Ring Signatures for Anonymous Sourcing in Journalism [*]

Jayshree Sarathy[†] Catherine Kerner[‡]

June 28, 2019

Abstract

The use of anonymous sources in journalism is often essential for conveying newsworthy information to the public. As a check against opaque institutions, anonymous sourcing is a critical tool employed by the press to maintain an informed public. In recent years, however, anonymous sources have become more reluctant to share information with the media. Mass surveillance and aggressive prosecution of information leaks have weakened the protections that journalists are able to provide to vulnerable informants. Anonymous sourcing, as a fundamental tool of the free press, is therefore less effective than ever before.

In this paper, we examine the workflow and impacts of an existing cryptographic technology, ring signatures, when applied to anonymous sourcing. Using ring signatures makes it possible for the public to verify that a message originated from a member of a publicly known set or ”ring” of parties, such as the set of senior White House officials, without knowing which individual produced the signature. We demonstrate that ring signatures provide strong protections for sources and provable deniability for reporters.

Moreover, we argue that the ring signatures protocol allows for better adherence to the journalistic ethical standards of anonymous sourcing. In light of recent disinformation campaigns and growing public mistrust of media, it is important for news outlets to address the fear that using anonymous sources could result in spurious or falsified reports. Through a case study of the New York Times anonymous op-ed, *I Am Part of the Resistance Inside the Trump Administration*, we find that although ring signatures do come with risks, they offer the benefit of enabling the public to independently verify the existence of authentic sources and corroborating reports. We demonstrate that ring signatures, if used in conjunction with traditional methods, can lead to clearer, more ethical and more accountable anonymous sourcing.

[*] Draft version 9/5/19. This paper has not been peer reviewed. Please do not copy or cite without authors' permission.

[†] Department of Computer Science, Harvard University. jsarathy@g.harvard.edu

[‡] Departments of Computer Science and Philosophy, Harvard College. cmkerner@college.harvard.edu

1 Introduction

The journalistic practice of using anonymous or unnamed sources has long been a crucial yet controversial tool of the free press. Protecting a source's identity is often the only way to expose information to the public; yet, anonymous sourcing is dangerous for sources as well as reporters, and has led to spurious reporting (Meyers, 2011; Duffy & Freeman, 2011). In our age of mass surveillance, public mistrust of media, opaque institutions, and aggressive prosecution of information leaks, the trade-offs of using anonymous sources are becoming increasingly pronounced (HRW, 2016). There is a growing need to address the ethical and practical risks of anonymous reporting for sources, journalists, the free press, and the public.

This paper proposes and evaluates modifications to the process of anonymous sourcing, centered on the incorporation of a preexisting technology: ring signatures. Ring signatures (Rivest, Shamir, & Tauman, 2001) are a cryptographic protocol for anonymous authentication. They allow a source to authenticate [1] a message while hiding within a publicly known set or "ring" of parties. Presented with a ring signature and the corresponding ring, it is possible for anyone to mathematically verify that a piece of information in a report originated from a member of the ring, without being able to pinpoint the exact source. For example, if a CEO of a company received an electronic, ring-signed message calling for internal reforms, she would be able to ensure that the message came from an employee in her company, without knowing precisely who sent it.

We outline a detailed workflow, in which a source wishing to leak information anonymously constructs a ring, signs a message as a member of the ring, and sends the information securely (using additional protections such as encryption, anonymous communication channels, and secure timestamping) to a journalist. The journalist validates the ring signature and decides whether or not to publish the information. We also describe the use of linkable ring signatures (Liu, Wei, & Wong, 2004) to allow multiple distinct sources from within the ring to corroborate the original source's account.

We are not advocating that ring signatures replace traditional methods of anonymous sourcing. Indeed, we find that the ring signatures protocol has several drawbacks, such as limiting the journalist's ability to investigate the motives of the source and implicating ring members without their consent. Rather, we argue that the ring signature protocol, used either by itself or as a supplement to other methods, is useful for high-risk situations in which traditional anonymous sourcing falls short. In these situations, using ring signatures gives sources quantifiable anonymity, offers reporters provable deniability, and enables transparency and verifiability for the public. In particular, our analysis of a case study (the NYT Anonymous Trump Official's Op-ed from September 2018) from the perspective of the Society of Professional Journalists' Ethics Guidelines for Anonymous Sources demonstrates that using ring signatures can enable clearer and more ethical anonymous reporting.

[1] In cryptography and security, authenticity is used to specifically convey that a message came from the stated sender. Oftentimes, authenticity is discussed along with the notions of unforgeability and integrity, which mean that a message has not been falsified by an adversary and has not been altered en route to its recipient.

2 Motivation

Anonymous sourcing is a critical tool for allowing the press to convey information to the public. With the rise of mass surveillance and aggressive prosecution of information leaks, however, traditional anonymous sourcing is no longer as effective as it once was. The diluted guarantees of anonymous sourcing, together with increasingly opaque institutions, are causing a chilling [2] of our free press. In this paper, we propose a method of fortifying anonymous sourcing against recent threats using preexisting tools in cryptography. As an extra benefit, we show that this method can enable clearer and more ethical usage of anonymous sources.

The guarantee of anonymity for sources has become much weaker in the last decade due to heightened prosecution of information leaks, starting with the Obama administration and continuing with the Trump administration. Before 2008, only three cases of whistleblowing or secret leaking had even been prosecuted by a presidential administration, but while President Obama was in office, his administration brought at least nine cases to court. The Trump administration seems to be just as strict (Risen, 2016). Sources can no longer rely on journalists to protect their identity, for courts may compel journalists to reveal this information under subpoena. Thus, those leaking secrets anonymously find themselves at greater risk of de-anonymization.

