Privacy Preserving Reputation Systems based on Blockchain and other Cryptographic Building Blocks: A Survey

Omar Hasan, Lionel Brunie, Elisa Bertino

Privacy Preserving Reputation Systems based on Blockchain and other Cryptographic Building Blocks: A Survey

Omar Hasan[a], Lionel Brunie[a], and Elisa Bertino[b]

[a]University of Lyon, CNRS, INSA-Lyon, LIRIS, UMR5205, F-69621, France
omar.hasan@insa-lyon.fr, lionel.brunie@insa-lyon.fr
[b]Department of Computer Science, Purdue University, IN 47907, USA
bertino@cs.purdue.edu

Abstract

The purpose of a reputation system is to hold the users of a distributed application accountable for their behavior. The reputation of a user is computed as an aggregate of the feedback provided by fellow users in the system. Truthful feedback is clearly a prerequisite for computing a reputation score that accurately represents the behavior of a user. However, it has been observed that users can hesitate in providing truthful feedback, for example, due to the fear of retaliation. Privacy preserving reputation systems enable users to provide feedback in a private and thus uninhibited manner. In this survey, we propose analysis frameworks for privacy preserving reputation systems. We use these analysis frameworks to review and compare the existing systems in the literature. An emphasis is placed on blockchain-based systems as they are a recent significant development in the area. Utilizing blockchain as a building block, privacy preserving reputation systems have been able to provide properties such as trustlessness, transparency, and immutability, which were absent from prior systems. The results of the analysis reveal several insights and directions for future research. These include exploiting blockchain to its full potential to develop truly trustless systems and to implement some important security properties and defenses against common attacks that are so far ignored by a majority of the systems.

1 Introduction

Reputation systems are an essential tool for determining the trustworthiness of users in environments where pre-established trust in users does not exist. Reputation of a target user is computed by aggregating the subjective feedback provided by source users. These are users who have previously interacted with

the target user and have consequently gained personal experience regarding her actions in the context of the given application. It is expected that actions perceived as legitimate will lead to high positive feedback and thus an aggregation of a positive reputation score. Inversely, a target user acting dishonestly will elicit negative feedback resulting in a low reputation score. Any users concerned about the legitimacy of future actions of a potential transacting partner, can consider the computed reputation score of the user as an indication of her trustworthiness. Reputation systems thus assist in holding users accountable for their actions despite the initial absence of trust in the users.

E-commerce marketplaces and sharing economy based platforms are some popular applications where reputation systems are employed. Sites and mobile applications such as `ebay.com`, `airbnb.com`, and `uber.com` are significant examples. Additionally, systems by Liu et al. [1], Azad et al. [2], Bag et al. [3], and Schaub et al. [4] are some of the academic proposals for managing reputation in e-commerce and retail environments. Let's consider Airbnb (`airbnb.com`), which is an online marketplace for vacation rentals. The platform enables independent hosts to offer their private lodgings to guests for short stays. The reputation system of the platform plays a critical role since the guests seeking satisfactory accommodations can only rely on the reputation of the hosts and their offerings stemming from reviews provided by previous guests. Similarly, hosts concerned about lending out their lodgings to well-behaving guests also need to depend on the reputation system.

Another application that relies on reputation systems is mobile participatory sensing, where users sense various environmental conditions with their mobile devices and submit sensing data to a central entity for analysis. Reputation is used to discourage users from furnishing corrupted information. Systems by Jo and Choi [5], Ma et al. [6], and Mousa et al. [7] are examples of reputation systems that target this application area. A related application is participatory sensing in Vehicular Adhoc Networks (VANETs), where vehicles collect and upload information about road conditions. Reputation systems by Zhao et al. [8], Lu et al. [9], and Chen et al. [10] aim to hold the vehicles and their owners accountable for submitting false data. One more notable application area, among several others, that counts on reputation systems is the Internet of Things (IoT). Trusting corrupted devices in the IoT can lead to comprises in network security [11]. Recent systems by Azad et al. [11, 12] are instances of reputation systems serving this application domain.

It has been documented that users may hesitate to provide truthful feedback [13, 14]. Reasons range from fear of retaliation to negative reviews [13, 14] to concerns about revealing sensitive personal information [15]. Returning to the example of Airbnb, we note that its reputation system escrows the feedback until both parties have submitted their opinion. This is done in hopes of preventing tit for tat retribution by the hosts and the guests. However, the truthfulness and the impartiality of user feedback can still be impacted due to the personal nature of the reviews [16]. Hiding the identities of the users has been recommended as a solution [16]. Moreover, it has been observed that the lack of anonymity on Airbnb "causes people to feel pressure to post reviews that lean positive" [17].

Privacy preserving reputation systems are designed to allay the fears of feedback providers by protecting the confidentiality of their individual feedback. The implication is that providing feedback in a private manner encourages the raters and un-inhibits them to submit honest and accurate feedback. Another approach that privacy preserving reputation systems take to motivate users to submit feedback and to do so truthfully is by guaranteeing their anonymity. Operating in an anonymous manner in the system signifies that a third party is unable to attribute sensitive personal information to the user or to profile the user in the long term. Privacy preserving reputation systems are therefore an important category of reputation systems for scenarios where user privacy or anonymity needs to be upheld.

The field of research of privacy preserving reputation systems is fairly mature. All academic reputation systems cited above are in fact privacy preserving. Reputation systems that respect user privacy were first proposed in the mid 2000s. Some notable original works include those by Pavlov et al. [18], Kinateder and Pearson [19], and Dingledine et al. [20], among others. However, privacy preserving reputation systems continue to evolve to cater for emerging application areas, such as Social IoT (Azad et al [11]), Industrial IoT-enabled retail marketing (Liu et al. [1]), and Intercloud (Dou et al. [21]). Moreover, the advent of the blockchain technology has recently fueled further research in this area. The outcome of utilizing blockchain as a building block is privacy preserving reputation systems that offer novel properties such as trustlessness, transparency, and immutability. For example, Schaub et al.'s [4] system does not require users to trust third parties or any fellow users in order to guarantee their security and thus provides trustlessness. This property was absent from prior systems. Another important reason for continued research in the area of privacy preserving reputation systems is the fact that a number of issues still remain open. As we discover in this survey and discuss in Section 10, these issues include lack of implementation of important security properties and lack of defenses against common attacks.

Despite the maturity of the topic, to the best of our knowledge, no comprehensive survey has been conducted so far on privacy preserving reputation systems. Section 11 provides a summary of the related work. We believe that a survey is needed to cultivate a uniform perspective to the rich literature in this area. Moreover, we believe that the current moment is opportune to present this survey due to the recent emergence of systems based on blockchain as well as novel systems for upcoming application domains. In this survey, we analyze 40 privacy preserving reputation systems published between the years 2003 and 2020 inclusive, while placing an emphasis on recent systems based on blockchain.

Reputation systems that respect user privacy have always mostly relied on cryptographic building blocks and their combinations to provide strong security guarantees. These building blocks include Secure Multi-Party Computation (SMPC), secret sharing, homomorphic encryption, zero-knowledge proofs, cryptographic signatures, and others. Blockchain is a recent addition to this repository of cryptographic building blocks utilized by privacy preserving reputation systems. We study blockchain-based systems as well as systems based

on other building blocks and security mechanisms in this survey.

1.1 Contributions

This survey makes the following contributions:

- Identification of the various dimensions of privacy preserving reputation systems. An analysis framework that allows for the decomposition and comparison of privacy preserving reputation systems in a normalized manner.
- Identification of the security requirements of privacy preserving reputation systems that cut across multiple types of such systems.
- Summary of the building blocks utilized by current privacy preserving reputation systems.
- Definition of broad categories of the privacy preserving reputation systems proposed in the literature according to their security mechanisms.
- Fine-grained analysis and comparison of 40 privacy preserving reputation systems using the proposed analysis frameworks.
- Detailed review of several significant and representative privacy preserving reputation systems in the literature.
- Discussion of the analysis results that lead to multiple insights and identification of avenues for future work in this field of research.

1.2 Organization

The rest of the paper is organized as follows. Section 2 develops an analysis framework that delineates the various dimensions of reputation systems. Section 3 identifies the dimensions and the requirements of privacy preserving reputation systems. Section 4 defines two broad categories of privacy preserving reputation systems with respect to their security objectives. Section 5 presents an overview of some of the major building blocks that serve as the foundation for privacy preserving reputation systems. Section 6 defines broad categories of the privacy preserving reputation systems proposed in the literature according to their security mechanisms. Section 7 presents a fine-grained analysis of privacy preserving reputation systems in the literature according to the frameworks established in Sections 2 through 5. Section 8 and Section 9 describe in greater detail some of the systems. Section 10 provides an overall discussion and reveals a number of insights into the field of research. Section 11 summarizes the related work and Section 12 concludes the survey.

