phone would have been an invasion of Dama's privacy. Dama tried to grant his father access to his phone through Touch ID. He may have thought that this allowed his father full access, but it did not. Given these considerations, we are left with four different outcomes that depend on whether posthumous harm is possible and whether Dama wanted his father to have access to the contents of his phone.

If posthumous harm is possible and (a or b):

(a) Dama wanted his father to have access to the contents of his phone, then Apple is harming Dama by the policy. Apple is not harming Dama by violating his privacy, but by disrespecting his wishes.
(b) Dama did not want his father to have access to the contents of the phone, then Apple is in the right and unlocking the phone would posthumously harm Dama by violating his privacy.

If posthumous harm is not possible (c or d):

(c) then Dama is not harmed even if he wanted his father to have access to his phone.
(d) and Dama did not want his father to have access, then his wishes were justifiably protected by virtue of Apple's policy; regardless, unlocking the phone would not have harmed Dama.

Both (a) and (b) suggest that Dama is harmed, but in different ways. By contrast, (c) and (d) suggest that Dama is not harmed, but (d) suggests that one can have justification for protecting the wishes of a deceased individual, even if not doing so would not cause harm to that individual. As we have stated, it seems that Dama did want his father to have access to the phone, so his privacy is not violated in either case. But he could be harmed if we accept posthumous harm as a possibility.

Each of these outcomes is isolated to the harm incurred by Dama in this specific case rather than the policy's implications as a whole. As we previously acknowledged, it seems reasonable to suggest that an adequate understanding of the systems could have prevented this case entirely. Given that this case is an instantiation of how users can misunderstand the effects of Apple's authentication technology and security policy, it is worth considering whether this case just happens to be one of misfortunate or if changes to the technology or policies ought to be made.

Recommendations

In providing recommendations regarding policy, we keep in mind how the policy affects everyone. Apple hoped to make data more secure by creating this authentication technology. One result includes Dama's father's inability to access his deceased son's phone. We think that it would be problematic to take isolated cases like this and use them in arguments in favor of leniency in security policy. As we explained, Apple's security technology necessitates a nonintervention policy. Given the technological infeasibility of Apple unlocking Dama's phone, if one argued that Apple should make their policy more lenient in cases like this then they must also argue that Apple should make their devices less secure. We simply cannot recommend such a course of action: these security measures are too important for personal security, and one unfortunate case does not merit a large enough step backward in user security.

One isolated incident, in which the security measures harmed someone who suffered a tragedy, does not stand up against the many negative consequences of switching to a less secure system. If we return to our analysis of Dama's specific case, we can compare the largest possible harms suffered in this case with the numerous likely harms that would be suffered in making iPhones less secure. And as we have just stated, the harms to Dama and his family would not be resolved, but future possible harms for similar incidents would be. Given our considerations, Apple's current policy is justified and no better policy exists. Apple's policy may cause certain people to feel unhappy or disappointed, but that is a consequence we must accept in the interest of existing users. We consider the heightened security of the rest of Apple's users to be of paramount importance.

Discussion Questions

1. Is the lifting of fingerprints really a significant enough concern for most people to disallow the use of fingerprint scanners as primary authentication?
2. Will we ever be precise enough to avoid issues of impersonation when it comes to biometrics on mobile devices?
3. Are there any interesting legal considerations for not allowing fingerprints to serve as a primary authentication?
4. What responsibilities do users have to understand the technologies they use and the policies associated with those technologies?
5. Is the act of allowing harms like those caused to Dama's family (and possibly Dama himself), in favor of maximizing security for most users, wrong? Has there been a greater harm suffered by the Fabretti family that makes you think Apple's policy is wrong?

References

Apple (2015). iOS security: iOS 10.

Ashbourn, J. (2000). Biometrics: Advanced identity verification. *Springer, 18.*

Bhutani, T., and Bhutani, B. (2013). No to fingerprint security system. *International Journal of Computer Science and Management Research*, 2(10), 3583-3587.

Byrne, B. (2014). iPhone 6 touch ID fooled by fake fingerprints. *ValueWalk.*

Cook, T. (2016). A message to our customers. *Apple.*

Gold, S. (2013). Meeting the biometrics payment security challenge. *Biometric Technology Today*, 2013(10), 5-8.

Goldman, D. (2016). Grieving father pleads with Apple to unlock his dead son's iPhone. *CNN Tech.*

Parker, R. (1973). A definition of privacy, *Rutgers Law Review*, *27*, 275.

Ratha, N., Connell, J. and Bolle, R. (2001) An analysis of minutiae matching strength, *AVBPA*. 2091. 223-228.

Thomson, J. (1975). The right to privacy. *Philosophy & Public Affairs*, *43* (1), 295-314.

Warren, S., and Brandeis, L. (1890). The right to privacy. *Harvard Law Review*, *4* (5), 193-220.

Author Biographies

Nathan Lemay is a Software Engineering student at California Polytechnic State University, San Luis Obispo. His current interests involve consumer security and privacy with a particular focus on vehicles. Nathan shares his passion for security and privacy through his involvement in Cal Poly's security and ethical hacking club, White Hat, by serving as an officer for four years. He is also the recipient of the 2016 Accenture Outstanding Junior Computer Science Award awarded based on faculty and department recommendations for academic excellence and community contributions.

Tristan Noack is a fourth-year Philosophy major at California Polytechnic State University in San Luis Obispo. His primary philosophical interests are: Comparative Aesthetics, Philosophy of Education, Evidentialism, Philosophy of the Internet, and Existentialism. He plans to seek a career in policy analysis before applying to graduate school for Philosophy, Public Policy, or Asian Art History.

Cybersecurity Case Library

Volume 1, Issue 1

Blockchain Databases and the Ongoing Drive to Secure Information

Grant Blake, Brian Kinnee

Abstract

This paper explains and explores blockchain databases, a category of information technology with important implications for cybersecurity professionals. Blockchain databases allow for new ways of storing and distributing data as well as timestamping digital transactions and using current generation network technologies. Because blockchains promise increased security to users and new levels of accounting for digital activities, the use of such databases has spread rapidly worldwide among both individuals and institutions. This paper begins with a simple explanation of blockchain database technology, which is followed by more complex examples. We then examine a real-world application that highlights the use of blockchain databases in the realm of healthcare and public health. Throughout the case study, we provide commentary on the significance of blockchain technology. This paper builds on longstanding academic conversations about the nature and function of databases (e.g., Bowker 2000, Manovich 2000; Waterton 2010, Hayles 2012) as well as on more recent conversations about the economics and culture of algorithms (e.g., Gillespie 2014; Mackenzie 2014; Bilic 2016). It also builds on critical literature on blockchains specifically (e.g., Levy 2017).The central contribution of this paper is to further raise awareness of blockchain databases and to highlight the links between blockchains and cybersecurity. The central contribution of the paper is to further raise awareness of blockchain databases and to highlight the links between blockchains and cybersecurity.

Introduction

How can we create, share, store, and retrieve information securely in a networked world? This question remains at the heart of contemporary discussions of cybersecurity. In this case study, we set out to explain blockchain databases, a category of recent information technology with important implications for cybersecurity professionals, policymakers, and scholars of technology and society.

Blockchain databases were first developed in the early 2000s to create a secure digital currency, or cryptocurrency, called Bitcoin. Bitcoin is a currency that was developed for online, virtual transactions. There are no physical "Bitcoins," nor is the currency backed by any larger governing body like the US Federal Reserve (Nakamoto, 2008). A key feature of Bitcoin is decentralization: unlike pre-digital currencies, no single person, government, or regulatory body oversees Bitcoin's supply and production. Bitcoin transactions and the records of such transactions

are not concentrated in one particular place, institutionally backed, or managed by intermediaries. Another key feature of Bitcoin is the use of blockchain databases to record every transaction on the Bitcoin network.

In a broad sense, a blockchain database is a large, distributed list or ledger (Lewis, 2016a). Computers from all over the world can use the internet to access the ledger, read entries on it, and add new entries to it. Once an entry is added, it cannot be removed by anyone and will remain in the ledger. This unique feature of blockchain technology makes it an extremely useful tool for verifying and documenting transactions while also guaranteeing that such transactions will not be modified after the fact.

Let's consider a simple analogy: an exchange between an officer and a driver. The officer may seek to verify our driver's identity, proof of registration, and proof of insurance. The driver, in return, may try to verify the officer's badge number, the justification for making the traffic stop, and that the details of the stop have been fairly and accurately recorded. In this analogy, the two parties rely on official markers of identification to aid in the exchange, but the details of the stop and who was involved can be disputed later, in traffic court for instance. What the two parties lack is a way of authenticating the alternate party's identity and the particulars of the stop. They also lack a way of preventing the other party from altering information about the exchange. This is what blockchain databases attempt to do: authenticate the particulars and make them unalterable.

Moving Deeper into Blockchain

Now let's consider a more complex exchange: a record of payments between bank accounts, recorded in chronological order. Each of the entries within the record will contain the bank account numbers of the sender and receiver, the amount being sent, a signature, and a date. The immutability of this record of payments is the advantage of blockchain: the technology prevents someone from modifying or duplicating payments after being recorded in a blockchain database.