The rise of mass surveillance has also greatly diluted the promise of anonymity. Twenty-first century technology gives investigators the ability to track communication between sources and reporters (Savage, 2017). Sources have reason to hide their identity, even from reporters themselves, due to the increasing risk that their identity is exposed regardless of a breach of trust by the reporter (CPJ, 2013). Journalists covering national security issues are forced to take extreme measures such as meeting sources in secret, making sure they are not followed, and using burner phones and non-network-connected laptops (Risen, 2016). The rise of surveillance has robbed journalists of the ability to promise anonymity for sources, which has made it much harder for them to obtain and convey sensitive information to the public.

While increased surveillance, criminal investigations, and prosecutions of leaks have weakened the effectiveness of anonymous sourcing, anonymous sourcing is more needed than ever. The government has established new programs that further limit government officials' contact with the media. These include over-classification of sensitive information, restriction of leakage of unclassified information through the Insider Threat program, and limiting officials' contact with reporters. This aggressive crackdown on all fronts, combined with the diluted guarantees of anonymous sourcing, has caused a chilling effect on our free press and its ability to expose the inner workings of the government to the public (HRW, 2016).

It is also important to note that the use of anonymous sources comes with major ethical concerns. While anonymous sourcing is often necessary for groundbreaking stories, it is also considered poor practice. Relying on such sources is generally a last resort (SPJ, n.d.). Anonymous sourcing has two main issues. First, using unnamed sources can diminish the quality of the reporting itself. Even when maintaining the confidentiality of a source, reporters tend to describe the source in terms that are more vague than necessary, leading to unclear reports. Second, the anonymity of a source

[2] In legal contexts, a chilling effect refers to the "inhibition or discouragement of the legitimate exercise of a a constitutional right, especially one protected by the First Amendment to the United States Constitution, by the potential or threatened prosecution under, or application of, a law or sanction" (*https://www.yourdictionary.com/chilling-effectlaw*, n.d.).

denies readers the information needed to fully evaluate the source's credibility and motivation. Readers object that anonymous quotes are unreliable and question the credibility of articles that hide their sources. Uninhibited employment of anonymous sources can damage a news outlet's public perception. In the face of mass disinformation and public mistrust of media, traditional methods of anonymous sourcing [3] would benefit from updates that facilitate clarity and public verifiability.

The current climate is one in which anonymous reporting is increasingly important yet increasingly dangerous. The severity of the consequences for those involved and the heightened challenge of keeping the secret of a source's identity pose threats to the effectiveness of our free press. Given that information flows are more restricted than ever, and that anonymous sourcing is one of the only tools that journalists have to fight against tightholds over information, it is clear that anonymous sourcing must be re-imagined to remain resilient against mass surveillance and prosecution. In this paper, we explore whether existing tools in cryptography can improve protections for sources, journalists and readers, while also enabling more ethical anonymous reporting.

3 Related Work

Ring signatures were originally introduced for the same application that we consider in this paper: anonymous sourcing in journalism. The authors posed the following scenario:

> Suppose that Bob (also known as "Deep Throat") is a member of the cabinet of Lower Kryptonia, and that Bob wishes to leak a juicy fact to a journalist about the escapades of the Prime Minister, in such a way that Bob remains anonymous, yet such that the journalist is convinced that the leak was indeed from a cabinet member. (Rivest et al., 2001)

But almost twenty years after the original ring signature construction was published, they have, in a sense, remained on the shelf. Despite several theoretical papers that have outlined new variants and constructions, and a few practical implementations in cryptocurrencies such as CryptoNote, ShadowCash, and Monero, ring signatures have not been used for their original purpose of anonymous sourcing in journalism. To our knowledge, they have not even been seriously considered or discussed within the journalism community.

In this paper, we take a second look at using ring signatures in journalism (Rivest et al., 2001). Due to various factors in our current climate that have added new complexities to anonymous sourcing, we believe it is time to critically analyze the potential impact of ring signatures in this context. Our contribution is not a new theoretical construction, but rather a set of practical workflows from the perspectives of the source, journalist, and the public, along with an analysis and case study of how ring signatures can mitigate the risks and ethical concerns of anonymous sourcing.

[3] By traditional methods, we mean long-standing as well as current practices for acquiring and communicating with unnamed sources. This includes meeting with an anonymous source in person, sending and receiving letters, speaking on the phone, using a trusted intermediary, or communicating through secure or insecure digital applications. Traditional methods also include communication with off-the-record sources or observation of the anonymous source's situation and environment in order to make judgments on the truthfulness of the report.

4 Technical Background

In this section, we formally define the cryptographic building blocks that we use in our scheme, including public key infrastructure, digital signatures, ring signatures, and linkable ring signatures.

4.1 Public Key Infrastructure

Public Key Infrastructure (PKI) allows parties (people and computers) to verify the identity of other parties with whom they wish to securely exchange data. PKI creates a chain of trust so that identities within a network can be verified.

At the core of PKI is a Certificate Authority (CA), whose role is to verify the identities of parties. The CA is often called the root of trust: if the CA is trusted, then any identity certified by the CA can also be trusted. The CA issues a digital certificate to each verified party containing a public key, digital signature algorithm, a date range during which the certificate can be considered valid, and a digital signature of the CA itself.

Public keys are freely available in public directories. In practice, there are a few competitive CAs who have the majority of the market share. It is through the public directory of one of these providers or their competitors that a source would look up the public keys of the individuals they would like to include in their ring.

4.2 Digital Signatures

Digital signatures (Goldwasser, Micali, & Rivest, 1988) are mathematical schemes for authenticating digital messages or documents. They offer the following properties.