2 An Analysis Framework for Reputation Systems

In this section, we develop an analysis framework that identifies the various dimensions of reputation systems. The framework presented in this section is not a novel contribution because several prior works (such as the surveys by Braga et al. [22], Hendrikx et al. [23] and Hoffman et al. [24]) have developed more comprehensive frameworks for reputation systems. However, since privacy preserving reputation systems are fundamentally reputation systems, we need to establish a uniform framework to analyze and compare their non-privacy features as well. The analysis framework for issues specific to privacy is presented in Section 3.

Some fundamental concepts in reputation systems are as follows:

Source User (Rater). A user u is said to be a source user or rater of a user t if u has feedback about t in a given context.

Target User (Ratee). When a source user assigns feedback to a user t, or a user q initiates a query to determine the reputation of user t, the user t is referred to as the target user or the ratee.

Querying User (Querier, Inquirer). When a user q initiates a query to determine the reputation of a user t, the user q is referred to as the querying user, the querier, or the inquirer.

Reputation. The reputation of a target user is any function that aggregates the feedback of its source users. In Section 2.4, we present some possible realizations of the aggregation function.

The analysis framework for reputation systems is graphically represented in Figure 1. In the following sections, we present the various dimensions of reputation systems.

2.1 Architecture

The architecture of a reputation system is one of the key factors in determining how the following activities are conducted: 1) Feedback collection; 2) Feedback aggregation (reputation computation); and 3) Reputation dissemination. We discuss below three architectures for reputation systems: centralized, decentralized, and hybrid.

Centralized. Centralized reputation systems are characterized by the existence of a fully or partially trusted central authority. The central authority receives feedback from users, aggregates it to compute the reputation, and disseminates the reputation scores. One of the benefits of a centralized solution is that it is straightforward to implement. Additionally, the central authority is universally trusted, therefore users can be assured that

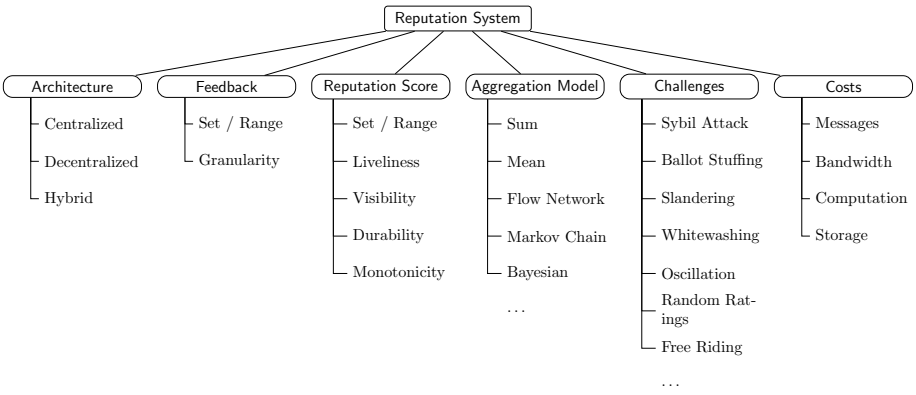

Figure 1: Analysis framework for reputation systems.

the feedback collection, aggregation, and dissemination are being done correctly. However, if the central authority fails, the whole reputation system is compromised. Thus the central authority is a single point of failure and a high-value target for attackers. Centralized reputation systems are also unable to cater for decentralized environments such as peer-to-peer networks, ad-hoc networks, decentralized social networks, etc.

Decentralized. Decentralized reputation systems are suitable for decentralized environments as they do not assume the presence of a central entity. In decentralized reputation systems, a central location for submitting and aggregating feedback, and disseminating reputation does not exist. Feedback is commonly stored locally by the node who generates it, for example in response to its experiences with another party. Computing reputation of an entity in the system requires finding all or a portion of the nodes who carry feedback about that entity. Once the feedback providers have been located, the aggregation may be done at a single location after receiving all feedback, or a more sophisticated protocol may be employed to aggregate the feedback in a distributed manner.

Hybrid. The hybrid architecture merges elements from the centralized and the decentralized architectures. Some activities are carried out in a centralized manner whereas others in a decentralized fashion.

2.2 Properties of Feedback

Set / Range. The set or range that the feedback belongs to, for example, $\{-1, 0, 1\}$, $[0, 1]$.

Granularity. The feedback of a rater may reflect the experience with the ratee for a single given transaction or the feedback may reflect the cumulative experience with the ratee over multiple transactions.

2.3 Properties of Reputation

Set / Range. The set or range that the reputation score belongs to, for example, $\mathbb{R}$, $[0, 1]$.

Liveliness. As noted by Schiffner et al. [25], reputation liveliness implies that a reputation system does not offer users the possibility to reach a final state of reputation in which bad behavior no longer damages their reputation. For example, for a reputation score in the set $\mathbb{R}$, there is no maximum limit, whereas, for a reputation score in the interval $[0, 1]$, the reputation can reach the maximum value of 1. If the reputation system is also monotonic (defined later in this section), the reputation cannot decrease and the user can retain the maximum score forever.

Visibility. The visibility of a reputation score may be global or local. Global visibility implies that all nodes in the system view the same reputation score of a certain entity. Whereas with local visibility, the reputation score available to a subset of the nodes may be different than elsewhere in the system. Local visibility is generally a concern in decentralized or personalized reputation systems, where a different subset of feedback providers may be included for computing the reputation of an entity at different instances.

Durability. Reputation durability refers to the transience of a reputation score. Once a reputation score is computed, it may be stored permanently for subsequent access by nodes through a simple retrieval operation. Recalculation of the score is mandated only when new feedback becomes available. Alternatively, the reputation score may be transient and re-computed every time a node wishes to learn the score. The latter approach requires repeated computation of the reputation, however, it does not require storage of the scores by a trustworthy entity.

Monotonicity. Monotonic reputation implies that the reputation score increments in only one direction. For example, consider a reputation system in which a user can receive integer feedback between 1 and 5 for each transaction, and reputation is considered as the sum of feedback. The reputation in such a reputation system can only increase upwards. The reputation of a user cannot be decremented.

2.4 Feedback Aggregation Models

There are a number of models for aggregating feedback to obtain reputation scores. We describe some interesting models below. A comprehensive survey of feedback aggregation models (also called reputation computation engines) is provided by Jøsang et al. [26].

Sum and Mean Model. One of the most common methods of aggregating feedback to obtain the reputation score is simple summation. The advantage of this approach is that it is very straightforward and easy to

understand for the users of the reputation system. A related method is to compute the reputation score as the mean of the feedback values. Reputation represented as mean has the benefit of being normalized and thus the reputation of different users may be compared objectively.

Flow Network Model. A class of reputation systems (such as the Advogato (advogato.org) [27] reputation system) are constructed using the concept of flow networks. The users are considered as the nodes of a network and the feedback that they assign each other is considered as the flow in the network. The reputation of a node is computed as a function of the flow that the node receives from other nodes. A salient characteristic of such reputation systems is that a node cannot assign more flow to other nodes than it has received itself. This prevents a node from creating multiple pseudonyms for malicious purposes, since the total incoming flow and hence the reputation of the pseudonyms would be only as high as the original node itself. It is assumed in the Advogato reputation system that the amount of flow available in the network is constant and regulated by trustworthy nodes adjacent to the source.

Markov Chain Model. Several reputation systems (such as EigenTrust [28] and PowerTrust [29]) draw on the Markov chain theory. Feedback from one node to another is considered as the probability of transition from the source to the target node. The reputation of a node is computed as the probability of arriving at that node by following random transitions from a known trustworthy node. The reputation systems based on the Markov chain theory also offer the advantage that a malicious node does not benefit from creating multiple pseudonyms for malicious purposes. This is due to the fact that even if the malicious node assigns maximum feedback to each of its pseudonyms, the probability of reaching those pseudonymous nodes from a trustworthy node would be no higher than reaching the original malicious node.

Bayesian Model. The reputation score in a Bayesian reputation system is generally represented by a beta distribution in which the two free parameters α and β correspond to the number of positive and negative feedback respectively. The reputation score is computed by statistically updating the given beta distribution. The observable difference in the statistical properties of fair and unfair ratings enables filtering out unfair ratings [30].

2.5 Challenges faced by Reputation Systems

Reputation systems can be classified by the challenges that they address and their success in resolving them. In this section, we discuss some of the challenges other than privacy that reputation systems have to contend with.

Sybil Attack. The Sybil attack [31] on a reputation system operates as follows: An attacker creates multiple identities in the system in order to gain an

unfair advantage over honest users who own a single identity. The attacker may use its multiple identities to mount attacks including self-promotion and slandering.

Self-Promotion, Ballot Stuffing. Self-promotion is the act of raising one's own reputation through unfair means. Self-promotion may be carried out by a user individually or in collusion with other members of the system. A self-promotion attack is particularly effective in systems where users may assign each other additional feedback after every transaction. Two users may repeatedly transact with each other, and after each transaction assign each other positive feedback. This attack is also known as ballot stuffing, which implies that a user submits more feedback than he is entitled to.