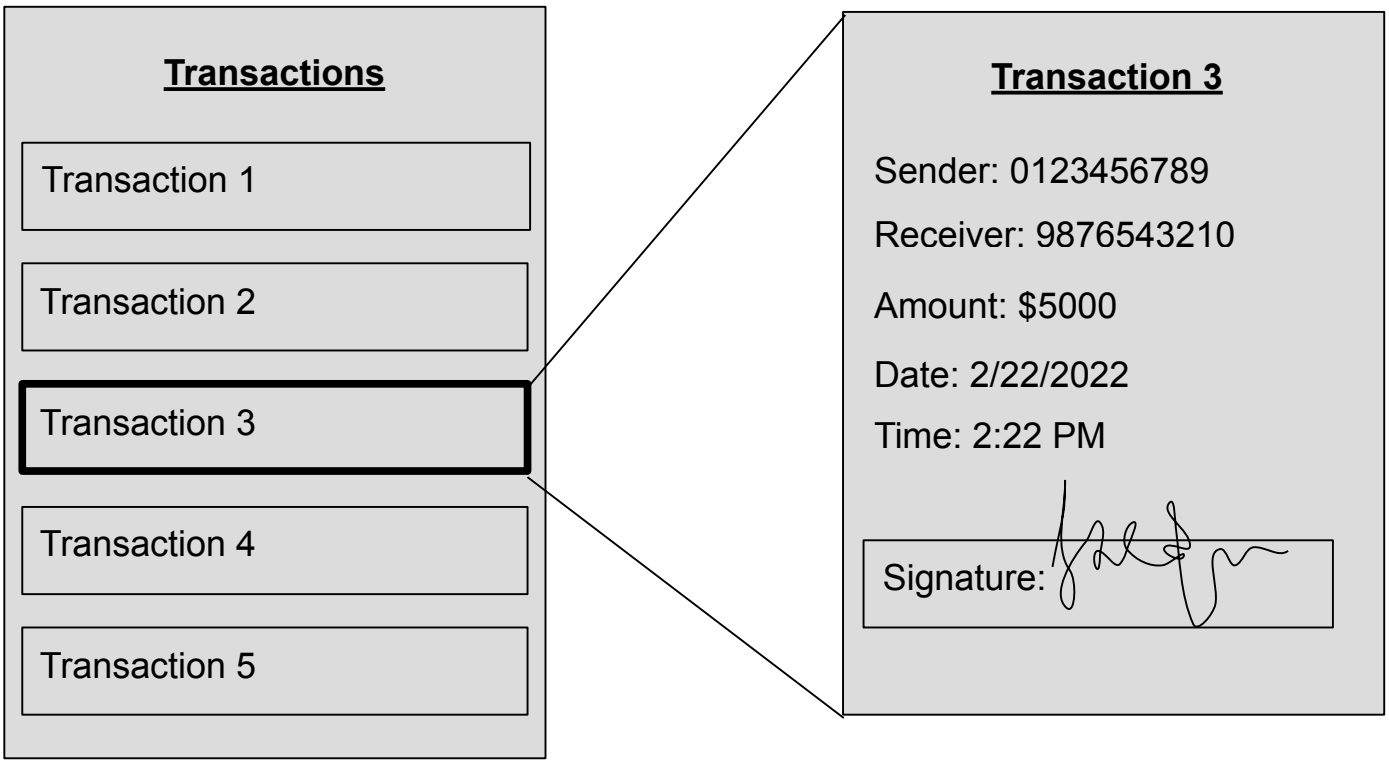
A blockchain secures its data by making it extremely difficult to change an entry after it has been added to the list. With a traditional list, the owner has the power to revise entries at will. For example, any user of a web browser user may delete his or her internet history; likewise, paper ledgers can have transactions erased or altered by hand. In contrast, the users of a blockchain do not want anybody to have the power to modify the list of transactions. To enforce this property, each user of the blockchain is given an up-to-date copy of the entire list. When a new entry is added, every user is notified. This makes it increasingly challenging for one user to modify an entry, as all the other users will have matching lists that dispute any malicious, after-the-fact modifications.

In terms of how they work, blockchain databases utilize signatures. The signature in a transaction has multiple purposes. The first is to verify that the sender is actually who she or he claims to be, much like a handwritten signature (Nakamoto, 2008). If a user tries to add a transaction to the list or ledger, but has a strange-looking signature compared to the account owner's usual signature, the other users of the system will notice and reject the proposed addition. The second purpose of the signature is to verify that all transactions are being added to the list in the exact order that they occur (Lewis, 2016a). For example, a user that tries to add a transaction with a strange-looking date (too far into the future or past) will quickly get denied by the other users of the blockchain.

In reality, the signatures in a blockchain are much more important and complex than a handwritten signature. They are generated and later verified by strong cryptographic algorithms using information from every transaction. The signature of a transaction is used to both verify the sender's identity and to keep the list of all transactions in chronological order.

In our banking example, the users of the blockchain database are not only the people making transactions but also those that verify transactions and maintain the list. In a more general blockchain, all people who access the list to read from it or write to it are considered users. Since blockchains are digital ledgers, they are accessed and operated through networked computers that can communicate directly with each other. Some computers connected to the network are responsible for verifying and timestamping new transactions. These computers are often called nodes or miners. Other computers broadcast new transactions to all nodes connected to the network.

The transactions recorded in a blockchain can represent more than just an exchange of currency. Almost any digital asset, including digital media, digital agreements called "smart contracts," and even personal records and information can be securely exchanged and documented using blockchain databases (Lewis, 2016b).



In general, the information broadcasted to the nodes is not the full details of the transaction, but rather digital proof of the transaction, called a digital signature (Brakeville, & Perepa, 2016). Digital signatures are produced using a cryptographic technique known as public-key cryptography, a scheme in which all parties possess a pair of passwords: one publicly known and one kept private (Lewis, 2016a). These passwords, also called keys, are usually hundreds of randomly generated characters, making them virtually impossible to guess in an efficient manner. The user creating a transaction uses his or her private key to "sign" the details of the transaction. Other users connected to the network can verify that this transaction was sent by the correct person using their public

key. Using a key that is not the exact pair to the private key that generated the signature will quickly let other users know that the signature is invalid. The downside to using digital signatures, though, is that users must deal with extremely long, random passwords and find ways to manage their public and private keys.

Nodes connected to the blockchain network "listen" constantly for new transactions being broadcasted. When a transaction is received, the node verifies it and places it into a bundle with hundreds of other transactions called a "block" (Lewis, 2016a). This bundling is a complicated process and involves solving a set of mathematical equations used commonly by the entire blockchain network. The solutions to these equations are usually extremely long, random numbers, so it takes individual nodes a few minutes to bundle up transactions into a block. Once the block is successfully created, it is broadcasted to all other nodes in the network, which verify each transaction one last time. As more and more nodes connect to the network, the computational work required to solve these problems uses a substantial amount of electricity. This is one aspect that groups must think about before implementing a system that utilizes blockchain technology. Bitcoin solves this problem by rewarding Bitcoin to each node that finishes a block. Blockchain databases, like other information technologies and infrastructures like data centers, can be resource-intensive solutions to questions of information security that produce real world environmental effects.

Information from previous entries in the list or ledger are used to secure new entries, so previous entries get more secure as the blockchain grows. In the original discourse surrounding Bitcoin, they claimed: "To modify a past block, an attacker would have to redo the proof-of-work of the block and all blocks after it and then catch up with and surpass the work of the honest nodes" (Nakamoto, 2008). Blockchain databases, in this regard, are said to produce extremely secure ledgers that can be trusted by users connecting from all over the world to engage in various forms of exchange and transaction. The technology has begun to transform financial transactions in profound ways, and it has found use in a wide variety of other contexts.

Blockchain in Healthcare

Because blockchain databases promise increased security to users and new levels of accounting for digital activities, the use of such databases has spread rapidly worldwide to include disparate kinds of individual and institutional users (Mettler, 2016) . One lively site of current experimentation with blockchain database technologies is in the realm of healthcare. The U.S. Department of Homeland Security includes healthcare and public health within its 16 critical infrastructure sectors (Underwood, 2016). Between storing and distributing data, timestamping transactions, and accounting for digital activities, healthcare and public health have become heavily networked and are a key area of concern for cybersecurity professionals both in the U.S. and internationally.

The recent work happening in Estonia around electronic health records offers an illustration of blockchain use beyond the financial sector (Gault, 2016). In early 2016, the Estonian Government implemented a blockchain-based system to store the health records of its citizens. "The system contains information on diagnoses, doctor's visits, tests, hospital treatments, medications prescribed by a doctor, etc." (E-health- Estonian digital solutions for Europe, 2016). Any Estonian citizen can gain access to medical information by confirming his or her identity with a government-issued ID card. This ID card functions as each user's "password" to the blockchain.

Documenting healthcare is notoriously complex. Healthcare systems must record, maintain, and secure huge sets of highly sensitive data about patients. In the United States, one person's medical records can be scattered among multiple doctors. It is a complicated process to access them (Azairia, Ekblaw & Vieira, 2016). Meanwhile, Estonia's healthcare system is now fully linked and interoperable. Citizens and doctors alike can access and exchange records between any two networked devices in Estonia with the aid of blockchain database technology. Since the health information is publicly distributed through blockchain, Estonian citizens have the power to make sure that all information kept on them is accurate and up-to-date.

The Broader Significance of Blockchain Database Technologies

Blockchain database technologies offer a new approach to information security and the documentation of digital activities. Blockchain technology could theoretically be used in more areas beyond the realms of finance and healthcare (Yli-Huumo, Ko, Choi, Park, & Smolander, 2016). One potential application of blockchain technology, and a particularly provocative one, involves solidifying the democratic processes. In democratic elections, voters must trust that their votes will not be changed before they are counted. The task of quickly counting large volumes of votes securely is challenging. In addition, networked voting machines are not invulnerable to cyberattacks. Blockchain technology could help to mitigate such vulnerabilities because each citizen's vote could be recorded on a public list and made extremely difficult to change after submission, with the aid of digital signatures and keys. Using blockchains in this manner offers the potential to integrate new tools into democratic processes, to create new patterns of civic participation, and to reimagine the role of databases in public life.

As social phenomena, blockchain database technologies have broad significant and heterogeneous consequences. For cybersecurity professionals, these technologies represent some of the latest developments within the longer history to advance the knowledge and practice of encryption (Gault, 2016). Such databases are also beginning to introduce new kinds of public expectations and standards into the cybersecurity equation. For policymakers, blockchain database technologies raise important unanswered questions about how to regulate new kinds of virtual transactions. Some might say that blockchain database technologies pose a threat to traditional institutions and actors (e.g., banks, medical records managers, election supervisors). For scholars of technology and society, blockchain database technologies intersect directly with enduring conversations about the nature and function of databases (e.g., Bowker 2000, Manovich 200, Waterton 2010, Hayles 2012) and more recent conversations about the economics and culture of algorithms (e.g., Gillespie 2014; Mackenzie 2014; Bilic 2016). The central contribution of this paper has been to raise further awareness of blockchain databases and to highlight the links between blockchains and cybersecurity.

In Blockchain: Blueprint for a New Economy—a book dedicated entirely to potential future applications of the blockchain—Melanie Swan suggests that sometimes data is managed best without blockchain technology. For example, Swan points out problems that do not require "sequential, public, and distributed data storage" can usually be solved by using other technologies such as cloud storage or distributed computing models (Swan, 2015, p.68). Expanding on this point, blockchain database technologies are certainly not appropriate for every information security need; but for those wanting to authenticate the particulars of virtual transactions and make them unalterable, blockchain database technologies are quickly becoming a key tool. The goal of this paper has been to raise further awareness of these technologies and to highlight their links to cybersecurity.