> *Authentication.* The message was sent by its stated sender.
> *Non-repudiation.* The sender cannot deny having sent the message.
> *Integrity.* The message is unaltered.
> *Unforgeability.* It is unfeasible for an adversary to forge a valid signature for a given message.

Formally, let m be the message, pk the signer's public key, sk the signer's secret key, n be the security parameter, and σ the signature. Then, a digital signature scheme is a triple of probabilistic polynomial-time algorithms (Gen, Sign, Verify):

$$\begin{aligned}
&\text{Gen}(1^n) \rightarrow (pk, sk)\\
&\text{Sign}(m, sk) \rightarrow \sigma\\
&\text{Verify}(\sigma, m, pk) \rightarrow \{\text{Accept, Reject}\}
\end{aligned}$$

If a signer, Alice, wants to send a signed message to her friend, Bob, she does the following. She first runs $\text{Gen}(1^n)$ to produce a public-private key pair (pk, sk). She posts her public key pk in a public directory and has a certificate authority validate that she is the owner of the public key. Then, to send her message m to Bob, she runs Sign(m, sk) to produce a signature σ. When Bob receives her message and signature, he runs Verify(σ, m, pk) to check that the message really came from Alice.

Correctness of the scheme guarantees that any valid signature will be accepted with probability 1. Security of the scheme guarantees that all invalid or forged signatures will be rejected with negligible probability.

4.3 Ring Signatures

Ring signatures (Rivest et al., 2001) are an 'anonymized' variant of digital signatures. Ring signatures allow the recipient to know that a message was signed by a member of a certain set of signers (the 'ring'), without revealing which member actually produced the signature. We define a ring to be the set of possible signers, the signer as the ring member who is responsible for producing the signature, and non-signers as the ring members who did not produce the signature. Each ring member, i, should be associated with a public key pk_i that defines the signature scheme and that has a corresponding private key sk_i.

Ring signatures are powerful because *no coordination* is required between members of the ring. Assuming that the signer knows the public keys of others, she can independently create a ring that includes herself and sign any message using her private key and the others' public keys. Additionally, the ring signature can be constructed even when members have used different independent public key signature schemes with different key and signature sizes. The original paper describes a ring signature construction for RSA and Rabin signature schemes.

Formally, let m be the message, $pk_i, i \in \{1, r\}$ be the non-signer's public keys, j be the index of the signer, and sk_j be the signer's private key. Then, the ring signature scheme is the pair of probabilistic polynomial-time algorithms (Ring-Sign, Ring-Verify):

$$\text{Ring-Sign}(m, pk_1, pk_2 \ldots, pk_r, j, sk_j) \to \sigma$$
$$\text{Ring-Verify}(m, pk_1, pk_2, \ldots, pk_r, \sigma) \to \{\text{Accept, Reject}\}$$

Correctness of the scheme guarantees that any valid signature will be accepted with probability 1. In addition to the security guarantees of traditional digital signatures, ring signatures also guarantee signer ambiguity– the recipient will have negligible probability in distinguishing the signer from non-signers within the ring.

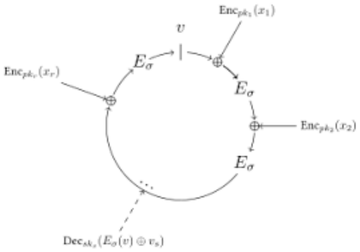

Figure 1. Ring signature scheme (Rivest et al., 2001)

The workflow of the ring signature is the following. First, parties $1, \ldots, n$ post their public keys, $pk_1, \ldots, pk_n$ in a public-key directory. They do not do this for the purpose of being included in a ring signature; this is a routine step for maintaining secure communication in general. Next, for a message m, a signer $i, i \in [n]$, can choose a subset Q of parties $1, \ldots, n$. Q is the ring. She generates a signature σ for m relative to the public keys for parties in Q. The signer does not need to inform or obtain consent from the other parties in Q; she can generate this signature on her own. It is important to note that the signer cannot create a ring signature for a ring Q unless she herself is a member of Q. At the same time, she cannot prevent anyone from constructing a ring Q that includes herself. She publishes her message m, the signature σ, and the ring Q. Anyone who wishes to verify the signature for a message m needs to know Q and the public keys of parties in Q.

Typically, the signer's anonymity is compromised when a subset of the others' secret keys are exposed. Recent constructions of ring signatures, however, have provided higher levels of anonymity. For example, *threshold anonymity* says that only if greater than t-out-of-n secret keys are exposed will the signer be revealed. The strongest guarantee is *anonymity under full key exposure* (Bender, Katz, & Morselli, 2006). This means that even if all the secret keys of members in Q are exposed, the signer will still remain anonymous.

4.4 Linkable Ring Signatures

Linkable ring signatures (Liu et al., 2004) are a variant of ring signatures with the additional property of linkability, which is defined below.

> *Linkability.* One can detect two signatures from the same signer.

Formally, let m be the message, $pk_i, i \in \{1, r\}$ be the non-signer's public keys, j be the index of the signer, and sk_j be the signer's private key. Let σ_1 and σ_2 be two signatures for different messages. Then, the linkable ring signature scheme is the triple of probabilistic polynomial-time algorithms (Linkable-Ring-Sign, Linkable-Ring-Verify, Linkable-Ring-Check):

$$\begin{aligned}
&\text{Linkable-Ring-Sign}(m, pk_1, pk_2 \ldots, pk_r, j, sk_j) \rightarrow \sigma \\
&\text{Linkable-Ring-Verify}(m, pk_1, pk_2, \ldots, pk_r, \sigma) \rightarrow \{\text{Accept, Reject}\} \\
&\text{Linkable-Ring-Check}(\sigma_1, \sigma_2) \rightarrow \{\text{Linked, Not Linked}\}
\end{aligned}$$

Correctness of the scheme guarantees that any valid signature will be accepted with probability 1. In addition to the security guarantees of traditional ring signatures, linkable ring signatures also guarantee linkability– the recipient will be able to detect two messages from the same signer with overwhelming probability.