Slandering, Bad-Mouthing. Slandering or bad-mouthing is the act of sabotaging an honest user's reputation by assigning them unwarranted low feedback. Motivation for such an attack may include retaliation, reducing a competitor's reputation, or malicious disruption of services. A slandering attack is highly detrimental to the target user in applications that are sensitive to the presence of even a small amount of low feedback, such as high-value monetary systems.

Whitewashing. A whitewashing attack occurs when a user with negative reputation quits the system and re-enters with a new identity and thus a fresh reputation. A reputation system is vulnerable to the whitewashing attack when: the pseudonyms in the system are not linked to real world identities; quitting the system incurs little or no loss; and creating new pseudonyms is cheap (in terms of limited resources, such as money, human effort, etc.). To mitigate the risk of whitewashing attacks, a reputation system may differentiate users who are newcomers from those who have been in the system for a long time. A user may only be allowed to build his reputation gradually by demonstrating good behavior consistently over a long period of time. This approach lessens the appeal of a whitewashing attack, since a user who re-enters the system with a new identity is not viewed as trustworthy.

Oscillation. In oscillation, an attacker initially builds good reputation in the system and then suddenly shifts behavior to take advantage of honest users who are misled into trusting the attacker due to the good reputation. This attack is advantageous only if the payoff of the attack is greater than the cost of building good reputation. One scenario is that an attacker engages in several low value transactions to accumulate reputation and then reverses its good behavior for a high value transaction. A reputation system may mitigate the risk of oscillation attacks by weighing feedback according to its age in the system.

Random Ratings. In a random ratings attack, an attacker submits randomly generated feedback to the system instead of providing an accurate evalua-

tion of the target's behavior. This attack is advantageous for the attacker in systems where feedback submission is incentivized.

Free Riding. Free riding is more of a user behavioral pattern that is generally detrimental to the system rather than an attack. In free riding, a user doesn't actively participate in the feedback collection part of the system, but still makes requests to learn the reputation of other users. The free riding user thus takes advantage of the system without providing any contribution.

2.6 Costs

The operations of a reputation system, which include feedback collection, feedback aggregation (reputation computation), and reputation dissemination, incur various computational costs. The costs of these operations can be measured as follows: 1) number of messages exchanged; 2) bandwidth consumed; 3) computational resources consumed; and 4) storage required.

3 An Analysis Framework for Privacy Preserving Reputation Systems

In this section, we propose an analysis framework that identifies the common dimensions and requirements of privacy preserving reputation systems. We conduct a fine-grained analysis and comparison of privacy preserving reputation systems in the literature using this framework in Section 7.

Three of the dimensions (adversary, reputation binding, and trust model) are described in the following subsections. Whereas, the other two dimensions are described independently in later sections: In Section 4, we discuss the security objectives of privacy preserving reputation systems proposed in the literature. In section 5, we list the building blocks that serve as the foundation for privacy preserving reputation systems.

The analysis framework for privacy preserving reputation systems is graphically represented in Figure 2.

3.1 Adversary

The goal of a reputation system is to compute the reputation from the inputs of the participants. All participants of the protocol are expected to pursue this and only this goal. An honest participant is one who conforms to this expectation. However, there may exist dishonest participants who have ulterior motives. Those motives may include learning the inputs of other participants, tampering with the output, disrupting the protocol, etc.

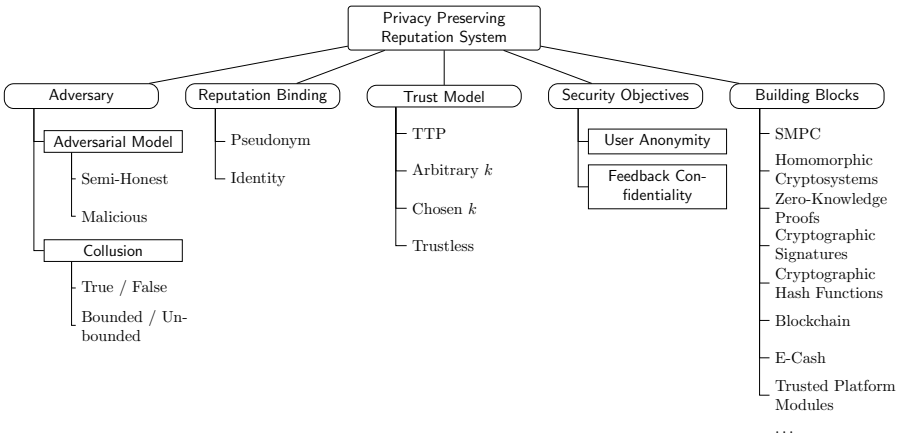

Figure 2: Analysis framework for privacy preserving reputation systems.

3.1.1 Adversarial Model

We list below two standard adversarial models [32] that characterize the behavior of dishonest users. A privacy preserving reputation system is considered secure under one of these models if it can show correctness and meet its privacy requirements under the given model.

Semi-Honest. In the semi-honest model, the users do not deviate from the specified protocol. In other words, they always execute the protocol according to the specifications. The adversary abstains from wiretapping and tampering of the communication channels. However, within these constraints, the adversary passively attempts to learn the inputs of honest users by using intermediate information received during the protocol and any other information that it can gain through other legitimate means.

Malicious. Malicious users are not bound to conform to the protocol. Users under a malicious model may deviate from the protocol as and when they deem necessary. They actively attempt to achieve their objectives. They may participate in extra-protocol activities, devise sophisticated strategies, and exhibit arbitrary behavior. Specifically, malicious users may 1) refuse to participate in the protocol, 2) provide out of range values as their inputs, 3) selectively drop messages that they are supposed to send, 4) prematurely abort the protocol, 5) distort information, and 6) wiretap and tamper with all communication channels. A malicious adversary may have either or both of the following objectives: 1) learn the inputs of honest users, and 2) disrupt the protocol for honest users. The reasons for disrupting the protocol may range from gaining illegitimate advantage over honest users to completely denying the service of the protocol to honest users.

3.1.2 Collusion

A dishonest user may act alone or multiple dishonest users may act in agreement to achieve their ulterior motives. When multiple dishonest users work together, it is referred to as collusion. Privacy preserving reputation systems either consider that collusion can take place between users or consider that collusion does not take place.

Collusion can be bounded or unbounded. Bounded collusion implies that the number of dishonest participants in the system allowed to collude with each other is limited, for example, $\frac{1}{2}$ or $\frac{1}{3}$ of all n participants. Unbounded collusion places no limit on the number of dishonest participants who can collude with each other, thus $n-1$ of the participants can be dishonest and colluding, except for the one honest participant whose privacy needs to be preserved.

3.2 Reputation Binding

A privacy preserving reputation system can be either pseudonym-bound or identity-bound.

In a pseudonym-bound system, the reputation of the user is associated with her pseudonym. If she changes or creates a new pseudonym then she looses her reputation. This can be disadvantageous for several reasons. This implies that reputation is not transferable between a user's multiple pseudonyms. Moreover, a dishonest user can drop a pseudonym with bad reputation and re-enter the system with a new pseudonym and a fresh reputation.

On the other hand, in an identity-bound system, the reputation of a user is bound to her real identity. Even if she changes pseudonyms, she maintains her reputation. This is often made possible by verifying the true identity of a user before issuing a new pseudonym.

3.3 Trust Model

The security and privacy guarantees that users receive in a privacy preserving reputation system often require that they trust certain entities, such as a central authority, or some fellow users in the system. The trust implies a belief of the trusting user that the trusted entity or the trusted fellow users will behave in an expected manner in order to ensure their security and privacy. We identify four different types of trust models that privacy preserving reputation systems are based on.

Trusted Third Party. A Trusted Third Party (TTP) for a set of users is an entity whom every user in the set trusts completely for certain actions. In this model, all users of the system must trust the designated TTP entities in the system. A user in a system who needs to be fully trusted is also considered as a TTP.

Trust on arbitrary k fellow users. A user in the system is required to place her trust in k different fellow users for the security guarantees, where $k \leq$

n, and n is the total number of users participating in the protocol. These k users are selected by the system without taking the user's preferences into account. Thus from the perspective of the user, the set of trusted users is selected arbitrarily. Generally, only a partial level of trust is required in each of the trusted users in this model.

Trust on chosen k fellow users. In this trust model, a user in the system also places her trust in k distinct fellow users. However, these fellow users are chosen by the user herself. The user may select the trusted users based on the level of their subjective trustworthiness in order to maximize the privacy guarantee. This model requires that a user is able to determine the trustworthiness of fellow users and choose accordingly from a pool of available users.

Please note that there is a difference between choosing fellow users for establishing security guarantees versus choosing feedback providers for personalizing the reputation score of the target entity. In the first case, a user chooses fellow users who specifically influence the security and privacy guarantees that she would receive in the reputation system. In the latter case, there is no intentional impact of the selection on the security guarantees. The "Trust on chosen k fellow users" model addresses choosing k users specifically for the purposes of security in the reputation system.