Discussion Questions

1. When is trust vital and where is blockchain now?

An important feature of blockchain technology is to manage problems of trust in a networked world, along with the high volume of digital transactions taking place worldwide. What kinds of online transactions and virtual exchanges matter most to you? Do you know if blockchain databases have become materially integrated into them?

2. What if blockchain technology became more widespread?

A wide-scale implementation of blockchain database technologies could have significant benefits to societies worldwide but could also render earlier information security practices and also some kinds of cybersecurity professionals obsolete. How might the widespread implementation of blockchain technology benefit and/or harm the cybersecurity professions?

3. What are the broader economic implications of blockchain?

If blockchain databases come to replace long-established institutions (such as banks), then the technology has the potential to upend certain fields and industries, along with their associated workforces. What are the broader economic implications of blockchain and is the technology a new democratic form of automation and outsourcing?

References

Azaria, A., Ekblaw, A., Vieira, T., & Lippman, A. (2016). MedRec: Using Blockchain for Medical Data Access and Permission Management. *2016 2nd International Conference on Open and Big Data (OBD)*. doi:10.1109/obd.2016.11

Bilić, P. (2016). Search algorithms, hidden labour and information control. *Big Data & Society*, *3*(1), 1-9. doi:10.1177/2053951716652159

Bowker, G. C. (2000). Biodiversity Datadiversity. *Social Studies of Science*, *30*(5), 643-683. doi:10.1177/030631200030005001

Brakeville, S., & Perepa, B. (2016). Blockchain basics: Introduction to distributed ledgers. International Business Machine.

E-health- Estonian digital solutions for Europe. (2016). e-Estonia.com.

Gault, M. (2016). *Estonian government innovation and the future of data.* Guardtime Blog & News.

Gillespie, T. "The Relevance of Algorithms." *Media Technologies: Essays on Communication, Materiality, and Society.* Cambridge: MIT, 2014. N. pag. Print.

Hayles, N. Katherine. *How We Think: Digital Media and Contemporary Technogenesis.* Chicago, IL: U of Chicago, 2012. Print.

Lewis, A. (2016a). *A gentle introduction to blockchain technology*. Bits on Blocks.

Lewis, A. (2016b). *A gentle introduction to smart contracts*. Bits on Blocks.

Levy, Karen E. C. "Book-Smart, Not Street-Smart: Blockchain-Based Smart Contracts and The Social Workings of Law." *Engaging Science, Technology, and Society* 3 (2017): 1. Web.

MacKenzie D. (2014) A sociology of algorithms: High-frequency trading and the shaping of markets. Available at: www.maxpo.eu/Downloads/Paper_DonaldMacKenzie. pdf

Manovich, L. "Database as a Genre of New Media." *AI & Society* 14.2 (2000): 176-83. Web.

Mettler, M. (2016). Blockchain technology in healthcare: The revolution starts here. *2016 IEEE 18th International Conference on e-Health Networking, Applications, and Services (Healthcom)*. IEEE.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Bitcoin*.

Swan, M. (2015). *Blockchain*. Sebastopol, CA: O'Reilly Media, Inc.

Underwood, S. (2016). Blockchain beyond Bitcoin. *Communications of the ACM, Volume 59 Issue 11*.

Waterton, C. "Experimenting with the Archive: STS-ers As Analysts and Co-constructors of Databases and Other Archival Forms." *Science, Technology & Human Values* 35.5 (2010): 645-76. Web.

Yli-Huumo J., Ko D., Choi S., Park S., Smolander K. (2016). Where is current research on blockchain technology?—A systematic review. PLoS ONE 11(10).

Author Biographies

Grant Blake is a Computer Engineering and Mathematics double major at Cal Poly. His work focuses primarily on secure computing and computational mathematics. After graduating, Blake aspires to earn an advanced degree in one of these fields of study. Trained as both a mathematician and an engineer, he strives to contribute to future research in secure computing technologies.

Brian Kinnee is a student in the Science, Technology, and Society (STS) program at Cal Poly. Much of his current work investigates the sociocultural impacts of information technologies. Kinnee also writes about healthcare technologies, human-computer interaction, and design. This is Kinnee's first publication within a larger series on recent developments in computing.

Cybersecurity Case Library

Volume 1, Issue 1

Mitigating Cybersecurity Threats in Autonomous Vehicles

Alexander Carrasquillo, Naomi Palmer, Abineet Singh

Abstract

Technology and data are rapidly expanding into every avenue in our modern lives. This is occurring with our vehicles as well, especially with the expansion of autonomous or self-driving cars. This paper considers policies to minimize the risk of malicious exploitation of self-driving vehicles. Exploitation is unapproved access to systems exposed by vulnerabilities found by hackers in the hardware and software design of these vehicles. These exploitations primarily occur on the networks of vehicles such as infotainment systems, which can be used to take over a vehicle's CAN (Controller Area Network). While the U.S. Department of Transportation (DOT) has developed policies in regards to thorough testing of autonomous vehicles (AV), there has been no assessment of the legal ramifications of failure to meet these standards and whether or not adherence to those standards are in fact successful in improving security. There is very little data present on the security of self driving vehicles. Reasons for this could be because self-driving cars are not fully developed or mainstreamed, or because no reported malicious hacks have taken place so far. Since autonomous vehicles are still in development there currently is not a clear-cut solution in regards to cybersecurity policy for these systems. However, we have identified commonalities between the DOT's proposed AV policy and the Federal Drug Administration's (FDA) current drug approval process; this has given us a basis to analyze and propose recommendations for future AV policies.

Introduction

Autonomous vehicles utilize software in conjunction with hardware such as cameras, RADAR, and LIDAR to adapt to surrounding traffic conditions. Popular examples of AV platforms include Google's Waymo Project, Tesla's fleet of Autopilot Vehicles, and Uber's new self-driving initiatives. This technology can reduce accidents and increase driver safety overall, but unforeseen security challenges will likely emerge. AVs are essentially classified into five levels of autonomy: these levels have features that increasingly reduce aspects of human control such as parking and adaptive cruise control at levels one and two; full safety feature control at levels three to four; and adaptation to changing road environments, unique to level five autonomy (DOT, & National Highway Traffic Safety Administration (NHSTA), 2016). The development of this new technology is associated with a high degree of software overhead. According to the DOT (2016), a modern vehicle already has around 100 million lines of software code. Presumably, AVs will require at minimum this much code, and possibly even more. According to Steve McConnell (2004), author of Code Complete, error density (number of errors per thousand lines of code) scales up with project size (p. 652). More errors imply additional points of vulnerability upon which a cybersecurity attack can take place. A cyber-attack occurs when a hacker attempts to access information from a

device, disrupt functionality of a technology, or damages how a particular technology works (Hathaway, Crootof, Levit, Nix, Nowlan, Perdue, & Spiegel, 2012). One such notorious cyber-attack known as Ransomware could emerge in AVs. This is where hackers use extortion by breaching data or injecting malware to lock users out of their computer system until a ransom is paid (Boatman, 2016.). This type of attack could possibly threaten users while driving or be used to completely disable a vehicle until demands are fulfilled. This is just one example of the emerging possibilities in car hacks, but fortunately precautions are emerging to mitigate such risks.

Process

Usually the primary objective of an attack on a vehicle is to exploit a vulnerability in order to obtain access to the Controller Area Network (CAN) bus—an internal communications network that enables devices and systems within a vehicle to communicate with each other via packets or electrical signals (Paganini, 2016). A CAN performs on every device in a vehicle ranging from the brakes and steering column to windshield wipers and speakers via microcontrollers. What makes this area so vulnerable is the lack of native security or encryption. That is, it assumes that any incoming signal is from a valid source. A vulnerable point of entry into a CAN is located at the On Board Diagnostics Port (OBDII) that is located under a vehicle's dash, as this port is used for diagnostics and repairs. Once a vehicle is compromised its CAN can be manipulated to perform any action in a car as hackers can sniff (detect) specific packets and, in return, spoof or alter them on various vehicle modules (Paganini, 2016).

This attack was performed by Chinese White Hat hackers from Tencent Laboratories who breached the computer systems of the Tesla Model S by manipulating the system's CAN bus in order to take control of everything from the side mirrors, sunroof, and door locks to completely overriding the computer console of the vehicle (Akan, 2016). It should be noted that the Tencent researchers spent several months learning about the workings of many Tesla cars, and the attack required that the car have a web browser open while connected to a malicious WIFI network (Akan, 2016). Without all of these conditions, the Chinese White Hat hackers may not have been able to succeed in their attack. Further, this is the first reported instance in which the CAN bus on a Tesla Model S has been breached.

This however was not the only incident where a vehicle's CAN bus has been compromised. The highly controversial Chrysler hack was the first true red flag when it comes to cybersecurity in vehicles. This notorious hack happened when 2014 Jeep Cherokees were breached by two security engineers: Charlie Miller and Chris Valasek. The engineers found vulnerabilities within the vehicle's Uconnect Infotainment System and, once entered, they proceeded to send spoofed commands to the vehicle's CAN bus. The hackers were able to remotely control the stereo, radio, brakes, and windshield wipers while the car was driving on a highway. Their findings were posted on Wired and Youtube, which sparked a massive outcry causing three Jeep owners to file class action lawsuits against Chrysler and Harman International, the developers of Uconnect (Bisson, 2015).

In order to begin addressing these potential cybersecurity threats, especially in wake of the infamous Chrysler hack, manufacturers collaborated to develop the Automotive Information Sharing and Analysis Center (Auto

ISAC) in July 2015. The primary goal of Auto ISAC is to establish "An industry-operated environment created to enhance cybersecurity awareness and coordination across the global automotive industry" (Auto ISAC, 2016). They seek to establish the best cybersecurity practices in order to protect data and assets, as well as establish general standards for properly identifying, assessing and managing cybersecurity risks (Auto ISAC, 2016).

Significance

When it comes to autonomous vehicles, there has been no record of hacking outside the scope of research for testing purposes. Thus, many will wonder why we deem it necessary for companies to pour so much of their time and resources into the cybersecurity division. After all, the Chinese Tesla Hack would have been very difficult to execute without the exact conditions that made it possible. It takes someone with an advanced technical background and a strong understanding of the underlying vulnerabilities to execute an attack.