5 Guidelines for Anonymous Sources

In addition to technical background, we provide background on ethical and practical guidelines for using anonymous or unnamed sources. The ethical guidelines come from the Society of Professional Journalists, and the practical insights are from the New York Times. Later in the paper, we argue that using ring signatures can allow journalists to better adhere to these guidelines.

5.1 Society of Professional Journalists (SPJ) Ethics Guidelines for Anonymous Sources

The Society of Professional Journalists maintains the following guidlines for anonymous sources.

1. Identify sources whenever feasible.

> Reporters should use every possible avenue to confirm and attribute info before relying on unnamed sources. They should identify the source as clearly as possible without breaching the confidentiality of the source (SPJ, n.d.).

2. Always question sources' motives before promising anonymity. Clarify conditions and keep promises.

> Reporters should make sure that confidentiality is not used to undermine others, attack opponents, or push personal agendas. They must weigh the value of the news to the public with the cost of not being transparent. They should get approval of their supervisors and editors and make sure they comply with the policy of their news outlet. They should aim to publish only with verification of multiple sources (SPJ, n.d.).

5.2 New York Times (NYT) Journalistic Practices

In June 2018, the New York Times released a short article to shed light on their journalistic practices surrounding anonymous sourcing. According to this article, the NYT is rigorous in deciding what is acceptable to publish. First, anonymous sources are only used for information that is credible, newsworthy and unable to be shared with the public in any other way. Second, at least one person other than the reporter is required to know the identity of the source, and the number and seniority of people with whom this information is shared scales with its centrality to the report. To address concerns of anonymous sourcing leading to untruthful information, the New York Times notes that many of their journalists find that anonymous sourcing often improves, rather than detracts from, information quality: sources may be more honest if they cannot be punished for what they say. At the same time, New York Times reporters and editors are careful to investigate the motivations and credibility of the anonymous source, and to reveal as much of this background as possible to readers without compromising the confidentiality of the source (Philip, 2018).

6 Workflow

We now describe the workflow of our proposal in detail. First, we state the assumptions needed for the scheme. Then, we list the protocols that the source, journalist, readers, and courts will follow. Throughout the protocols, we highlight points where important policy decisions must be made. We conclude the section by discussing practical implementation of the proposal.

6.1 Assumptions

We assume that the members of a group or institution, such as the group of senior White House officials, or the Board of Directors for a certain company, all have public keys. These public keys should be listed in a public directory, along with the corresponding signature scheme for each key, and they should be certified by a certificate authority. We also assume that all journalists have certified public keys in a public directory.

6.2 Source's Protocol

Let Sam be a source who wishes to leak sensitive, newsworthy information to the public. Sam could, of course, directly post his ring-signed information online for the public to access. However, we assume that he prefers to go through a journalist, Jane, because that will give his information appropriate credibility. Before creating the ring signature, Sam must think through the following questions and make decisions based on his desired level of anonymity, the information he wishes to leak, and the specifics of his organization.

First, how many members should be in the ring? The larger the size of the ring, the more anonymity Sam will have. The smaller the ring, the more specific the journalist can be in describing where the information came from, so the information may be taken more seriously by readers. For example, a ring that can be described as "White House officials" will lend Sam a good amount of protection, but may not be as powerful as a smaller ring of "White House cabinet members." Thus, Sam must weigh his desire for anonymity with the newsworthiness of his information.

Second, who should be included in the ring? The answer to this question depends heavily on the contents of Sam's information. If Sam's information contains something the President said when only Sam and cabinet members were in the room, then a ring that contains no cabinet members offers him no anonymity. Sam must also consider that including people in the ring connects them in some way, likely without their consent, to his message.

Once Sam has decided on a list of members to include in the ring, he can construct his ring signature. First, he collects the public keys of the ring members from the public directory, making sure that all keys are valid under a certificate authority. Next, he encrypts his message using the journalist's public key and signs the encrypted message using the public keys of his chosen ring members and his own secret key. Finally, he uses an anonymous communication tool such as Tor to send his message to Jane. He also makes sure to securely time-stamp his signed message using a blockchain transaction or a widely-trusted time-stamping authority [4].

[4] A time-stamping authority is similar to a certificate authority. It certifies the time-stamp of electronic signature.

6.3 Journalist's Protocol

Upon receiving the message, Jane verifies the authenticity and time-stamp of the message. She checks all the public keys of ring members in the public directory to makes sure they are valid under a trusted certificate authority. Based on the identities of the ring members, she comes up with the most specific descriptor that applies to all of them. For example, she might label them as "White House officials within three ranks from the President." Then, she decrypts the message using her private key.

At this point, Jane also has a number of decisions to make. Some are specific to using ring signatures. Is the set of members small enough to make the contents of the message credible? Is the set of members relevant enough to the message to make the information credible? Is the set of members large enough to afford the source a reasonable amount of anonymity?

She must also ask several other questions that are common with traditional methods of anonymous sourcing. For example, does independent evidence exist to support the information? Can she obtain the information in this message from a named source? Is the information newsworthy enough to justify using an anonymous source? Would citing this information in her reporting follow the guidelines of her supervisors and her media outlet? Note that Jane should maintain the ethical standards when deciding whether or not to publish the information as she would with any other information she might receive.

If Jane and her supervisors decide to use the information in her report, she should describe the source using the most specific descriptor of members of the ring (ie. members of the Senate Judiciary Committee). Then, she should include the ring signature and information (both of which can be represented mathematically as a sequence of zeroes and ones) in her citation of the information. She may choose to publish the public keys of the ring members (which, again, can be represented as zeroes and ones) for greater transparency, or to keep this information private unless requested under a subpoena.