As an example, consider the systems by Hasan et al. [33] and Gudes et al. [34]. In the system by Hasan et al., the selection of k fellow users is made in the context of preserving privacy. The trust model of this system can thus be classified under the chosen k users category. In contrast, in the system by Gudes et al., even though a user selects a subset of fellow users, the system's trust model cannot be classified as the chosen k users model. The reason being that the selection of users in this latter system is made purely for personalizing the reputation score.

Trustless. In the trustless model, the users in a system do not need to trust any entities or any fellow users. Thus, this model does not expect users to have pre-existing trust toward fellow users or entities in the system. The users need to rely solely on the underlying algorithms and protocols of the system in order to receive the security guarantees.

However, we note that even though the users do not need to directly trust any entities or users in this model, there may exist a requirement of trustworthiness for the overall correct and secure functioning of the system. Trustless systems are based primarily on the blockchain technology. As an example, the Bitcoin blockchain requires that a majority of all participants in the system act honestly in order to ensure integrity.

The trustless model may be considered a special case of the "Trust on arbitrary k fellow users" model, where k is at least greater than half of the total number of all participants in the entire system (not just a protocol instance). A blockchain system functions by building consensus among

peers. In case of Bitcoin, if a majority of peers are dishonest, consensus cannot be achieved and the entire system malfunctions. Thus, the breach of the trustworthiness requirement in such systems does not simply threaten the security of a given user but the integrity of the entire system. It is therefore in the collective interest of all honest users in the system to prevent any breach of trustworthiness.

4 Security Objectives of Privacy Preserving Reputation Systems

We have identified two broad categories of privacy preserving reputation systems with respect to their security objectives. The goal of the systems in the first category is to preserve the anonymity of the users. The systems in the second category do not aim to hide the identity of the users but focus on preserving the confidentiality of the feedback that the users provide. The two categories of privacy preserving reputation systems are defined as follows:

1. **User anonymity oriented privacy preserving reputation systems.** The true identity of the users is hidden in these systems. The feedback providers thus remain anonymous. A user is represented in the system by one or more pseudonyms which are unlinkable to his real identity. This setup allows the user to anonymously carry out transactions with others and submit feedback. There is no need to guard the confidentiality of the submitted feedback since the anonymity of the users prevents it from being linked to them.

2. **Feedback confidentiality oriented privacy preserving reputation systems.** These systems do not attempt to hide the identity of the users beyond assigning each user a single pseudonym. Moreover, these systems do not conceal the act of a user assigning feedback to another user. However, the value of the submitted feedback and any other related information is considered private. This type of systems is necessary since complete anonymity is not always possible due to the nature of real world transactions. For example, even if anonymity is preserved online on an e-commerce site, the exchange of physical items sold and bought through the site would reveal the real identities of the participants. Preserving the confidentiality of the feedback values is a practical alternative to enable users to submit truthful feedback without the fear of retaliation.

The security objectives of a privacy preserving reputation system can be further subdivided as those fulfilling privacy and those fulfilling integrity or correctness. The privacy objectives are concerned with hiding information about users, for example, preserving the anonymity of the rater and the ratee. On the other hand, the integrity objectives are aimed toward maintaining the correctness of the functions of the reputation system while preserving the privacy of

the users. For example, integrity objectives include preventing a malicious user from manipulating the reputation aggregation function to forge an unmerited good reputation.

Figure 3 illustrates the classification of the security objectives of privacy preserving reputation systems. In Sections 4.1 and 4.2, we describe several individual security objectives of user anonymity and feedback confidentiality oriented privacy preserving reputation systems respectively. A particular reputation system may pursue a few or more of these objectives depending on the stringency of its security requirements.

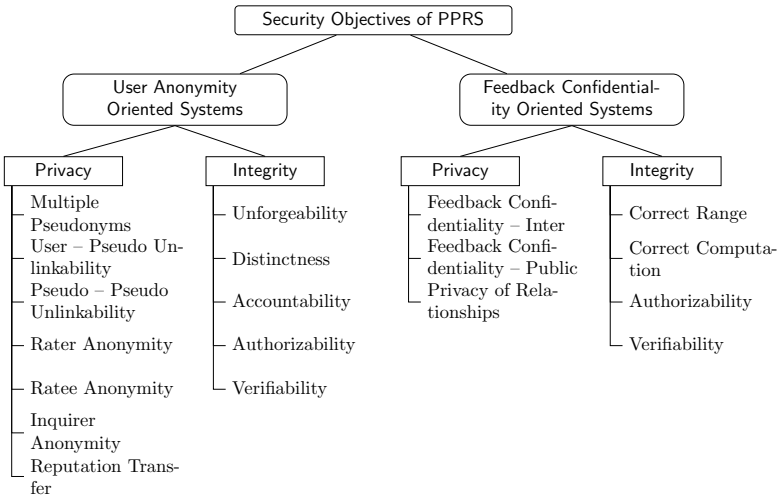

Figure 3: Security objectives of privacy preserving reputation systems.

4.1 User Anonymity Oriented Privacy Preserving Reputation Systems

4.1.1 Privacy Objectives

Multiple Pseudonyms. A user is able to assume multiple pseudonyms in the system. As noted by Anwar and Greer [35, 36], the variation in the pseudonyms of a user may be on a per context or a per transaction basis. In the first case, a user may adopt a different pseudonym for each context in the system, for example, a tutor could use different pseudonyms for different subjects in an e-learning system. Alternatively, a user may choose a different pseudonym for each transaction in the system.

User-Pseudonym Unlinkability. User-pseudonym unlinkability implies that the true identity of a user is not linkable to any pseudonym that he uses in the system. Androulaki et al. [37] identify this requirement as follows: Given a pseudonym P that does not belong to a corrupted party, the

adversary can learn which peer owns P no better than guessing at random among all non-corrupted peers that appear consistent with P.

Pseudonym-Pseudonym Unlinkability. Pseudonym-pseudonym unlinkability implies that two different pseudonyms that belong to the same user cannot be linked to each other. The adversary is unable to tell whether two given pseudonyms belong to the same user. This property is specified by Androulaki et al. [37] as follows: Given two pseudonyms P_1, P_2 that do not belong to corrupted parties, the adversary has no advantage in telling whether P_1, P_2 belong to the same peer or not. This requirement should hold as long as there are at least two non-corrupted peers who appear consistent with both P_1 and P_2 (because if there is only one such uncorrupted peer, clearly both pseudonyms belong to the same one).

Rater Anonymity. A user is able to rate another user without his true identity being revealed. The purpose of rating anonymously is to prevent the adversary from linking the rater to his interaction with the ratee and the rating that he submitted. Schiffner et al. [25] state this property as follows: A pseudonym P_1 that interacted with a ratee R should not be linkable to the pseudonym P_2 that rated R.

Ratee Anonymity. A user is able to receive a rating without his real identity being disclosed. A ratee may not wish to be associated with his past transactions and ratings since they could influence the ratings for his future transactions. According to Schiffner et al. [25], this property implies that a ratee R can use a different pseudonym for each transaction.

Inquirer Anonymity. A user is able to inquire about the reputation of another user, however, others are not able to learn whose reputation he is querying or even the fact that he is inquiring about another user's reputation. Users wish to query the reputation of other users anonymously in order to prevent the adversary from compiling a profile of their interactions and interests.

Reputation Transfer and Aggregation. A user is able to transfer reputation among multiple pseudonyms that he owns without letting the adversary draw associations between these pseudonyms. Consequently, a user is able to aggregate the reputation of his multiple pseudonyms into the reputation of one pseudonym.

4.1.2 Integrity Objectives

Reputation Unforgeability. A user is unable to show reputation higher than the cumulative reputation of his pseudonyms. A user is also unable to borrow good reputation from another user.

Distinctness. It is possible to prove that the reputation of a target user is an aggregate of votes or feedback that come from distinct users while simultaneously hiding the identities of those users. The advantage of this

property is that one or a few dishonest users are not able to submit multiple votes or feedback (ballot stuffing) for artificially raising the reputation of the target user.

Accountability. If and only if a user commits a predefined adversarial act, such as ballot stuffing, then his pseudonym becomes linkable to his real identity. This property ensures that anonymous users are still accountable for adversarial actions.

The properties of authorizability and verifiability are discussed in Section 4.3.

4.2 Feedback Confidentiality Oriented Privacy Preserving Reputation Systems

4.2.1 Privacy Objectives

No Inference from Intermediate Information. This property requires that a rating assigned by a rater to a ratee is never revealed to any other party including the ratee. The system must protect the confidentiality of the feedback such that the feedback is neither divulged explicitly nor inferred from any intermediate information gained by the adversary during a reputation query. The system may define the confidentiality of the feedback as deterministic or probabilistic. In the first case, the adversary is unable to learn any information about the feedback. However, in the latter case of probabilistic confidentiality, the amount of information leakage depends on certain variables, such as the number of raters, the reputation score, etc.