AV manufacturers should be concerned about hackers with sufficient knowledge in long-range wireless communications and who may be capable of performing simultaneous mass system attacks from a remote location. For example, hackers could target the telematics unit, which handles mobile communications such as for GPS navigation (Papp, 2016). If the hackers manage to exploit the telematics unit in a way that grants access to critical systems such as engine or brake control, there would then be a means to cause several vehicles to crash. These types of dangers are precisely why legislators should discuss fair policies regarding the security of AVs.

Resolution

As the reality of self-driving cars reaching the market approaches, many discussions over cybersecurity and other safety concerns are coming to the forefront. Self-driving cars are not a new phenomenon. For years, companies like Google and Tesla have been testing these vehicles on the streets. However, companies are beginning to exit testing phases and are moving toward mass deployment for consumers. For example, in Pittsburgh, the transportation company Uber has recently released a new autonomous vehicle system that is available to users (Davies, 2016). This growing use by laypeople opens increased opportunities for exploitation by hackers and is why policy is an important component for resolving AV concerns, especially those pertaining to cybersecurity.

Within the past few months, the DOT has put into place a policy regarding autonomous vehicles and the requirements companies must meet in order to have their vehicles on the road. This new regulation outlines specific concerns, identified by the DOT, that manufacturers should address before vehicles are on the roads. For the purposes of this paper, we are mostly concerned with the regulations regarding cybersecurity (DOT, & NHTSA, 2016).

Within the autonomous vehicle policy there are several sections, one of which focuses specifically on cybersecurity within AVs. The main focus of this section revolves around requiring manufacturers to conduct thorough risk assessments to test the integrity of their systems (i.e., vehicles). Within the policy, risk assessment is identified as reducing the risk to safety. Manufacturers expect this process to occur while in product development, as well as after products have been manufactured to ensure that any issues with the system are identified and managed in a timely manner. The shortfall of this policy, identified by the authors, is that the DOT's policy is not currently

written to deal with regulatory standards for manufacturers. Recall, the main reason regulatory standards have not been set is due to the evolving nature of this topic, and lack of research in cybersecurity. Before any standards are set, research must be conducted in order for the DOT to have something to base the standards off of (DOT, & NHTSA, 2016). The current policy addresses the concern over cybersecurity and vulnerabilities of code; however, it does not lay out how cybersecurity attacks will be addressed from a legal perspective, which leaves the issue of how to address bugs that are found via hacking.

Recommendations

The cybersecurity section of the DOT's AV policy is a work in progress; the policy is relatively undeveloped and has an air of uncertainty because of the speculative inferences being made. However, other policies exist today that might shed light onto possible expectations of and concerns for this autonomous vehicle policy in the coming years.

The most relatable policy identified is the FDA's regulations of drugs, which drug manufacturers are meant to meet before their drugs reach the marketplace. For FDA purposes, manufacturers conduct clinical trials to test how safe and reliable their drugs are, and any initial concerns that come up in these trials are self-reported (FDA, 2016a.). The FDA's evaluation of regulating drugs bares similarities to the DOT's policy on AVs: both require that manufacturers self-report any issues they find with their products. Another similarity between these two policies is the type of testing required. The FDA requires manufacturers to conduct clinical trials to test their product, and, similarly, the DOT requires companies to do extensive trials to test the integrity of their products (DOT, & NHTSA, 2016; U.S. FDA, 2016a.). Furthermore, both of these federal agencies have similar missions and goals driven by national interest that promote public health (DOT, 2015; FDA, 2016b.). This further adds to why the FDA's policy is a good choice for the DOT to take into account when considering future changes to its AV policy. Addressing the similarities between these two agencies' policies is important because the FDA's policy has been in effect much longer than the DOT's, which provides potential insight into flaws and strengths of the AV policy that the DOT and manufacturers alike may experience.

The FDA has been criticized over the years about its methods of drug approvals; some critics claim the process is inefficient and slow, while others say it is too quick of a turnaround for proper drug approval. This long debate has led to many policy reforms within the FDA (Henderson & Rowland, 2005). The constant dual criticism of the FDA's drug approval process affects the administration's credibility and is something the DOT can use as an example when it comes to their own policy. The mechanics and software that go into the system are complex, and the integrity of the system must be maintained at all times to ensure safety for users. That being said, we believe that the DOT should continue to add to its policy in the years to come, and, as more research is devoted to this topic, the DOT should adjust their policy accordingly. A big next step for the DOT should be to set up regulatory standards that manufacturers must meet in order to have a stronger system of accountability and to keep potential risks at a minimum. Keeping in mind the criticisms that the FDA has faced, such as the perceived inefficiency of its regulations and the changing of policy to satisfy critics, the DOT should support its regulations with empirical evidence provided by researchers in order to have a logical backing to its choices. Organizations such as the Auto

ISAC can play a pivotal role in shedding light to the DOT in regards to best practice standards for cybersecurity, especially if they work from the ground up with supply chains to find vulnerabilities.

Conclusion

Even though self-driving vehicles have not quite reached their optimal state of development, companies like Tesla in particular are pushing out massive updates, strengthening hardware and increasing features for their autonomous systems (Kang, 2016). With the emergence of these updates and evolving levels in autonomy, not only must the consumer be aware of the possible risks that come with using such advanced technology, but our lawmakers need to adapt as well. This new technology is emerging in an environment that has not completely adapted to this change, so policy and safety regarding cybersecurity is critical, especially since policy has a tendency of lagging behind newer technology. As the FDA is flexible in its policy on drug testing and development, so too should legislation written by these lawmakers. This assessment aims to provide security for the consumer while promoting technological growth of AVs. Initiatives taken by organizations like Auto ISAC can aid the DOT and NHTSA in creating policies that can adapt along with the growth in AV technology, and lawmakers would be wise to follow suit.

Discussion Questions

1. As emergence in AVs grow what assessments should consumers make before committing to buying and driving such vehicles?
2. What are some possible strategies that legislation may need to consider upon while creating security laws, without stifling the expansion and development of AV technology?
3. How can organizations like the DOT, NHTSA and Auto ISAC collaborate to create policies that can find a sufficient balance?

References

Akan, E. (2016). Tesla car hacked by chinese tech company. *Epoch Times.*

Automotive Information Sharing and Analysis Center. (2016). *Automotive cybersecurity best practices.*

Bisson, D. (2015). Inside a connected car's points of vulnerability. *Tripwire.*

Boatman, Kim. (2016). Beware the rise of ransomware. *Norton*.

Davies, A. (2016). Here's what it's like to ride In Uber's self-driving car. *Wired.*

Francillon, A., Danev, B., & Capkun, S. (2010). Relay attacks on passive keyless entry and start systems in modern cars (Unpublished doctoral dissertation). ETH Zurich, Switzerland.

Greenberg, A. (2016). Radio attack lets hackers steal 24 different car models. *Wired.*

Greenough, J. (2016). 10 million self-driving cars will be on the road by 2020. *Business Insider.*

Hathaway, O., Crootof, R., Levitz, P., Nix, H., Nowlan, A., Perdue, W., & Spiegel. (2012). The law of cyber-attack. *California Law Review, 100*(4), 817–885.

Henderson, D., & Rowland, C. (2005). Once 'too slow,' FDA approvals called 'too fast'. *Boston Globe*

Kang, C. (2016). Self-driving cars gain powerful ally: The government. *The New York Times.*

McConnell, S. (2004). Code complete: A practical handbook of software construction. *Microsoft Press*.

Paganini, P. (2014). Car hacking: You cannot have safety without security. *INFOSEC Institute*.

U.S. Department of Transportation. (2015). *About us.*

U.S. Department of Transportation, & National Highway Traffic Safety Administration. (2016). *Federal automated vehicles policy: Accelerating the next revolution in roadway safety* (p. 1-116).

U.S. Food and Drug Administration (2016a). *How FDA evaluates regulated products: Drugs.*

U.S. Food and Drug Administration. (2016b). *What we do.*

U.S. Government Accountability Office (2016). *Vehicle cybersecurity: DOT and industry have efforts under way, but DOT needs to define its role in responding to a real-world attack.*

Author Biographies

Alex Carrasquillo is a 4th year Electrical Engineering transfer student with a keen interest in low-level programming and security. He is a computer science and physics tutor for the Cal Poly Multicultural Engineering Program (MEP).

Naomi Palmer is currently pursuing a B.A. in Political Science, and double minoring in Statistics and Ethics, Public Policy, Science and Technology (EPPST). She plans on graduating in Fall 2017, and then continuing her education towards attaining a Master's of Public Policy (MPP) degree. She hopes her Political Science degree combined with her minors in Statistics and EPPST will provide her with the relevant knowledge to excel in her pursuit of an MPP degree.

Abineet Singh is a 4th year Software Engineering transfer student and an aspiring White Hat Security Engineer. His ideal career is to be a Security or Privacy Project Manager as he want to be able to fulfill a role in industry where he can work with engineers to implement security policies and act as an intermediary between technical and client side roles. He feels CSCL is a fantastic opportunity for students interested similar career paths as he is. He will apply what he learned from this project to future projects and research, especially since he is hoping to work with more Autonomous Vehicle technology in the future.

Cybersecurity Case Library
Volume 1, Issue 1

On Ethics and Doxing

Nick Gonella, Lorenzo Nericcio

Abstract

The spread of information technology has created a new era of ethical problems. Specifically, the purpose of the paper is to explore one such moral problem resulting from doxing, the act of revealing someone's private information on a public Internet site (Douglas, 2016). Doxing is not a new practice; it has been around since the beginning of the Internet. However, modern cases will allow us to better understand the full extent of the issues surrounding doxing. In the course of the paper, we will present the history of doxing and give accounts of two modern cases: ex-InterActive Corporation (IAC) Director of Corporate Communications, Justine Sacco, and members of the Ku Klux Klan (KKK). Next, we will explore the moral implications of each case. We will draw conclusions on these moral problems using two ethical theories: consequentialism and deontology. Each theory will provide principled grounds for discussion, as well as recommendations toward improving the technological state of affairs to avoid future breaches of privacy.