On the other hand, if Jane is hesitant to include the information in her report, she can simply treat the information as a lead for further investigation without directly referencing it in her report.

6.4 Public's Protocol

If Jane includes the public keys of the ring members in her citation, any reader can independently verify the authenticity of an anonymously-sourced quotation or piece of information. To do so, the reader can simply run the publicly known Ring-Verify algorithm on the public keys, ring signature, and information cited in the report.

If the algorithm outputs Accept, then the reader can trust that the information truly originated from someone within the ring. The reader can even verify the descriptor assigned to the ring members by the journalist by looking up the public keys in the public directory.

Otherwise, if the algorithm outputs Reject, then the reader should be wary of trusting the anonymously sourced information in the report.

6.5 Court's Protocol

With traditional anonymous sourcing, the court could subpoena Jane for her knowledge of the anonymous source. If the ring signature protocol was used, however, the court can only ask Jane to hand over the list of public keys defining the ring that she may not have chosen to release. In many cases, this is sufficient for verifying that some information truly came from an authentic source described in the journalist's report. Unless there is concrete evidence that Jane had further contact with the anonymous source, courts cannot compel Jane to reveal information that she provably cannot have (since the ring signatures provably guarantee signer-ambiguity), which means Jane is protected from being held in contempt of the court.

Given the list of ring members, it is important to consider whether or not the court has a right to subpoena ring members for their private signing keys, which would allow investigators to identify the source (unless the ring signature scheme that was used guaranteed anonymity under full key exposure). Revealing a private key could have severe consequences for an implicated ring member who has used it to encrypt other important information unrelated to the secret-leaking case, so we assume that non-signers will not disclose their private keys unless the information is severely damaging to their reputations.

Usually, the Fifth Amendment protects ring members from being forced to reveal private keys. The Fifth Amendment says that no person "shall be compelled in any criminal case to be a witness against himself" (*U.S. Const. ammend V.*, n.d.). Private keys are long, random sequences of letters and numbers that are hard to remember. Thus, public key encryption applications usually allow owners to access their public keys through a second password chosen by the owner. This password is ideally committed to memory, grants the signer access to the private key, and only exists on hardware in hashed form that is computationally intensive to reverse. Under these conditions, legal scholar Scott Brady argues that private keys are safe from compulsory disclosure because a source being forced to reveal such a password to a private key would thereby be forced to "disclose the contents of his mind" (Brady, 2012). This assumes the password is not written down or stored in a physical form, in which case it is vulnerable to subpoena.

This argument as applied to n members of a signing ring needs to be extended slightly. The Fifth Amendment protects only information that is incriminating. In truth, revelation of a private key is only incriminating for one ring member: the signer of the message. However, since it is provably unfeasible to identify signers from non-signers, one could argue that it is unconstitutional to demand a private key of someone when its revelation has a $\frac{1}{n}$ probability of being self-incriminating. Perhaps more convincingly, it is impossible to say that compulsory revelation would not be self-incriminating. However, the question of compelled private key disclosure in the context of ring signatures remains open.

6.6 Extension with Linkable Ring Signatures

According to the SPJ ethics guidelines, Jane should use multiple sources to corroborate any information gained from an anonymous source. Linkable ring signatures, a preexisting variant of traditional ring signatures, are well-suited for enabling independent corroboration while still preserving the source's anonymity. Under a linkable ring signature scheme, the Jane can check whether two messages she receives from the same ring were signed by distinct members of the ring, Sam and another individual. If so, Jane may attribute greater credibility to the information and have better justification for its use in her report.

6.7 Practical Considerations

We do not advocate that sources, journalists, and the public perform the mathematical computations themselves to generate and verify ring signatures. It is unrealistic that anyone in the system has sufficient knowledge of the cryptography to create or verify a ring signature, and making this a requirement in practice would preclude any use of the system. Instead, we imagine the ring signature functionality to be built into existing digital communication tools. In particular, secure communication applications such as Tor, Signal, Telegram or WhatsApp are already heavily used by sources and journalists who wish to maintain privacy. We believe that ring signature functionality (both signing and verifying) could be easily incorporated into these tools. The verification functionality could also be implemented in browser extensions to cater to online readers of news media.

A secure application with ring signature functionality could be used as follows. The source could form the ring by searching and selecting desired ring members from online public key directories, accessed through the application. The application would then allow the source to ring-sign the message with her private key and her chosen ring members' public keys with the tap of a button. The application would also need to implement a verification feature. Per the verification protocol, the feature would require access to the public keys of all ring members, including the signer's. A journalist, upon receiving the ring signature, would be able to use this feature to verify that the signature originated from the appropriate ring. If readers had access to a similar feature, they would have a means of independent verification right at their fingertips through the application.

7 Further Issues to Consider

7.1 Truth of the Report

One risk of using ring signatures is that replacing personal interaction with a digital message introduces a lack of accountability and opens up the system to abuse. Yet, even when the identity of a source is known to a reporter, the truth of an anonymous report is always questionable. The current best practices for reporters discourage the use of anonymous sources whenever possible, encourage cross-referencing them with accountable sources, and insist on a consideration of the source's motives (SPJ, n.d.). When using ring signatures for anonymous sourcing, the judgment of the reporter is paramount, and we expect journalists to still adhere to the current best practices.

The fact still remains, however, that when using ring signatures, the reporter has limited means of questioning the motives of a source. One solution is for sources to include an explicit description of their motives for anonymity, and for journalists to categorically deny publication of information if the motive is not included. It is the source's responsibility to ensure that this field is as complete as possible, since that increases the chances that their information will be published. Both the motive and the message will be available to the journalist and other authorities (editors, supervisors) who currently make the decision of whether to permit using an anonymous source (Philip, 2018).