No Inference from Public Information. The reputation score of any user is by definition public and any other user in the system is authorized to learn this score. The issue is that a dishonest user may use this public information to derive the private feedback of honest users. For example, in a basic additive reputation system, the adversary simply needs to observe the reputation score before and after the latest user submits his feedback to learn its value. The requirement of confidentiality of feedback, with no inference from public information, implies that the adversary is unable to learn information about the feedback even from publicly available information.

Privacy of Relationships. A user may have relationships with multiple users in the system. These other users may include fellow users who have rated the same ratees. The relationships between the users could be in various contexts, for example, the context of trust in preserving each others privacy. This requirement implies that information about the relationships of a rater is not revealed during the course of a reputation query. This information includes the amount of trust that the rater has in the fellow users.

4.2.2 Integrity Objectives

No Out of Range Feedback. A dishonest user is unable to submit out of range feedback. A dishonest user may take advantage of the fact that the feedback is confidential and submit out of range feedback in order to mount an attack such as bad mouthing or ballot stuffing. A system enforcing this property does not permit out of range feedback even though the feedback is hidden.

No Incorrect Computations. A dishonest user is unable to carry out incorrect computations. A reputation query may require users to perform certain computations, for example, the summation of some values. This property requires that a dishonest user is unable to submit erroneous results for these computations.

4.3 Integrity Objectives Common to Both Types of Privacy Preserving Reputation Systems

Authorizability of Ratings. The requirement of authorizability of ratings implies that only the users who have had a transaction with the ratee are allowed to rate him. This property prevents users who have not transacted with a ratee from assigning him feedback and thus possibly reduces the impact of attacks such as bad mouthing and self promotion.

Verifiability by Ratee. The requirement of verifiability by ratee as identified by Kerschbaum [38] suggests that a ratee R should be able to identify all published feedback linked to his identity and verify that they are related to a recorded transaction and the correct transaction partners. Moreover, a ratee R should be able to identify all published feedback linked to his identity and verify that the inquirer has computed its reputation score according to them.

5 Building Blocks for Privacy Preserving Reputation Systems

In order to achieve their security objectives, privacy preserving reputation systems utilize various building blocks, which are generally cryptographic in nature. These building blocks include secure multi-party computation, homomorphic cryptosystems, zero-knowledge proofs, blockchain, etc. In this section, we present an overview of some of the major building blocks that serve as the foundation for privacy preserving reputation systems.

5.1 Secure Multi-Party Computation

Secure Multi-Party Computation (SMPC) is the study of protocols that take inputs from distributed entities and aggregate them to produce outputs, while

preserving the confidentiality of the inputs.

One of the well-known secure multi-party computation protocols is secure sum [39], which takes inputs from entities and computes their sum. The protocol assumes that there are three or more sites and there is no collusion between them. It is also assumed that the value to be computed, $v = \sum_{l=1}^{s} v_l$ lies in the range $[0..m]$. The sites are numbered as $1 \ldots s$. Site 1 generates a random number R uniformly chosen from $[0..m]$. It then sends $R + v_1 \ mod \ m$ to site 2, where v_1 is site 1's local input. Site 2 does not learn any information about v_1 since $R + v_1 \ mod \ m$ is distributed uniformly across the range $[0..m]$ due to R. For sites $l = 2 \ldots s-1$, the protocol proceeds as follows: Site l receives $V = R + \sum_{j=1}^{l-1} v_j \ mod \ m$. Site l learns nothing since the value is distributed uniformly across $[0..m]$. Site l computes $R + \sum_{j=1}^{l} v_j \ mod \ m = (v_l + V) \ mod \ m$. Site l then sends this value to site $l+1$. Eventually, site s also performs the above step. Site s sends the result back to site 1, who subtracts R from it to obtain the sum. Site 1 does not learn any of the private values due to the uniform distribution of the received result over the range $[0..m]$.

The protocol may be used to compute reputation as the sum or the mean of the feedback values provided as private inputs by the participants of the protocol. However, the above protocol in its original form suffers from problems such as no resistance to collusion and no defense against out of range feedback.

5.2 Homomorphic Cryptosystems

Let $E_u(.)$ denote an encryption function with the public key PK_u of user u in an asymmetric cryptosystem $\mathcal{C}$. The cryptosystem $\mathcal{C}$ is said to be additive homomorphic if we can compute $E_u(x+y)$, given only $E_u(x)$, $E_u(y)$, and PK_u. In other words, a cryptosystem is additive homomorphic if we can compute the encryption of the sum of two plaintexts, given only their ciphertexts and the encrypting public key.

The Paillier cryptosystem [40] is a well-known additive homomorphic cryptosystem. Similarly, a multiplicative homomorphic cryptosystem such as the ElGamal Cryptosystem [41] allows computation of the encryption of the product of two plaintexts from their ciphertexts and the encrypting public key.

Homomorphic cryptosystems are useful for privacy preserving reputation systems as they allow operations on multiple source feedback values in their encrypted form. The operations can be carried out while obfuscating individual values.

5.3 Zero-Knowledge Proofs

A zero-knowledge proof [42] is an interactive proof that allows a prover to convince a verifier that a statement is true without revealing any information other than the fact that the statement is valid.

As an example, consider a prover who knows an RSA modulus n and its two large prime factors p and q. A verifier knows only n. Factoring n is considered intractable therefore the verifier cannot learn p and q. An interactive proof

would be zero-knowledge if it allows the prover to convince the verifier that he knows the factors of n without revealing any information about p and q.

Zero-knowledge proofs are used in the context of feedback confidentiality to guarantee the integrity of the obfuscated values by mitigating the risk of out of range feedback. In anonymity preserving reputation systems, zero-knowledge proofs may be used to provide a proof that a user's pseudonym is valid without revealing the identity of the user.

5.4 Cryptographic Signatures

A blind signature scheme (for example, the scheme introduced by Chaum [43]) is a cryptographic signature scheme in which an entity signs a message provided by a user, however, the entity does not learn the content of the message.

Group signatures are another type of cryptographic signatures. As discussed by Michalas and Komninos [44], group signatures allow a group of users to create a signature with the following properties: 1) Only members of the given group can sign a message representing that group; 2) The receiver is able to verify that a signed message originates from that group; 3) The signature reveals the group but not the identity of the individual signer; and 4) The real identity of a user, who behaves maliciously in a predefined manner, can be revealed by collectively opening the signature.

Cryptographic signatures are used in privacy preserving reputation systems in several ways. Some systems in the literature use signatures as proof of identity, providing a way to ensure that a pseudonym belongs to a user while keeping the identity of the user secret. Signatures are also used during the transfer of reputation between two pseudonyms. Finally, some systems use signatures to ensure the integrity of the source's feedback.

An application of signatures in privacy preserving reputation systems is in the form of anonymous credential systems. Using this building block, an organization grants credentials to pseudonymous identities of users. Verifiers are able to verify the authenticity of credentials in the possession of users. However, neither the organization or the verifier is able to link the credential to the true identity of the user.

5.5 Cryptographic Hash Functions

A cryptographic hash function maps data of arbitrary size to data of fixed size. The key properties of a cryptographic hash function include collision-resistance, hiding and puzzle-friendliness.

Cryptographic hash functions in the privacy preserving reputation systems literature are mostly used to produce verifiable short term pseudonyms, or a transaction identifier that allows a source's feedback to be linked to a specific transaction.

5.6 Blockchain

A blockchain is a distributed data structure that was introduced as the foundation of the Bitcoin cryptocurrency. A blockchain can be considered a public distributed ledger that is composed of a set of blocks linked by cryptographic hashes. The blocks are chronologically ordered. Each block comprises of the record of a set of transactions or operations that have recently taken place between the users. Through an implicit consensus mechanism, all users eventually agree on the state of the public distributed ledger. A new block is proposed for appending to the blockchain by competing users. The user who wins the right to append the new block by first solving a cryptographic puzzle receives an award in order to incentivize the continuity of the blockchain. This Proof of Work (PoW) mechanism is specific to Bitcoin, however, several other consensus mechanisms have been proposed as well. Examples include Proof of Stake (PoS) used by Ethereum, and Proof of Authority (PoA) used by VeChain. The new block and the user's right to append it are verified by the peers. Only correctly formed blocks are accepted thus guaranteeing the security of the blockchain.

A blockchain offers some advantageous properties that can be utilized by systems using it as a building block. A blockchain stores an immutable record of information, which means that the information once recorded is not modifiable and its integrity and persistence are guaranteed. Additionally, a blockchain provides transparency since all information is public and each block of information is appended in an auditable manner. Moreover, a blockchain offers decentralization since there is no trusted third party or any super nodes involved in its maintenance. Every node in the network is able to verify the integrity of the blockchain as well as compete toward earning the right to appending a new block. This decentralization also leads to the property called trustlessness, which enables users to cooperate and collaborate without needing to trust each other.