Introduction

The spread of information technology has created a new era of ethical problems. Currently a major issue on the public agenda is online privacy, specifically the moral implications of breaches of privacy. As the public nature of the Internet increases, so does the impact of privacy loss. The new phenomenon of doxing, that is, the sharing of private information about a person in a public online setting, is a pertinent example of such a public breach (Douglas, 2016). This case study will focus on the situation surrounding Justine Sacco, a former Director of Corporate Communications at IAC, and Operation Hoods Off, the doxing of KKK members. We will first focus on the technological issues in the security systems that were breached, which allowed personal information to be released to the public. Second, we will consider the ethical problems surrounding these violations of privacy.

These ethical problems will be examined using two ethical theories to deliver rigorous, principled results on the nature of the moral problems caused by doxing. So, the treatment will proceed: (I) we will investigate the state of technological affairs that allowed the situation we are examining to occur, as well as explain the history and context for the moral issues; (II) we will detail the moral ramifications of these technological issues and identify all parties victim or responsible using the moral theories of deontology and consequentialism; (III) we will make suggestions toward the improvement of the technological state of affairs such that it better accommodates the results of our ethical analysis.

Narrative

In order to understand the social and ethical implications of doxing, one must first understand what doxing is. In this section, we will present a working definition of doxing relevant to this case study. We achieve this by examining the history of the community that created and propagated the action. Then, we will apply our definition to modern examples of doxing and explore the implications of the practice.

I. Rise of Usenet

Doxing originated from the Usenet boards of the early pre-Internet in the 1990's (Garber, 2014). In the years leading up to the Internet, collections of computers would connect to each other through several purpose-built intranets, or closed networks of computers (Liener et al., 2012). Usenet, however, differed from the other Intranets of the time, as it was not created for a specific organization, but rather to be used by anyone who had a computer and a subscription for a Usenet server. People would connect to specific servers, which had to be discovered by word of mouth, to download files and articles.

One of the biggest uses of the Usenet were news servers that were maintained by several key players in the creation of the Internet. Usenet was the first place doxing was employed (Giganews, 2016, Liener et al., 2012). Users would connect to centralized Usenet servers and download, read, or respond to any of the articles present. Servers could also share files, often ones that were either copyrighted or illegal to have, between cross-continental communities. The display of differing opinions in combination with the pseudo-anonymity of not speaking in person led to frequent instances of conflict on these news boards (Baker, 2001). Prolonged back and forth posts, also called flames, in the news groups, would frequently start up as a consequence of differing opinions. Unfortunately, these flames had a tendency to escalate to more and more "ad hominem" style arguments, often attacking the character of the people posting, rather than the content of the posts that started them. Most of the time, these conflicts would be relatively benign, even humorous, in how caught up people would get in their own arguments (DeLaney, 1994). However, these flames could also escalate to more serious consequences. Occasionally, users would attempt to drive out their opponents from the newsgroup by de-anonymizing them, often posting documents that show the true identity of the opposing poster. These actions were the first iteration of doxing, named after the "docs" which were posted to expose someone.

II. Doxing and Hacktivism

Doxing continued to be a facet of the Usenet community, and eventually Internet culture itself, but did not come into the full light of the public until the rise of another social phenomena: hacktivism. Hacktivism is the use of computer hacking to further political goals, often by a civilian group (McCormick, 2013). Hacktivist groups, like the infamous Anonymous, used several common computer hacking techniques to further their political ideologies. Among the tools used by Anonymous was doxing: the hacker would retrieve any documents on a target they could find and post them to the popular site WikiLeaks in order to attack an entity they found to be in the wrong. These attacks deviated from the original kind of doxing because the targets of these attacks were already public entities. This new kind of doxing, focused on making private information about a target public, rather than just exposing their identity. Often, the purpose of this information leak was to allow the public to further antagonize a target; thus shielding the attacking group from blame and furthering their agenda.

Doxing can then be defined as maliciously making information public that the victim party would prefer to be kept private. Many would imagine that this action would require either specialized access to information or the ability to execute some highly advanced hacking techniques. However, with just a cursory Google search, one can find several articles on websites that facilitate the process of doxing (Emeka, 2013). The process generally revolves around first doing some Googling to find two correlated pieces of information about the target, for example username and an actual name. From there, the websites mentioned earlier can cross-reference these pieces of data in order to find out more information about the target. Often, the databases used for doxing are free or require very little cost to use. From these websites, one can retrieve home addresses, phone numbers, and any other measures of sensitive data. This process is made easier with the rise of social media, as many people post massive amounts of personal data without thinking through the potential consequences with making it publicly available or otherwise not understanding how public data is (Ryan, 2014).

Introduction to Cases

I. An Unfortunate Tweet

For an example of modern doxing, one needs to look no further than the case of Justine Sacco. Ms. Sacco was a Director of Corporate Communications for IAC, the holding company OKCupid and Vimeo, when she became the subject of an Internet storm in 2013 (Ronson, 2015). On December 20th, Justine Sacco boarded an 11-hour plane ride from London, England to Cape Town, South Africa. Before she got on the plane, she sent a tweet to her 170 followers: "Going to Africa. Hope I don't get AIDS. Just kidding. I'm white!"(Ronson, 2015). She then proceeded to turn her phone off for the flight and fall asleep. Like many people, Justine misunderstood the technology she was using, specifically the lack of privacy involved in the tweets she sent. She assumed that her small circle of friends would see this message and then it would simply fade away.

However, what followed could only be described as an Internet mob. Through a series of retweets, tens of thousands of people saw her supposedly private tweet and began to display their outrage over the ignorant message. People further doxed Justine Sacco by finding her personal information, including her place of work. This resulted in IAC becoming involved and issuing a public statement disavowing the statements made by Ms. Sacco. IAC then fired Ms. Sacco on the spot before she could even land and turn her phone on to see what happened. Needless to say, Sacco was greeted by the awful surprise of a massively publicized, outraged Twitter feed and sudden unemployment.

Looking into the case of Justine Sacco brings up another important part of doxing and the Internet as a whole. In the words of Rooney Mara in the popular movie The Social Network, "The Internet's not written in pencil ... it's written in ink." (Rudin, Cean, De Luca, Brunetti, & Fincher, 2010). Once something is on the Internet it never really goes away as anyone can save and keep what was posted on their local computer. Additionally, services exist such as the Internet Archive Wayback Machine, which takes snapshots of every major webpage on the Internet and saves them for future viewing. It is estimated that over 470 billion web pages are part of the Wayback Machine, all available for free public viewing (Wayback Machine, 2016). It is likely that an incriminating comment will be preserved forever.

II. Anonymous and the Klan

Another case of doxing as hacktivism occurred in 2015 when Anonymous released the names of over 100 KKK members in Operation Hoods Off (Franceschi, 2015). To mount this attack, Anonymous used a technique known as social engineering, which involves using deception to bypass normally very technically difficult problems. In this case, the Anonymous members socially engineered their way into several Facebook groups that local KKK chapters were using to organize meetings. From this, the attackers were able to identify names and other personal details of people heavily involved in Klan activity. After gathering enough information, the Anonymous group collectively posted all of these details to the Internet for all to see. Many would classify this as a more traditional case of doxing, as the purpose of the action was to remove anonymity, rather than trying to expose private information about an already public figure.

After having looked at some examples of modern doxing, one can return to the definition of doxing and attempt to further understand it. Doxing ultimately comes down to a power imbalance, with the victim party becoming helpless to stop the release of their private information. In this sense, the modern doxing harkens back to its original purpose in the Usenet newsgroups. Doxing can be seen as a particularly effective and terrifying tool, since it leverages mob mentality to do most of the malicious work.

Significance

The moral implications of doxing are far-reaching and complicated; let us begin by introducing two moral theories we can use to unpack some of the moral wrongs committed in a doxing case. The two theories, roughly outlined here, will be consequentialism and deontology. A consequentialist determines the morality of an action in terms of the beneficial outcomes versus the detrimental outcomes, which can be defined in many ways. For these case studies we will suppose pleasure and individual freedom as our benefits, and pain and loss of freedom as our harms (Sinnott-Armstrong, 2015). A deontologist believes that individuals have strong duties or obligations—and violating them is morally wrong in and of itself, regardless of the consequences. We can assume that recognizing others' humanity, protecting privacy, and respecting others' freedom are strong obligations for a deontologist (Alexander, 2015). While the issues of privacy, individual liberty, and mob-justice that arise in doxing cases will doubtlessly concern both a deontologist and a consequentialist, the use of each theory will illuminate some of the more controversial issues in each case. Further it will open them to more productive discussion.

I. Justine Sacco

The case of Justine Sacco is a moral muddle and there is no doubt that people will disagree on whether she deserved what came to her, or whether her employer did the morally correct thing. First, let us consider what happened under a consequentialist framework and then move to deontology. We can start by identifying the moral agents at play in this scenario: Justine Sacco and the group of people who publicly shamed her. We will propose a hypothetical, mischievous Internet poster—whom we will name John—and who is a moral agent partially responsible for Sacco's fate.

A. Think about the consequences

First, it is clear that Justine Sacco's remark might have caused some pain. Pain, of course, may come in the form of psychological pain. A person of African descent, or someone deeply concerned with social justice, may have read her Tweet (recall, "Going to Africa. Hope I don't get AIDS. Just kidding. I'm white!") and been terribly distressed. Certainly, many people took this as a cruel and dismissive remark, treating both the horrors of epidemic AIDS and the history of white colonial oppression of African populations as a joke. Thus, we can say that Sacco bore some moral responsibility in this case: she was responsible for the psychological pain felt by the people who read and were distressed by her Tweet.

Next, we can consider John's responsibility in this ordeal. The responsibility borne by the people who shamed Sacco is likely to be great, as they caused her to be in immense amounts of psychological pain. This severely limited her individual freedom by ruining her reputation. In spreading the word about her Tweet, John may have caused for himself some small amount of pleasure, perhaps the pleasure of making a satirical joke and contributing to some perceived act of justice. To John, it surely might have seemed that Sacco deserved what was to befall her. Let us say that John tweeted, "White privilege at its finest—burn in Hell, Sacco," or something of the kind. Let us further say that he got a significant amount of retweets on this post, and thus felt some pleasure in his public approval. Given that it is not at all clear, or even plausible, that Sacco's shaming actually contributed to race equality movements or the distribution of medical resources to AIDS victims, we can restrict the moral benefits of this case to pleasure gained by posters such as John. Given that posters like John are what caused Sacco's demise, they are clearly responsible for the pain she ended up feeling. As such, we should possibly consider John as partially responsible for his contribution.