7.2 Availability of Public Keys

The ring signature scheme relies on the availability of public keys of all potential members in a ring. In fact, our protocol assumes that every person in an organization of interest has a long-term public key that is accessible through a public directory. However, this assumption is far from reality. Further research needs to be conducted to know the proportion of senior White House officials, for example, who have public keys posted in a directory, but it is safe to say that the answer is much lower than we would hope for.

Unavailability of public keys diminishes the usability of the protocol. For example, the best way for a source within Facebook's Board of Directors to maintain anonymity and credibility would be to include all of the board members in a ring. If there are only a few board members with available keys, the source will not have adequate anonymity. This exacerbates the challenge of constructing a ring of sufficient size and legitimacy to both protect the source and convince readers of the truth of the report.

The use of the ring signature protocol in a few high-profile whistleblowing cases could, in fact, dissuade people from publishing public keys in the first place. If high-ranking government officials knew that publishing a public key could implicate them in a whistleblowing scandal, they would likely be disincentivized from publishing one. This may change in the future, however, as more people use public keys for general secure communication. In that case, the utility of listing one's public key in a directory (for receiving secure messages in day-to-day business activities) may outweigh the slight risk of being included in someone else's ring signature.

7.3 Timing Attacks

Though a ring signature itself does not reveal which member signed the message, some information about the signer can still be leaked from side channels. In particular, the timing of public key listings, which is publicly available metadata, can give a clue to the signer of the message. If the source, Sam, created and listed his public key only after he realized he had some information to leak, ie. shortly before a ring-signed message was sent to the journalist, one would assign greater suspicion to Sam than to other members in the ring who had listed their public keys long before the message was sent. Such timing attacks weaken the guarantees of security for our ring signature protocol.

7.4 Implication of Ring Members

One of the major ethical issues with ring signatures is that a source can include ring members without their consent. The legal question here is: does the government have the right to investigate all implicated ring members? There are two cases to consider when addressing this question.

The first case is when the parties included in the ring perfectly match the language used to describe the ring: for example, a ring signature for a report from a “senior White House official” that includes all senior White House officials in the ring. This scenario is identical to that of a traditionally sourced report that references a ”senior White House official.” By the provable anonymity guaranteed by the ring signature scheme, such a ring signature does not implicate anyone who would not have been under scrutiny given the description in the article alone. The implications are then the same as they would be in a traditional anonymously sourced report.

In the second case, the ring members might only have a vague common descriptor, such as “a government official”. Since the number of government officials is quite large, such a ring would likely exclude some officials, and an observer would naturally be suspicious as to why only certain government officials were included. Now, the question is whether or not membership in this ring is sufficient grounds for government investigation. Should this depend on the size of the ring? The sensitivity of the information?

While it seems reasonable, though impractical, to allow investigation of all implicated members of the ring, it may be unreasonable to take more extreme action on the grounds of ring membership alone. As we concluded in a previous section, private keys are generally safe from compulsory disclosure under the Fifth Amendment. It is unclear whether these protections apply in the context of ring signatures.

7.5 Reassuring Skeptical Users

It may be difficult to convince persons unfamiliar with or skeptical of ring signatures of their strong guarantees. An application that incorporates ring signatures as a feature would need to effectively communicate the reliability of the protocol to its users, particularly to assure sources that their anonymity is irrevocable and to assure readers that a report provably originates from members of a ring.

8 Impacts

8.1 Impact on Sources

Using this scheme is beneficial to the source in that it provides her with a quantifiable, mathematical guarantee of anonymity. Even if the journalist to whom the source reveals her information is under surveillance, or receives a subpoena from court to reveal the source’s identity, the source’s anonymity cannot be compromised.

The drawbacks of using this scheme are in the difficulties of its implementation. It may be challenging for the source to choose a ring that matches her desired balance between anonymity and impact. It may be especially difficult to find certified public keys for all desired ring members. In addition, it may be time-intensive or confusing for the source to follow so many steps in order to convey her information to a journalist; she may prefer the simplicity of talking to a journalist in person.

8.2 Impact on Journalists

Receiving information via ring-signed messages has several benefits for journalists. First, the protection afforded by ring signatures to sources may counteract the chilling effect on information flows due to government surveillance of journalists. Because ring signatures offer unconditional anonymity, sources will be more open to conveying information to journalists.

Second, ring signatures protect journalists from breaking their promises of confidentiality. Even if a reporter receives a subpoena from court to reveal her source, she provably cannot reveal the particular signer of the information she received. Thus, journalists have more assurance of maintaining their credibility and their jobs, even when reporting on highly politicized or controversial stories.

Journalists will, however, face the challenge of not knowing the identity of their source. They will have to work harder to understand the different perspectives involved, to judge the veracity of the information, and to provide a well-researched report.

8.3 Impact on Public

The use of ring signatures in anonymous reporting makes it possible for readers in the general public to verify for themselves where the information is coming from, assuming that the ring membership is released. Specifically, it provides an objective definition of the type of source being used. Whereas terms like "government official" can be interpreted to refer to a wide range of informants, the ring membership will outline exactly which kinds of sources could be responsible for the report. Making the ring membership available provides a concrete source of truth, giving an unsatisfied reader additional assurance about the origin of the report.

More generally, the safe avenue for information flows from sources to journalists provided by ring signatures will likely allow the public to access more information about the inner workings of corporations, institutions, and governments. In our current climate of opacity, in which important information that affects citizens is kept increasingly hidden from the outside world, ring signatures could provide a massive benefit in maintaining an informed public.