Certain blockchain systems, such as Ethereum, build on the principles of blockchain to implement the smart contract technology. A smart contract is a set of programmed rules that are agreed upon by a group of users in advance. The correct execution of the program and the enforcement of the rules is then guaranteed by all nodes in the system who are maintaining the blockchain. Smart contracts allow users who do not have any pre-existing trust in each other to be able to conduct transactions with guaranteed compliance to the mutually agreed upon set of rules. They can rely on the underlying blockchain system to prevent deceitful behavior from any of the parties.

Privacy preserving reputation systems can benefit from blockchains in multiple ways. A blockchain can be used for its immutability, transparency, and auditability properties to create a reputation system that enables users to verify the integrity of the computation of the reputation scores. The decentralized system by Schiedermeier et al. [45] is an example of such utilization of blockchain. Moreover, a privacy preserving reputation system can use smart contracts to transparently enforce the rules for updating the reputation of a user. This is the case in the reputation framework for participatory sensing systems by Jo

and Choi [5], where a smart contract manages the reputation of a participant user based on their sensing data and the corresponding feedback.

5.7 E-Cash

Some privacy preserving reputation systems (such as the one proposed by Androulaki et al. [37]) with user anonymity as their security objective, use E-cash as one of the building blocks. E-cash, a predecessor of blockchain-based cryptocurrencies, is a digital currency first proposed by Chaum [43, 46]. As discussed by Belenkiy et al. [47], E-cash provides the following features:

Anonymity. It is impossible to trace an e-coin (the monetary unit of e-cash) to the user who spent it. This property holds even when the bank (a central entity who issues the e-coins) is the attacker.

Unforgeability. The only exception to the anonymity property is that e-cash does not guarantee the anonymity of a user who tries to double-spend an e-coin. In this case, the bank can learn the identity of the dishonest user. A forged e-coin allows the bank to trace down the user who forged it.

Fungibility. A user can use the e-coins received for services provided as payment for services received from any other user in the system.

Endorsed e-cash [48] adds the following property to e-cash:

Fair Exchange. Fair exchange means that a buyer gets the item only if the seller gets paid and vice versa.

E-cash protocols are often used in privacy preserving reputation systems in the following manner: the votes of the raters are represented by e-coins of an E-cash system. The quantity of the coins that a target has received is considered its reputation. A trusted third party acting as a bank is needed to enforce the integrity of the system.

5.8 Trusted Platform Modules

A Trusted Platform (TP) [49, 50] is described as a secure computing platform that preserves the privacy of the user by providing the following three functionalities:

Protected Storage. Data on the TP is protected from unauthorized access.

Integrity. The TP can prove that it is running only the authorized software and no malicious code.

Anonymity. The TP can demonstrate that it is a genuine TP without revealing the identity of the user. The TP uses a pseudonym attested by a PKI Certification Authority (CA).

A Trusted Platform comprises of a Trusted Platform Module (TPM), which is a hardware device with cryptographic functions that enable the various security functionalities of the TP. The TPM is unforgeable and tamper-resistant.

Kinateder and Pearson [19] implement a privacy preserving reputation system, in which a TPM at a node enables to demonstrate that it is a legitimate member of the system without disclosing its true identity.

6 Categorization of Privacy Preserving Reputation Systems according to their Security Mechanisms

In this section, we identify broad categories of the privacy preserving reputation systems proposed in the literature. These categories are based on the general mechanisms that these systems rely on in order to guarantee privacy and other critical security properties, for example, authorizability, verifiability, etc.

We also briefly discuss the contributions of the systems that belong to each of these categories. Each system is further analyzed in depth and compared in Section 7. Five of the listed blockchain-based systems are discussed in detail in Section 8. One or two interesting systems are selected from each of the other categories and discussed in detail in Section 9.

Please note that these categories are not mutually exclusive and a system may belong to multiple of these categories. For example, the system by Schiedermeier et al. [45] can belong to the category of blockchain-based systems as well as SMPC-based systems. However, we place a system under a single category based on its main novel idea. For example, even though Schiedermeier et al.'s work uses SMPC, the novel idea and the main contribution is rather the use of a blockchain-based public ledger as the sole communication medium between the parties of the SMPC protocol. The blockchain-based protocol provides transparency and verifiability properties that are usually missing from SMPC-only systems. The system by Schiedermeier et al. is therefore categorized as a blockchain-based system.

In this survey, we have included the systems that we are aware of in this field of research as well as those discovered using the following approach. We searched for articles on Google Scholar published during the period of 2000 to 2020. The search phrases included the keyword 'reputation' along with one of the keywords 'privacy', 'anonymous', and 'anonymity'. For each relevant article found, we studied its list of references to find other potential systems. Moreover, we also looked at the article's "Cited by ..." list on Google Scholar to discover later relevant papers that cite the given article.

6.1 Blockchain-based Systems

These systems rely on a blockchain or smart contracts as an integral building block for achieving their security objectives.

Schaub et al. [4] introduced the first blockchain-based trustless privacy preserving reputation system. The system does not need to rely on trusted third parties, arbitrary trusted nodes, or subjective trust relationships in order to guarantee security. Using blinded tokens issued by service providers, raters anonymously submit feedback, which is recorded on a public immutable blockchain. Issuing a token requires spending the system's cryptocurrency, which provides an incentive to mine and maintain the blockchain and also discourages ballot-stuffing. Bazin et al. [51] present a system, which in addition to protecting rater privacy, enables retrieval of a self-reported reputation score directly from the target service provider. The validity of the reputation score is verifiable and only a constant number of messages need to be exchanged for its retrieval.

Azad et al. [2, 11] propose privacy preserving reputation systems for online marketplaces and for the Social Internet of Things environment. Self-enforcing computation is a property of their latter system, which implies that the computation process is independent of any trusted third party and it allows verification of the integrity of the scores in an autonomous and public manner. Bag et al. [3], describe a system for computing personalized global reputation of a target, which considers only the feedback from a set of trusted participants. This is done without disclosing the identities of the members of the trusted set and their feedback. The systems by Azad et al. and Bag et al. rely on a public bulletin board for communication, which according to the authors may be realized by a blockchain.

Dou et al. [21] propose a distributed trust evaluation protocol with privacy protection for the Intercloud environment. A distinctive feature of the protocol is that it can continue to function even if some of the feedback providers go offline. Lu et al. [9] present a privacy preserving trust model based on blockchain for vehicular adhoc networks. Vehicles can anonymously submit alerts about traffic conditions and neighboring vehicles can provide feedback about the validity of the alerts. The anonymous reputation of a vehicle reflects the feedback received regarding its contributions. Owiyo et al. [52] propose a decentralized privacy preserving reputation system based on blockchain that is claimed to provide low transaction overheads. Jo and Choi [5] describe a blockchain-based privacy preserving reputation framework for participatory sensing systems. The system includes a smart contract that manages the reputation of a participant based on their sensing data and the corresponding feedback. The smart contract and the underlying blockchain enable transparency and public auditability of the reputation scores.

Liu et al. [1] present an anonymous reputation system for retail marketing in the Industrial Internet of Things environment. The system, which also uses smart contracts on a Proof of Stake blockchain as a building block, is able to provide transparency and public verifiability under the malicious adversarial model. Schiedermeier et al. [45] describe a protocol for holding referendums in trustless networks, which can also serve as a reputation protocol. The protocol combines secure multi-party computation with a blockchain as the unique channel for communication between the parties. The protocol ensures trans-

parency, that is, maintaining a public trace of all operations performed and the information exchanged among the participants. Moreover, any participant is able to autonomously verify the correctness of the outcome of the referendum. Zhao et al. [8] propose a privacy preserving reputation management system that takes advantage of blockchain technology in the resource-constrained environment of mobile crowdsensing. The global reputation scores are updated by a smart contract based on the average of all feedback. The system overcomes the challenge of user dynamics, that is, frequent user turnover, by including a delegation protocol.

6.2 SMPC-based Systems

These systems use feedback score as direct evidence from witnesses to compute a reputation score. Their goal is to obfuscate the feedback score of the witnesses from the querier as well as from fellow witnesses. These systems use Secure Multi-Party Computation to achieve their goal. The reputation systems in this category focus primarily on feedback confidentiality as their security objective.

Pavlov et al. [18] introduced SMPC-based privacy preserving reputation systems by proposing a number of protocols for decentralized additive reputation systems. Two of their protocols are secure under the semi-honest and the malicious adversarial models respectively. The protocols draw their strength from witness selection schemes, which guarantee the inclusion of a certain number of honest witnesses as participants. Gudes et al. [34] and Gal-Oz et al. [53] present several schemes that augment their Knots reputation system [54] with privacy preserving features. A defining characteristic of the Knots reputation model is the notion of subjective reputation. The reputation of a target member is computed by each querying member using a different set of feedback, thus the reputation is subjective for each querying member. Nithyanand and Raman's system [55] complements an SMPC mechanism for privacy with a fuzzy technique and an Ordered Weighted Average (OWA) operator in order to compute local as well as global reputation scores.