Considered under this consequentialist system, it seems that we have a very clear answer to our moral question: even if Sacco's Tweet was insensitive and thoughtless, the amount of pain she felt and the amount of freedom she lost outweighs the pleasure felt by her tormentors. This result allows us to conclude that the people arose from the fact that Sacco's doxing was a form of mob justice, something consequentialists are very wary of, as it nearly always leads to more suffering than happiness. Perhaps, then, it would be appropriate to put in place a firm rule against mob justice, in this case doxing. Rules in consequentialism, which are guidelines by which the good might be maximized, and duties in deontology differ. For our purpose this is a good point to move on to the deontological considerations.

B. No disrespect

Let us consider the deontological view now. It is possible that, in trivializing the suffering of millions of Africans, Sacco breached an absolute moral duty: not to disrespect and dehumanize people for the sake of a joke. If this is a universal obligation, then it is clear that Sacco has done something that is simply morally impermissible outright, as deontologists are less concerned with the consequences of the action so much as the character of the action in and of itself. Thus, even if no one ever read or was hurt by Sacco's tweet, her having posted it was morally impermissible. The act was not one that demonstrated respect for people's freedom and dignity, one could easily argue. Further, it did not act in accordance with a rule that she would wish to be a universal law—this is deontology's first categorical imperative, and a rule we will use to evaluate actions under deontology.

This, however, does not excuse the doxers. John, in making his Twitter post, also violated a moral duty: not to ruin people's personal lives for the sake of public approval or perceived moral justice. Note that John's perception of justice is particularly relevant in this case. Though it may be justifiable to exact revenge against an agent who has perpetrated a moral harm themselves, it is certainly not true, in Sacco's case—the punishment fits the crime. Consider as well that there is no reason to suppose doxers wish to see the ruining of a reputation, ending of career, and endangering of personal safety—as was the case for Sacco after being outed—as a punishment for an offensive tweet. Though John had ample cause to judge Sacco for her action, his vigilante punishment was itself a breach of moral code. Even if Sacco was morally suspect herself, this does not imply that she should not to be treated as a human being who has freedom and dignity. He and his fellows are all subject to moral judgment now—perhaps even to a greater extent than Sacco was. In either case, however, one has a moral duty not to disrespect the humanity of another for the sake of a dumb Twitter post.

II. Doxing the Klan

Historically, the Ku Klux Klan as an organization is morally responsible for a great number of terrible things (SPLC, 2017). Their continued existence is a continuation of the very same convictions that motivated them in the past: the belief that those of Northern European descent are biologically superior to those of other ethnic and racial backgrounds, and the belief that the structure of society should reflect this "biological fact." Since Klan is known for its atrocities and hate crimes, the moral case is not so clearly decidable. Perhaps, because of the KKK's history, the same act that we found to be immoral in Sacco's case might be morally justifiable when turned against the Klan.

A. Maximizing freedom?

Consider the following argument (MJ):

(1) Consequentialism obliges us to maximize the happiness and freedom of as many people as we can.
(2) The presence of the KKK causes many minority groups to feel unwanted and afraid in the United States, thus limiting their freedom.
(3) Exposing the identities of KKK members, in this case doxing them, will cause their public presence to diminish.
(4) So, doxing KKK members will lead to the increased freedom of those oppressed by the KKK.
(5) Allowing (4) is permissible only if it leads to greater overall pleasure with no resulting significant harm.
Therefore,
(6) One is morally obliged to dox members of the KKK.

This argument seems to be sound at first glance, but premise (4) may be problematic. Perhaps (4) is true, but we should hesitate to draw the conclusion expressed in (6). It seems that mob justice is not a particularly reliable form of justice. We might be uncomfortable living in a society where this is the accepted moral norm for punishment. Thus, there are good reasons to hesitate to accept (5), as there may be a whole host of morally suspect outcomes to result from doxing that are not been taken into consideration.

B. A general maxim?

Argument (MJ) above left us at an impasse: if exposing the identities of KKK members maximized freedom, surely it is the morally justifiable act—but this seems wrong. This is where a deontological view may become handy. While it may have been consequently beneficial to expose the identities of KKK members, we are left with the suspicious conclusion that doxers should be allowed to carry on as they please, so long as their targets are morally suspect themselves. We may express this in the following principle:

(DX) One, morally, ought to dox people with unsavory political views.

Is (DX) a violation of human dignity, or the freedom of others? Seemingly so. Surely many cases could arise wherein people disliked by the majority are doxed. This may result in the suppression of many legitimate but unpopular views, or even some bizarre mob-ruled censorship program. If we employ the categorical imperative, as we did before, we can see that this action is not an example of a rule that one could consistently wish to be universally followed—surely instances where it is unjust would occur. Thus, the deontological answer is that doxing the Klan was not justified.

Resolutions

We have seen in the previous sections that understanding the given cases through the lenses of different theoretical frameworks yields different results. This is not cause for frustration or quietism, however. Each theory yields insights into just exactly what is morally wrong about each case, and as a result, yields clear conclusions about what might be done. While we need not go too far into each ethical theory, we should note that they are well equipped to handle a variety of cases and outcomes, and in doing so, gives us better tools than our blind intuition to make moral conclusions about difficult cases. It seems that in every case and under every theory, doxing is nearly always wrong. What then should be done about doxing? These conclusions will be explored in the next section—especially the ways in which our ethical considerations reflect on the technological issues previously explored.

Recommendations

Now that we have understood the moral ramifications of doxing, we move to consider what policy changes ought to be enacted to better suit the conclusions of our ethical arguments. Let us assume that a good moral principle to emerge from our discussion above is the following.

(PF) Doxing person A is a violation of A's personal freedom; and, as such, one ought not dox A.

An obvious recommendation that follows from (PF) is: do not dox people. But this is not very useful on its own. We might add that moral responsibility extends to those who protect the data exposed in a successful doxing attempt. If security firm SF protects A's data, and R doxes A, we may argue that SF bears some responsibility in allowing the doxing to occur. Hence, it seems reasonable to suggest SF follow a basic set of guidelines when dealing with S's sensitive information. A few examples of such guidelines are:

(i) Establish a policy for what data is to be shared with certain parties and what data has to be kept secret. This should reflect the level of sensitivity of the data.
(ii) Work with the clients to make sure they understand what their responsibilities are to keep the data they give you secure, e.g., using modern standards (AES-256/ RSA-4096) to encrypt all data sent between the client and SF.
(iii) Build a security first system for storage of sensitive data. This means designing a system with security in mind, rather than the more common practice of designing a system and then trying to secure it after the fact.
(iv) Responsibly disclose when a breach of trust has occurred to the client and work to properly rectify the situation in any way which fault fell upon the SF. There is good reason to believe that SF has an obligation to inform a client should their information be disclosed to unwanted parties.

Thus, were recommendations (i)–(iv) put in place, cases of doxing become much less likely. And moral responsibility for the doxing would become much less ambiguous; obviously those who dox are responsible, but so too is SF, should they fail to meet (i)–(iv).

Discussion Questions

1. To what degree is preventing doxing the responsibility of a company who manages data?
2. Can you think of a counterexample to an outright rule against doxing? In what instance might doxing be permissible?
3. Do you think corporations that manage security should be more concerned with the potential for harm done to their clients, or with maintaining a standing set of rights for those clients? In other words, if rights are violated online and no one is hurt, is it a crime?
4. If considered using consequentialism, what kinds of harms and benefits should we be interested in? If deontological, what sorts of duties do the people involved in each case bear?

References

Alexander, L. and Moore, M. (2015). Deontological ethics. *The Stanford Encyclopedia of Philosophy.*

Baker, P. (2001). Moral panic and alternative identity construction in Usenet. *Computer-Mediated Communication, 7*(1).

Delaney, D. (1994). Noticeable phenomena of Usenet. Retrieved from http://hack.org/mc/texts/net-legends.txt

Douglas, D. (2016). Doxing: A conceptual analysis. *Ethics and Information Technology, 18*(3), 199–210.

Emeka, I. (2013). Doxing: A way of tracing anonymous people. *Professional HackingTricks.*

Franceschi-Bicchierai, L. (2015). Anonymous hackers officially dox hundreds of alleged KKK members. *MotherBoard.*

Garber, M. (2014). Doxing: An etymology. *The Atlantic.*

Giganews. (2017). Giganews' Usenet history. Retrieved from http://www.giganews.com/usenet-history/

Southern Poverty Law Center. (2017). Ku Klux Klan. Retrieved from https://www.splcenter.org/fighting-hate/extremist-files/ideology/ku-klux-klan

Rudin, S. (Producer), Cean C. (Producer), De Luca, M. (Producer), Brunetti, D. (Producer), Fincher, D. (Director). (2010).The Social Network [Motion picture]. United States: Columbia Pictures.

Liener, B. M. et al. (2012). Brief history of the Internet. *Internet Society.*

McCormick, T. (2013). Anthropology of an idea: hacktivism. *Foreign Policy*, (200), 24.

Moody, C. (2015). Donald Trump gave out Lindsey Graham's personal cell number to America. *CNN Politics*

Ronson, J. (2015). How one stupid tweet blew up Justine Sacco's life. *The New York Times.*

Ryan, T. (2014). What are social media's responsibilities in doxing cases? *Bloomberg.*

Sinnott-Armstrong, W. (2015). Consequentialism *The Stanford Encyclopedia of Philosophy* Edition)

Wayback Machine (2016) Retrieved from https://archive.org/web/

Author Biographies

Nick Gonella is a third year Computer Scientist at Cal Poly, where he works as both a researcher and teaching assistant. He specializes in security and operating systems design. After he graduates, he hopes to continue on to graduate school and eventually teach in a university setting.

Lorenzo Nericcio is a Senior Philosophy major whose primary interests are in philosophy of mind and applied ethics. He hopes to pursue a PhD in philosophy, specializing in artificial intelligence ethics specifically, and the moral and political ramifications of emerging technologies broadly.