9 Case Study: Anonymous Trump Official's Op-Ed

In September of 2018, the New York Times published an anonymous op-ed describing the "quiet resistance" against President Trump within his own administration. The op-ed was attributed by the New York Times to "a senior official in the Trump administration whose identity is known to us and whose job would be jeopardized by its disclosure" (*I Am Part of the Resistance Inside the Trump Administration*, 2018). The op-ed created a storm in the White House, among the press, and of course, on Twitter. The media and the public frantically analyzed the language of the op-ed to try and identify the unnamed official. Many, including President Trump, were skeptical that such a source existed, even accusing the New York Times of fabricating the entire article (Michael, 2018).

How does the use of anonymous sourcing in this op-ed stack up against the SPJ ethics guidelines? First, we consider clarity in describing the anonymous source. According to the SPJ guidelines, journalists should identify an unnamed source as clearly as possible. Although the New York Times specified that its unnamed source was a senior official in the Trump administration, we believe this characterization was too vague, especially in the context of the op-ed itself. The op-ed was newsworthy because it exposed internal resistance high up in the Trump administration; its credibility relies on its authorship by a senior official in the administration. But despite the importance of the source's seniority to the article, the New York Times never mentioned what criteria was used to determine whether or not the source was a senior official. Such an explanation is needed, for the usage of this designation in political media is far from standardized. According to Politico, "if the source insists on being called a senior administration official, chances are he or she can get away with it, no matter if by most objective measures – title, salary, proximity to power – they are not really all that senior" (Allen, 2007). The use of such a fuzzy identifier, however, may have been the best that the New York Times could have done. With traditional methods of anonymous sourcing, it may be unfeasible for the editors of the op-ed to identify the unnamed source more clearly without compromising the source's confidentiality.

We now consider how the ring signature protocol could impact the clarity of the source's identity in the op-ed. In this case, the journalist would obtain the ring-signed message as well as the ring membership. The journalist could then publish the list of ring members with the op-ed, and identify the source as one of the ring members. Readers could independently verify that the op-ed was digitally signed by a member of the ring. If the ring is chosen well by the source, publishing it would enable readers to understand how large the internal resistance really is within the administration, while leaving enough ambiguity to hide the identity of the source. The publicly available ring would enable more precise and explicit identification of the source.

Second, we evaluate guarantees of anonymity for the source. In the case of the anonymous op-ed, how well did the New York Times maintain anonymity of its source? The identity of the source still remains hidden nearly a year after publication of the op-ed, so it seems that the New York Times did fulfill this aspect of the ethics guidelines. However, just because the source's identity has remained hidden so far does not guarantee its anonymity in the future. It is possible, for example, that some electronic communications between the source and journalist were captured through surveillance and could be revealed in the future. Or, the journalist could eventually be compelled by court subpoena to reveal her source. Without provable and quantifiable guarantees of anonymity, it is hard to assess whether the New York Times succeeded in protecting the identity of its anonymous source.

The ring signatures protocol would enable us to more concretely evaluate the anonymity provided to the source. The level of anonymity would, in essence, be inversely related to the size of the ring. Since the source chooses the ring himself, the New York Times journalists would not be responsible for the level of protection provided by the ring, nor would they be able to degrade or compromise this protection even when compelled by court. In contrast to verbal or contractual guarantees of confidentiality, the guarantee of anonymity provided by a ring signature is not easily broken by surveillance or subpoenas. Thus, in terms of protecting an unnamed source, ring signatures provide a clear benefit over traditional methods of anonymous sourcing.

Our third axis of evaluation considers the thorough investigation of the anonymous source. The SPJ guidelines state that journalists should question sources' motives before promising anonymity.

Current practices align with this ethical guideline: according to the New York Times Journalistic Practices (Philip, 2018), New York Times reporters thoroughly investigate the motivations and credibility of a source before promising anonymity. It is reasonable to assume that this practice was maintained in the case of the anonymous Trump official's op-ed.

On this axis, current methods of anonymous sourcing remain superior to the ring signatures protocol. This is because traditional anonymous sourcing enables in-depth, in-person conversations between the source and journalist, while the ring signatures protocol lets the source completely eschew this interaction. By sending a ring-signed op-ed, the source remains completely hidden even to the journalist, so the journalist would be unable to directly question the source and to glean additional information about the source's motives. While this offers protection to the source, it also hinders the journalist from doing her job. For the journalist to be able to provide a well-researched, high-quality report, she must understand the motives, perspective, and background of a source, ideally through direct communication. The barriers that ring signatures create for this aspect of anonymous sourcing is, therefore, a significant risk of using the ring signatures protocol.

Finally, we consider verification of a report by multiple sources. It is unclear whether or not the New York Times followed this guideline for the op-ed, as corroborating sources were not mentioned by the editor. Here, ring signatures could offer more transparency to the public. After receiving the ring-signed op-ed from the initial source, the New York Times journalist could issue a public call asking another member of the ring to verify the information in the op-ed. Given the strong protections provided by ring signatures, it is very possible that the journalist would have received corroboration from additional sources in the administration and could make these available to readers. Thus, the ring signatures protocol (in particular, using linkable ring signatures) offers a safe avenue for sources to anonymously corroborate reports, as well as a procedure for the New York Times to safely pass on the supporting accounts to the public.

In our evaluation of the anonymous sourcing of the Trump official's op-ed, we find that the New York Times came up short of the SPJ ethical guidelines on multiple levels. The New York Times did not clearly identify its unnamed source, failed to provide strong guarantees of anonymity for its source, and did not disclose whether it obtained verification using multiple sources. These deficiencies are mainly due to the limitations of current methods of anonymous sourcing, which we believe can largely be overcome with the ring signatures protocol. As we have described, ring signatures enable clearer identification of sources, stronger protection for sources, and safer, more transparent verification from multiple sources.