Hasan et al. [33, 56] present a system that operates under the more demanding malicious adversarial model and offers the chosen k trust model (discussed in Section 3.3) instead of the usual arbitrary k trust model for privacy preservation. Dimitriou and Michalas [57, 58] describe a decentralized privacy respecting scheme that is formally shown to be resistant to collusion against up to $n-1$ malicious participants. Dolev et al. [59, 60] propose SMPC-based reputation schemes that are more efficient than the previous ones in terms of the number of messages exchanged. Their schemes privately compute reputation scores with a communication overhead of $O(n)$ messages, where n is the number of participants in the protocol. Clark et al. [61] present a dynamic privacy preserving decentralized reputation system. They specifically address the problem of the dynamicity of the nodes in a network. Nodes may frequently leave along with their feedback, which then becomes unavailable for reputation computation in a decentralized manner. Clark et al. propose a privacy preserving reputation information delegation protocol to counter this problem.

6.3 Token-based Systems

These systems are a type of privacy preserving reputation systems in which a cryptographic token is issued to a pseudonymous user participating in a transaction. The token is implemented using a blind signature or another scheme. The token is issued either by a central entity (called the bank in the system by Androulaki et al. [37]) or directly by the ratee to the rater (as in the system by Kerschbaum [38]). A variation of the following approach is then employed in order to credit the ratee with a reputation point while preserving the privacy of the token depositing user. The token is deposited by the user to an account maintained by the central entity using a different pseudonym or even their real identity. The blinded nature of the token unlinks the user from the initial pseudonym while assuring the central entity of the legitimacy of the deposit. The number and the value of the tokens deposited reflects the reputation of the ratee.

The system by Androulaki et al. [37] addresses the difficulties outlined by Dingledine et al. [20] for building reputation systems in anonymous user networks. Androulaki et al.'s system achieves: 1) unlinkability between a pseudonym and the identity of its user; 2) no double-awarding or forging of a token; 3) no false accusations of forgery; and 4) non-transferability of reputation, that is, a user cannot borrow reputation from another user. The system by Kerschbaum [38] builds on the blinded token idea to achieve feedback confidentiality while enforcing the property of verifiability. Schiffner et al. [25, 62] improve upon Androulaki et al.'s work by introducing systems that support the properties of liveliness and non-monotonicity.

Zhang et al. [63] propose a reputation system that preserves the privacy of feedback providers and resists Sybil attacks. The system is based on the Camenisch and Lysyanskaya (CL) signature scheme. Busom et al. [64] describe a privacy preserving reputation system based on Chaum-Pedersen blind signatures that allows users to anonymously submit text feedback about a target entity. Fellow users can in turn anonymously endorse a text feedback that they find helpful. The system thus encourages honest feedback. Moreover, the system offers a privileged status for users who earn sufficient endorsements thus also incentivizing feedback submission.

6.4 Proxy-based Systems

These systems aim to maintain privacy through the use of a trusted third party as a proxy between the feedback providers and the reputation querier. The proxy may forward the anonymized feedback scores to the querier or the proxy may compute the aggregated reputation and only report that to the querier. Additionally, the querier and the feedback providers may interact directly, however, in this case, a feedback provider is generally issued an anonymous identity or an encryption key by the proxy to protect their privacy. The proxy may be composed of one or several central entities. Usually, the architecture of these systems comprises of one to three central entities that are considered not

to collude with each other in order to guarantee security. The proxy may be considered partially or fully trusted.

Ries et al. [65] propose an approach for privacy preserving computation of trust. A key contribution of this approach is that in addition to computing reputation based on encrypted private feedback, the querier can also evaluate the trustworthiness of the feedback providers. Petrlic et al. [66] propose a reputation management system that focuses on privacy (anonymity in reputation retrieval, and anonymity in rating) as well as robustness (authorization, authentication, integrity, and accuracy). A semi-honest Reputation Provider (RP) entity serves as an intermediary between the raters and the service providers. The RP manages the reputation of the service providers and helps enforce some of the above listed security objectives.

Mousa et al. [7] present PrivaSense, a privacy preserving reputation system for mobile participatory sensing applications. The system implements a sequence of registration and authentication phases orchestrated by independent central servers that ensure participants' anonymity and improve the system's resilience against Sybil and replay attacks. Ma et al. [6] propose a privacy preserving reputation management system for edge computing enhanced mobile crowdsensing. The architecture comprises of a Central Manager (CM), a Reputation Manager (RM), and a Central Authority (CA). Participants submit sensing data in homomorphic encrypted form. The encrypted deviation of a participant's data from the aggregated result is computed and the RM updates reputation according to the deviation.

6.5 Signature-based Systems

Inspired by cryptographic digital signatures and group signature schemes, Benthencourt et al. [67] propose a new cryptographic framework called signatures of reputation. In a scheme based on this framework, the verification of the signature of a user reveals her reputation instead of revealing her identity. This is in contrast to a conventional signature scheme where the verification of the signature of a user results in the confirmation of the identity of the user associated with the corresponding public key.

Guo et al. [68] build upon the notion of signatures of reputation to propose a fine-grained attribute-based privacy preserving reputation system. The system enables users to rate each other's attributes instead of real identities. The signature verification process provides authenticity of the reputation value of a user for a given attribute. Bethencourt et al.'s system is improved by the work of Anceaume et al. [69] and Lajoie-Mazenc et al. [70], who implement non-monotonic signature-based reputation systems. Whereas, Bethencourt et al.'s system can only support monotonic reputation.

Chen et al. [10] present a privacy and reputation-aware announcement scheme for vehicular adhoc networks where vehicles can report road conditions. The scheme is based on the Boneh-Boyen-Shacham (BBS) short group signatures. The scheme overcomes the problem of having to establish a secure channel for reputation score retrieval in prior systems.

6.6 Transitory Pseudonym-based Systems

Transitory pseudonym-based systems aim to obfuscate a user's identity by assigning them multiple short-term pseudonyms. The focus is on how to make the multiple pseudonyms of a user unlinkable with the user as well as with one another. Moreover, how to transfer reputation from one pseudonym to another while preventing observation and profiling is also addressed.

One of the first systems in this category is RuP (Reputation using Pseudonyms) by Miranda and Rodrigues [71]. In their system, a user is identified by a certified pseudonym that is valid only for a predefined time slot. The certified pseudonyms are issued by a TTP called Pseudonym Certification Authority (PCA). However, the link between the real identity of the user and the pseudonym is hidden from the PCA as well. The system also includes a scheme based on blind signatures that allows a user to transfer their reputation associated with an old pseudonym to a new one, without disclosing the link between them or their real identity. Another early work in this category is by Steinbrecher [72]. Their system enables simultaneous use of multiple pseudonyms by a user and permits them to regularly change their pseudonyms to achieve anonymity. To prevent an adversary from linking new and old pseudonyms, the system suggests using a set of non-colluding trustworthy third parties who make incremental changes to the pseudonym of the user.

Anceaume et al. [73] propose a privacy preserving distributed reputation mechanism. The system allows users to themselves generate pseudonyms in order to achieve anonymity. They introduce the concept of mailboxes, which are agents that replicate anonymous feedback, in order to provide resistance against network dynamicity and user misbehavior. Christin et al. [74] present IncogniSense, another improvement on the RuP scheme, which is claimed to achieve better protection against reputation manipulation and reduce the cryptographic overhead for the client.

6.7 Other Systems

In this category, we include systems that propose unique approaches and therefore cannot be placed in the above defined categories.

Kinateder and Pearson [19] introduced one of the very earliest privacy oriented decentralized reputation systems. The system requires a Trusted Platform Module (TPM) chip at each agent, which enables an agent to demonstrate that it is a valid agent and a legitimate member of the system without disclosing its true identity. This permits the agent to provide feedback anonymously. Bo et al. [75] present a privacy preserving reputation system, which offers incentives to users for feedback submission. A user who anonymously submits feedback can also anonymously receive a discount token (an incentive) from the ratee. The architecture of the system comprises of a Card Issuer (CI) entity and a Registration Center (RC) entity that are responsible for issuing smart cards and anonymous identities to users respectively.

6.8 Additional Literature

Due to the depth and breadth of research in this area, there are a number of other privacy preserving reputation systems in the literature that we have not been able to treat in detail in Section 7. Moreover, there are papers that do not present specific systems but discuss the problem of reputation and privacy in general. We summarize some of these two types of works below in chronological order.

Ismail et al. [76, 77] present a decentralized privacy preserving reputation system that enforces accountability regarding the legitimacy of the feedback provided. Cvrček et al. [78] develop a specification of requirements for a trust model for evidence-based reputation systems supporting pseudonymity. Voss et al. [79] present a decentralized system that is based on similar lines as Kinateder et al. [19]. They suggest using smart cards as the trusted hardware modules. A later system by Kinateder et al. [80] avoids the hardware modules, however, it requires an anonymous routing infrastructure at the network level. Hao et al. [81] present a scheme in which a user's anonymity is achieved by changing pseudonyms with the help of a TTP, while preventing disclosure of linkable information to the TTP and other users. They show that their scheme reduces the TTP's RSA decryption operations and message overhead by more than half as compared to the RuP scheme by Miranda and Rodrigues [71].