Cybersecurity Case Library
Volume 1, Issue 1

The Sovereignty of Prosecuting Cyber-Attacks on Cloud Computing

Dylan Howell, Kyle Libby

Abstract

When a cyber-attack occurs in the cloud, the global distribution of data and computing resources complicates the process of determining authority to investigate the incident and prosecute the offenders. In this paper, we present a hypothetical cyber-attack on a distributed cloud service to illustrate the complexity of determining sovereignty when an attack spans multiple legal jurisdictions. We explore this incident in the context of existing legal institutions to determine whether these frameworks are up to the demands of such a distributed attack and have formulated recommendations based on the results.

Introduction

Globalization has expanded the boundaries of how and where data can be stored and accessed across the world. Consequently, globalized data storage involves companies spanning several nations. There are existing supranational organizations, such as the United Nations, that have tried to establish jurisdiction among the anarchic relationships between nations. However, little has been developed in global cybersecurity law. This has led to an underlying question: who possesses the sovereign right to try those cases of cyber-attacks on globally distributed data? The implications are even greater because as the 21st century continues to globalize, attacks within these bounds will expand.

This paper explores the technical details of a hypothetical cyber-attack case against a globally distributed cloud service. The case will be used to examine the readiness of international legal institutions to handle a cyber conflict spanning multiple legal jurisdictions. First, we will provide background information to define institutions and technologies relevant to our case study. Next, the hypothetical case will describe why and how a hypothetical, yet plausible, attack would be carried against a cloud service. From this case, we will investigate the significance of the attack by examining its legal complexities against international legal institutions. Finally, we will provide technical and political recommendations for the issues discussed followed by a series of discussion questions.

Background

Though the emergence of cyber warfare and the increasing interconnectedness of the world are contemporary topics, background research provides information that may be applicable in establishing legal precedent. Sovereignty in cyberspace is an adventurous and contentious topic, and it is certain to be fraught with the jurisdictional disputes among states and their respective prosecutorial bodies. There are few established standards for this issue, but the concept of various international jurisdictions is not new. The United Nations has attempted to bring countries to heel with institutions for various crimes; specifically, the International Criminal Court (ICC). However, the ICC lacks clear and consenting jurisdiction (Hurd, 2014). Not all countries submit to its justice (e.g. the United States), and if they do, the guidelines are specific in terms of when the ICC can intervene.

The other international legal body, the International Court of Justice (ICJ), is typically assigned to handling interstate conflict and advising the United Nations on legal matters (ICJ, n.d.). It would seem that the ICJ would handle such an engagement due to its interstate nature. However, jurisdiction within the ICJ is compulsory to United Nations members. Enforcement requires a vote of the Security Council and is subject to vetoes from the U.S., France, Great Britain, Russia, and China. Adjudicating these states would be nearly impossible to enforce through this mechanism (Hurd, 2014). Litigation of a "cyber" nature, in the context of this paper, has yet to arise. However, cyber-attacks between states have occurred. There are some existing mechanisms within international law that may be enlightening to the adjudication of the following scenario. For example, the Budapest Convention and United Nations Convention on Transnational Organized Crime are two existing frameworks worth considering, yet they both fall short of fully encompassing potential international cyber-attacks (Jamil, 2012; Al Hait, 2014). Both Conventions will be discussed in further detail in the Significance section of this paper.

The volume of data has grown to a scale at which major web services must rely on a distributed data model to meet the demands of an expanding global network. This model, known as cloud computing, was described by the National Institute of Standards and Technology (NIST) as "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources" (Mell & Grance, 2011, p.2). For the end user, a cloud system provides services that can be accessed from anywhere over the Internet. Cloud computing is used by over ninety percent of companies today ("RightScale 2016 State of the Cloud Report," 2016). The global demands of services such as social media, file sharing, e-commerce, and banking make cloud computing a prominent technology of modern society.

One of the primary features of cloud computing is data redundancy, in which data is copied in multiple data centers to prevent loss or service interruption if one data center is inaccessible. Copies of data are often stored closest to where they are accessed to reduce latency. Latency is the time it takes to receive that data according to physical distance. Despite this, there is no knowledge of where data is stored or from where it originates in cloud computing (Mell & Grance, 2011). "In a globally replicated database […] data […] are copied to geographically separated datacenters. For example, copies might be made in data centers in Singapore, on the west coast of the U.S., on the east coast of the U.S., and in France" (Kadambi et al., 2011).[1]

[1] This does not hold true for all cloud systems. For example, the system we will be basing our case study off of, Azure only replicates data within the region of its originating data center (Microsoft, 2017).

For the purpose of this case study, we will focus on a cloud system that has the potential to house data from different companies spanning multiple jurisdictions. To draw a parallel to an existing technology, our cloud system will be modeled after Microsoft Azure: a cloud system that offers services managed in Microsoft data centers located around the world ("What is Azure," n.d.). Azure allows cloud services to be built and managed without having to focus on the actual infrastructure (Torre, Singh & Turecek, 2015). This boils down to abstracting away from physical server location. The technical complexities of managing data over geographically separated data centers are handled by Azure ("Azure Regions, n.d.). For an attacker, the treatment of globally distributed cloud services as one logical infrastructure means a breach to any one data center could yield a path to others. With this definition of a cloud system, we can model an adversary to take full advantage of the jurisdiction complexities of distributed data storage.

Chronology of Case

Matt Swann, a Principal Engineering Manager for a globally distributed cloud service, Microsoft SharePoint Online,[2] has provided us insights into how a data breach could be carried out in an internationally distributed system, such as the one described in Background Information. Although there are limitations to exploring a hypothetical scenario, Swann's expertise in cloud security has allowed us to craft a plausible attack using common practices in the cloud field. While a real incident would likely have any number of confounding variables, the true implications of this case study lie in the general impact of such an incident in the international sphere. The overall goal is to present a scenario that allows us to test current international cyber law against complexities that could occur—"it is vital to acknowledge that a breach has either already occurred or that it is only a matter of time until it will" (Microsoft, 2014).

In this scenario, an adversary (e.g. Russia) will be compromising a data center in Western Europe (e.g. Denmark). This data center holds information for several companies throughout the region, possibly outside of Denmark. All data is replicated entirely to another data center in a nearby country in the region (e.g. Belgium). The data centers are owned by a fictional company, "Contoso," in the United States and is operated by a subsidiary company of Contoso in the data center's country of residence.[3]

Swann states that an adversary carrying out a targeted data attack is a "human with a set of tasks, goals, and motivations." Swann has identified our adversary's tasks based on the intrusion kill chain (Hutchins, Cloppert, & Amin, 2010). First, the adversary will target a company that stores data on Contoso's cloud service and perform reconnaissance to determine where the data is located. Second, the adversary will penetrate Contoso's cloud service to gain access to present and future data owned by a company who stores their data on this cloud service. Third, after the adversary has found their way to the target data, they will install malware to allow persistent remote access. Finally, the adversaries will discreetly "exfiltrate" the data to their own private server (Hutchins et al., 2010).

---

[2] SharePoint Online is a cloud service that allows companies to build internal websites for information collaboration and sharing ("SharePoint Online," n.d.).

[3] According to Swann, this is similar to how Azure data centers are managed in some instances.

To solidify the plausibility of a cyber-attack with this level of risk, we have identified our adversary's motivations: to gain financial and competitive leverage over a company. Companies use cloud services to house some of their most sensitive data. Intellectual property such as "confidential product designs, source codes, patents, or trade secrets" reveal how a company operates and can be used to damage the company, its shareholders, or its customers (Cidon, 2015, p.1). These motives are important to provide as context, as it distinguishes the attack from cyber warfare or a state-sponsored attack; motivation is illuminating to law enforcement and the judicial process to follow. Once the adversary has targeted a company for leverage, they can begin their execution path.

Swann continues to say that "most attacks are not through malicious access, but rather an abuse of legitimate access." The most effective way to gain an initial foothold is to abuse credentials for an internal computer owned by a Contoso employee.[4] The adversary will steal credentials of an employee through a spear phishing campaign, which involves targeting individuals with a fraudulent email that tricks them into providing sensitive information or installing malware on their computer (Han & Shen, 2016). To achieve this, the adversary will spend much of their time in reconnaissance to gather information on Contoso and its engineers. Reconnaissance is defined as "identification and selection of targets" to find "mailing lists for email addresses, social relationships, or information on specific technologies" (Hutchins et al., 2010, p.4); they can use this information to determine where the data is, who may have access, and who is most likely to be deceived by spear phishing email.[5] The spear phishing campaign will end with one carefully selected candidate receiving a single malware email to minimize the risk of detection and reporting.[6] If the email is opened and the malware is installed and is signaling to the adversary, they have gained access.[6]

The initial foothold (employee's computer) in Contoso's internal network will serve as a "hop point to compromise additional systems and move laterally inside the network" (Hutchins et al., 2010, p.5). At each step, the adversary will evaluate the context of the network from their perspective by answering the following questions: "Where is the data they are searching for from where they are? How can they use an engineer's tools to get to the data? How can they use an engineer's credentials to elevate their permissions closer to those needed to access the data?" [6] The adversary will craft a path from their initial foothold through the internal network and to the location of the target data.

Swann specifies that an adversary will persist to have access to present and future data; they will hide a malicious process in one of the available computers that can signal to their server and await tasks. To avoid suspicion, the adversary can disguise their malware as a benign process on a victim's computer. [6] Once this process is installed and running, the adversary has remote access back to the compromised computer.

Finally, the adversary can complete their objective of "exfiltrating" company data to a server under their control (Hutchins et al., 2010). The target data will be selected and stored somewhere in the data center to be removed. To avoid suspicion, they will be sending out small chunks of data at random intervals.[6] This strategy hides messages

---

[4] (M. Swann, phone interview, October 26, 2016).
[5] The adversary chooses spear phishing because of its effectiveness. Studies have shown that 40% of people, despite skill level or knowledge, are vulnerable to phishing attacks (Dhamija, Tygar & Hearst, 2006).
[6] (M. Swann, phone interview, October 26, 2016).

to the malicious server within the massive amounts of network traffic going through Contoso.[6] The adversary has now completed their intrusion kill chain. (Hutchins et al., 2010).

During this hypothetical attack, the adversary compromised a data center in the Denmark jurisdiction—which may be holding data from other European jurisdictions—with data replicated to the Belgian jurisdiction. The data center is owned by a company that operates within the laws of the United States but is operated by a subsidiary company in Denmark. In addition, the adversary could locate any data center and repeat the lateral process to find a path to it after this first breach;[7] from there the legal implications will only grow more complex. Who will hold the data sovereignty to prosecute this attack? From this scenario, we can analyze the significance of the legal complexities of an attack in the cloud-computing realm.