The main drawback of the ring signatures protocol, with regard to the SPJ ethics guidelines, is that it limits journalists from thoroughly investigating the motives and credibility of anonymous sources. To address this limitation, the New York Times could use ring signatures in conjunction with traditional methods of anonymous sourcing. For example, if the journalist aims to provide more clarity and transparency for readers, but is not as concerned about the threat of a subpoena, she could first meet with her source in person to determine the source's motives and credibility, and then obtain ring-signed information from the source to include in her report. This combined protocol would improve the clarity and transparency of the report, without compromising on thoroughness or quality of information. In the context of the anonymous op-ed, it is evident that the ring signatures protocol, used in tandem with current methods, would enable the New York Times to maintain a higher ethical standard when using anonymous sources.

10 Conclusion

The use of anonymous sources is a fundamental tool of the free press for countering opaque institutions and maintaining an informed public. In recent years, however, the effectiveness of anonymous sourcing has come under threat. With the rise of mass surveillance and aggressive prosecution of information leaks, it is harder for informants to remain anonymous. In addition, growing public mistrust of the media makes it more important than ever to address the ethical concerns of using anonymous sources.

In this paper, we explored the impacts of ring signatures, an existing tool for anonymous authentication, on sources, journalists, and the public. We demonstrated that ring signatures provide strong guarantees of anonymity for sources, deniability for reporters, and both transparency and verifiability of reports for the public. We found that ring signatures do have a significant drawback, in that they hinder journalists from investigating the veracity of the source's report, but we described how the ring signatures protocol can be combined with traditional methods to mitigate this risk.

There are many areas for further exploration. First, the limitation described above points to the theoretical question of whether we can construct ring signatures that enable a back-and-forth conversation between the source and a journalist, which would allow the journalist to question the background and motives of the source. Second, we would like to conduct further legal analysis of the protocol, especially with regard to implicating ring members without their consent. Third, it would be illuminating to conduct a rigorous, quantitative analysis of the impacts of the ring signature protocol on institutional transparency, disinformation, and public trust in the media, in order to more concretely measure the benefits and drawbacks of the protocol. Finally, we would like to see ring signing and verification functionality incorporated into secure messaging apps and browser extensions. We hope that this will make the ring signatures protocol described in this paper accessible and intuitive to sources, journalists, and the public.

References

Allen, M. (2007, Jan). *Who's that senior administration official?* Retrieved from https://www.politico.com/story/2007/01/whos-that-senior-administration-official-002548

Bender, A., Katz, J., & Morselli, R. (2006). Ring signatures: Stronger definitions, and constructions without random oracles. In *Theory of cryptography conference* (pp. 60–79).

Brady, S. (2012). Keeping secrets: A constitutional examination of encryption regulation in the united states and india. *Ind. Int'l Comp. L. Rev.*, *22*(4), 317.

CPJ. (2013). *The obama administration and the press.* Retrieved from https://cpj.org/reports/2013/10/obama-and-the-press-us-leaks-surveillance-post-911.php

Duffy, M. J., & Freeman, C. P. (2011). Unnamed sources: A utilitarian exploration of their justification and guidelines for limited use. *Journal of Mass Media Ethics*, *26*(4), 297–315. doi: 10.1080/08900523.2011.606006

Goldwasser, S., Micali, S., & Rivest, R. L. (1988). A digital signature scheme secure against adaptive chosen-message attacks. *SIAM Journal on Computing*, *17*(2), 281–308.

HRW. (2016, Jan). *With liberty to monitor all — how large-scale us surveillance is harming journalism, law, and american democracy.* Retrieved from https://www.hrw.org/report/2014/07/28/liberty-monitor-all/how-large-scale-us-surveillance-harming-journalism-law-and

(n.d.). Retrieved from https://www.yourdictionary.com/chilling-effect#law

I am part of the resistance inside the trump administration. (2018, Sep). The New York Times. Retrieved from https://www.nytimes.com/2018/09/05/opinion/trump-white-house-anonymous-resistance.html

Liu, J. K., Wei, V. K., & Wong, D. S. (2004). Linkable spontaneous anonymous group signature for ad hoc groups. In *Australasian conference on information security and privacy* (pp. 325–335).

Meyers, S. (2011, Sep). *Business insider: 'we will grant anonymity to any source at any time for any reason'.* Retrieved from https://www.poynter.org/news/business-insider-we-will-grant-anonymity-any-source-any-time-any-reason

Michael. (2018, Sep). *Anonymous op-ed in new york times causes a stir online and in the white house.* The New York Times. Retrieved from https://www.nytimes.com/2018/09/05/business/media/new-york-times-trump-anonymous.html

Philip. (2018, Jun). *How the times uses anonymous sources.* The New York Times. Retrieved from https://www.nytimes.com/2018/06/14/reader-center/how-the-times-uses-anonymous-sources.html

Risen, J. (2016, Dec). *If donald trump targets journalists, thank obama.* Retrieved from https://www.nytimes.com/2016/12/30/opinion/sunday/if-donald-trump-targets-journalists-thank-obama.html?_r=0

Rivest, R. L., Shamir, A., & Tauman, Y. (2001). How to leak a secret. In *International conference on the theory and application of cryptology and information security* (pp. 552–565).

Savage, C. (2017, Jun). *Intelligence contractor is charged in first leak case under trump.* Retrieved from https://www.nytimes.com/2017/06/05/us/politics/reality-winner-contractor-leaking-russia-nsa.html

SPJ. (n.d.). *Anonymous sources.* Retrieved from https://www.spj.org/ethics-papers-anonymity.asp

U.s. const. ammend v. (n.d.).