Significance

One would expect a transnational breach of data centers to generate international legal confusion. Similar instances of international cyber-attacks have occurred, such as the Love Bug virus of 2000. With $8 billion in total damage, officials were quick to pin down the responsible parties in the Philippines, but they lacked the tools and framework to prosecute them successfully. These inefficiencies stem from a lack of legal procedures and the principle of "dual criminality," meaning the "accused individual may be extradited 'only if the alleged criminal conduct is considered criminal under the laws of both the surrendering and requesting nations'" (Esworthy, Gonzalez, & Gauger, 2015, p.2). Since most of the world's data exists on geographically dispersed servers and information access points, attacks of this kind can cause damage to corporate, government, and personal domains. The plausibility of such an attack, combined with the scale and influence of distributed infrastructure, demands a legal system to handle complex cyber-attacks.

The existing institutions of the anarchic system of international relations are not well equipped to handle a case such as this. While there may be domestic legal procedures available to affronted parties, the ICC, UN, and even INTERPOL do not have a clearly defined process for adjudication on the international level. However, INTERPOL does facilitate international law enforcement cases. According to International Criminal Police Organization, "most cyber crimes are transnational in nature, therefore INTERPOL is the natural partner for any law enforcement agency looking to investigate these crimes on a cooperative level" (International Criminal Police Organization, 2017). The only true international framework that exists to handle international cybercrime is the Convention on Cybercrime, also known as the Budapest Convention. Ratified in 2004, the Convention "represents the only international instrument and the best hope for countries to establish common minimum standards of relevant offenses, prevent criminals operating from jurisdictions with lower standards and enable expedited and 24/7 international cooperation between law enforcement" (Jamil, 2012, p.110). It is important to realize that while 49 countries have signed the Convention, there is still much room for improvement and expansion. Nevertheless, key countries, such as the United States and countries in Europe, have ratified it; however, Russia is noticeably missing from the roster.

---

[6] (M. Swann, phone interview, October 26, 2016).
[7] (M. Swann, phone interview, October 26, 2016).

The importance of the above scenario plays into the existing framework and shows a need for stronger resolutions (i.e., the Budapest Convention) but more amicable to other countries, especially those where cyber crime is rampant. Another recent development in cyber law comes from a recent UN consensus, including Russia, on the applicability of international law. It states that "the explicit affirmation that international law, particularly the principles of the UN Charter, is applicable to state activities in cyberspace, including activities of non-state actors attributable to states, will allow the international community and affected states to react to violations more effectively" (Wolter, 2013, p. 27).

Developing clear legal jurisdiction in our example case is tricky and poses a complex question: should prosecution occur in the physical location of the server, in the country of the company that owns the server, in the location of the victim, or in the country of the perpetrator? Deciding on standard legal operating procedure in this scenario will likely chalk up to it depends. The United Nations Convention on Transnational Organized Crime (UNTOC) has provisions for international jurisdictions for physical crimes and could be a framework to address issues that have been posited so far. However, cyberspace is a rapidly changing domain that is not clearly laid out in UNTOC, as it lacks a definition and set of procedures. Indeed, "cybercrime can be covered under its articles, when cyberspace is used as an environment for committing organized crimes" (Al Hait, 2014, p. 76). Either the country of the violated party or the country where the offense is committed can obtain jurisdiction, but this does not solve the problem we face in this case study. Where do we define the offense? The UNTOC allows the country of the victim to prosecute, even if the perpetrator is technically stateless as it resides in the affronted country (Al Hait, 2014). Through legal interpretation, the UNTOC outlines jurisdiction for crimes committed through, but not within, cyberspace. That remains murky and highlights the significance of this case study.

Recommendations

Swann provides technical recommendations[8] that can help defend against the issues discussed in this hypothetical attack. To thwart lateral movement within an organization, Swann's security team works to place additional steps and several traps between the adversary and company data. For example, partitioning the logical cloud infrastructure into regions limits the data available per breach and forces the adversary to repeat steps over again. [9] The goal is to exploit the Intruders Dilemma; the defender "only needs to detect one of the indicators of the intruder's presence to initiate incident response within the enterprise" (Bejtlich, 2009).

Swann proposes an effective means of attack preparation, which is to deploy regular red teaming—the testing of live cloud infrastructure for vulnerabilities as well as the detection and response capabilities of the teams managing the infrastructure security (Microsoft, 2014). Swann's team specifically follows an Assume Breach model, meaning their focus lies on detection, response, recovery, and further prevention of attacks. Human intervention is a necessary complement to automated prevention techniques (Microsoft, 2014). The goals are to become better equipped to mitigate weaknesses exploited in the intrusion kill chain, cut off attackers earlier in their chain, and minimize data breaches and their effects.

---

[8] For a more in depth view into the security behind cloud services, view Matt Swann's talk on SharePoint Online security at https://youtu.be/cS8S5hjw-cc

[9] (M. Swann, phone interview, October 26, 2016).

Regarding political recommendations, we must realize that we are in relatively uncharted territory regarding the expansion of jurisdiction in the cyber realm. The most potent action to remedy this confusion starts with creating definitions; legally, cybersecurity lacks unifying geopolitical standards. While there are overarching legal bodies that could be expanded to handle a similar case, those bodies do not have explicit guidelines in this scenario. While interpretation may be useful to some degree, potent action requires clear guidelines. The actual adjudication of such definitions would give rise to precedent, allowing for further development and civil understanding of how to handle this case.

For those concerned with increases in bureaucratic interference in the event of expanding definitions and responsibilities, there are existing agencies and committees that simply need to come to agreement on relevant definitions. This has not happened yet, mostly to avoid issues that can arise from poorly worded laws, as some countries may have different beliefs and plans for cyberspace. Finding a way to generate consensus is half of the battle when dealing with international relations.

Discussion Questions

1. The authors talk about definitions for cybersecurity and the importance of clearly laying out legal jurisdiction. What are some possible definitions for terms like cyber-attack, data breach, and cyber sovereignty?
2. Regarding external implications of this case study, how do you think that certain countries would react to an attempt to standardize legal procedures? (e.g. Russia, United States, China, etc.)
3. Why might an adversary choose the route that utilizes credentials and social engineering rather than other technical means in the intrusion kill chain? What are the tradeoffs of each path of execution?
4. What is at risk if an adversary performed a similar hypothetical attack on domains such as social media, public file storage, financial, or governmental platforms?

References

Al Hait, A. (2014). Jurisdiction in cybercrimes: A comparative study. *Journal of Law, Policy, and Globalization*, 22, 75-84.

Bejtlich, R. (2009). Defender's dilemma vs. intruder's dilemma.

Cidon, A. (2015). Protecting intellectual property in the cloud.

Dhamija, R., Tygar, J. D., & Hearst, M. (2006). Why phishing works. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 581-590). ACM.

Esworthy, M., Gonzalez, J., & Gauger, N. (2015). Cases without borders: *Unique challenges in international cybercrime investigations.*

Gilbert, S., & Lynch, N. (2002). Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services. *ACM SIGACT News, 33*(2), 51-59.

Han, Y., & Shen, Y. (2016). Accurate spear phishing campaign attribution and early detection. In *Proceedings of the 31st Annual ACM Symposium on Applied Computing* (pp. 2079-2086). ACM.

Hurd, I. (2014). *International Organization: Politics*, Law, Practice (Second). New York, NY: Cambridge University Press.

Hutchins, E., Cloppert, M. J., & Amin, R. M. (2010) Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains.

International Court of Justice. (n.d.). *The Court.*

International Criminal Police Organization. (2017). *Cybercrime.*

Jamil, Z. (2012). Global fight against cybercrime: Undoing the paralysis. *Georgetown Journal of International Affairs*, 109-120.

Kadambi, S., Chen, J., Cooper, B., Lomax, D., Ramakrishnan, R., Silberstein, A. & Garcia-Molina, H. (2011). Where in the world is my data. In *Proceedings International Conference on Very Large Data Bases (VLDB)*.

Mell, P., & Grance, T. (2011). The NIST definition of cloud computing.

Microsoft. (2014). Microsoft enterprise cloud red teaming.

Microsoft. (2017). Azure storage replication.

Microsoft. (n.d.). Azure regions. Retrieved from https://azure.microsoft.com/en-us/regions/

Microsoft. (n.d.). SharePoint Online. Retrieved from https://products.office.com/en-us/sharepoint/sharepoint-online-collaboration-software

Microsoft. (n.d.). What is Azure. Retrieved from https://azure.microsoft.com/en-us/overview/what-is-azure/

RightScale. (2016). RightScale 2016 State of the Cloud Report. Retrieved from http://www.rightscale.com/lp/2016-state-of-the-cloud-report

Torre, C., Singh, K., & Turecek, V. (2015). Microsoft Azure - Azure service fabric and the microservices architecture.

Wolter, D. (2013). The UN takes a big step forward on cybersecurity. *Arms Control Today, 43*(7), 25-29.

Author Biographies

Dylan Howell is a 4th year Computer Science major focusing his studies in computer security topics. Dylan is passionate about upholding privacy and civil liberties with secure and ethical software design. He is following a career path of developing software that enables philanthropic impacts on the world.

Kyle Libby is a Senior Political Science major with a concentration in Global Politics, graduating in Spring of 2017. In addition to research in cyber security, Kyle has written on comparative policing strategies, conflict resolution, as well as coauthored an article for the upcoming Oxford Encyclopedia of Foreign Policy Analysis, to name a few. Kyle's academic interests lie in international relations, foreign policy, and crisis management. He hopes to receive a Masters in International Relations and find a career in public service through the State Department, FBI, or other federal agencies.

CONTACT US

Bill Britton
Director, California Cybersecurity Institute
Phone: 805-756-2190
Email: bibritto@calpoly.edu

Jimmy Baker
Interim Director for Industry Outreach
Phone: 805-756-2948
Email: jbaker30@calpoly.edu

Martin Minnich
Program Manager, California Cybersecurity Institute
Phone: 540-903-4004
Email: mminnich@calpoly.edu

Paul Jurasin
Director, Digital Transformation Hub
Phone: 805-756-5582
Email: pjurasin@calpoly.edu

CAL POLY California Cybersecurity Institute