Nin et al. [82] present a reputation system that computes the reputation of a user based on whether she correctly follows a set of rules for making trust-based access control decisions. Fellow users are able to audit the decisions made by a user and provide feedback accordingly. The privacy objective of the reputation system is to keep the trust relationships between the users private. The anonymization is derived through the multiplicative homomorphic property of the ElGamal encryption scheme. Mármol et al. [83] describe TRIMS, a privacy aware trust and reputation model for identity management systems. The system addresses the problems that arise when a domain needs to decide whether to exchange information with another domain to provide a service to one of its users. According to the authors, this is one of the first approaches dealing with trust and reputation management in a multi-domain scenario.

Zhang et al. [84] present STARS, a software component that can serve as an add-on to an underlying non privacy preserving reputation system to achieve anonymity and traceability. Kellermann et al. [85] present a privacy respecting reputation system for wiki users. The system allows to assess the expertise and the reliability of authors contributing to a wiki in order to foster trust in the wiki content. Goodrich and Kerschbaum [86] introduce a privacy enhanced reputation-feedback method, in which two transacting parties provide feedback about each other to an escrow. Their system keeps the feedback escrowed and thus private and publishes only the updated reputation scores. A randomized feedback sampling mechanism provides privacy of individual feedback despite immediate publishing of the reputation scores.

Hasan et al. present a number of privacy preserving reputation protocols [56, 87, 88] that operate under the semi-honest adversarial model. The protocols

have the advantage of providing more efficient computation than prior SMPC-based protocols for the semi-honest adversarial model. Huang et al. [89] propose a privacy preserving reputation system for participatory sensing. The system includes an anonymization scheme based on the concept of k-anonymity, which prevents adversaries from de-anonymizing users while minimizing the impact on the usability of the application outputs. Au and Kapadia [90] present PERM, a reputation-based anonymous blacklisting system. One of the key building blocks used by the system is a signature scheme called BBS+ proposed by Au et al. [91], which is used in the system for binding scores to transaction identifiers.

Wang et al. [92] propose ARTSense, a framework that addresses the problem of anonymous reputation in mobile sensing. Their solution consists of a privacy preserving provenance model, a data trust assessment scheme and an anonymous reputation management protocol. The scheme does not require a trusted third party and both positive and negative reputation updates can be applied. Clauß et al. [93] define the concept of a k-anonymous reputation system, where a user remains k-anonymous when obtaining a new independent random pseudonym, even when other users may choose to retain their pseudonyms. Aldini et al. [94] explore the trade off between trust, privacy, and cost in incentive-based networks.

Brangewitz et al. [95] discuss a reputation system that enables the incorporation of reputation information into markets of composed services while preserving the privacy of customers who provide feedback. The system caters to the On-The-Fly (OTF) computing environment, where the goal is automated composition of flexibly combinable services. Michalas and Komninos [44] describe Lord of the Sense (LotS), a privacy preserving reputation system for participatory sensing applications. The system uses group signatures as a key building block. A user is able to anonymously submit a sensing report. Other users can vote about the validity of the submitted report. The system maintains the reputation score of a user while respecting their anonymity.

Zhang et al. [96] describe a privacy friendly weighted reputation aggregation protocol secure against malicious adversaries in cloud services. The problem addressed is as follows: A server has a vector that comprises of weights for the feedback of each of the feedback providers. The protocol is considered privacy friendly if the server and the feedback providers can cooperate to compute the weighted reputation while keeping all respective inputs private.

7 Fine-Grained Analysis and Comparison of Privacy Preserving Reputation Systems

In this section, we conduct fine-grained analysis of privacy preserving reputation systems in the literature according to the frameworks established in Sections 2 through 5. The analysis is presented in the form of Tables 1 through 6. The tables also permit side by side comparison of the systems.

We have analyzed 40 privacy preserving reputation systems in depth and

summarized their properties in the given tables. We report information about the systems as gleaned from the articles. In case of multiple variants of a system presented in the same article, we have selected the variant that provides the strongest security guarantees. The systems are grouped in the tables according to the category of their security mechanisms. The categories are ordered by the number of included systems and then alphabetically. Under each category, the systems are ordered chronologically to allow observation of the evolution of the systems.

Table 1 identifies the fundamental characteristics of each of the reputation systems according to the analysis framework developed in Section 2. The architecture of the systems and the properties of their feedback and reputation are presented.

Table 2 and Table 3 present the security related fundamentals of user anonymity and feedback confidentiality oriented systems respectively. In accordance with the analysis framework for privacy preserving reputation systems formulated in Section 3, the properties reported include the adversarial model, the extent of collusion resistance, reputation binding, the trust model, and the main security building blocks. Multiple adversarial models are listed if a scheme uses different adversarial models for different entities, for example, semi-honest for the server, and malicious for the users. We note strong collusion resistance if t out of the n users in the protocol must collude to breach security, where $t < n$, and t is variable. For example, $t = \frac{1}{2}n$, or $t = \frac{1}{3}n$. Alternatively, we note partial collusion resistance if a constant number of colluding entities, for example, two partially trusted colluding servers, are able to breach security. Multiple trust models are noted for the systems that rely on different models for their different security properties. The aggregation model is stated as open where the system is not constrained to one specific function.

The details of the security objectives of user anonymity and feedback confidentiality oriented systems are presented in Table 4 and Table 5 respectively. As discussed in Section 4, the security objectives of privacy preserving reputation systems include those aiming to enforce privacy and those targeting integrity or correctness.

The robustness of the reputation systems against the challenges discussed in Section 2.5 is summarized in Table 6.

8 Blockchain-based Privacy Preserving Reputation Systems

In this section, we describe in greater detail some of the blockchain-based privacy preserving reputation systems in the literature. We focus on their security mechanisms as well as their use of blockchain. Moreover, we highlight salient features that require further explanation or those that are not evident from the analysis in Section 7.

Table 1: Fundamentals.

		Feedback		Reputation					
System	**Architecture**	**Set / Range**	**Granularity**	**Set / Range**	**Liveliness**	**Visibility**	**Durability**	**Non-Monotonicity**	**Aggregation Model**
Blockchain-based Systems									
Schaub et al. 2016	D	$\mathbb{Z}$	S	$\mathbb{R}$	●	G	○	●	Open
Bazin et al. 2017	D	$\mathbb{Z}$	S	$\mathbb{R}$	●	G	●	●	Open
Azad et al. 2018	D	$\{-,+\}$	S	$\mathbb{Z}$	●	G	○	●	Beta reputation
Bag et al. 2018	D	$\{0,1\}$	M	$[1,10]$	○	L	○	●	Mean
Dou et al. 2018	D		S			G	●		Weighted mean
Lu et al. 2018	C	$\{-1,0,1\}$, $[0,1]$	S	$\mathbb{R}$, $[0,1]$	●	G	●	●	Polynomial
Owiyo et al. 2018	D		S			G	●		Open
Jo and Choi 2019	H	$\{-1,1\}$	S	$\mathbb{R}$	○	G	○	●	Sum
Liu et al. 2019	C	$[1,10]$	S	$\mathbb{N}$	●	G	●	○	Sum
Schiedermeier et al. 2019	D	$\{-1,1\}$	S	$\mathbb{Z}$	●	G	●	●	Sum
Zhao et al. 2019	C	$[0,1]$	S	$[0,1]$	●	G	●	●	Mean
Azad et al. 2020	D	$\{-1,1\}$	S	$\mathbb{Z}$	●	G	●	●	Weighted sum
SMPC-based Systems									
Pavlov et al. 2004	D	$\mathbb{R}$	M	$\mathbb{R}$, $[0,1]$	●	L	○	●	Sum, beta reputation
Gudes et al. 2009	D	$\mathbb{R}$	M	$\mathbb{R}$	●	L	○	●	Weighted sum, mean
Nithyanand and Raman 2009	D	$\mathbb{R}, \{0,1\}$	M	$\mathbb{R}$	●	L	○	●	Ordered weighted average
Gal-Oz et al. 2010	D	$\mathbb{R}$	M	$\mathbb{R}$	●	L	○	●	Weighted sum, mean
Hasan et al. 2013	D	$[0,1]$	M	$\mathbb{R}, [0,1]$	●	G	○	●	Sum, mean
Dimitriou and Michalas 2014	D	$\mathbb{Z}$	M	$\mathbb{Z}$	●	G	○	○	Sum
Dolev et al. 2014	D	$\{1,2,\ldots,10\}$	M	$\mathbb{R}$	●	L	○	●	Weighted mean
Clark et al. 2016	D	$[0,v_{max}]$	M	$[0,v_{max}]$	●	L	○	●	Mean
Token-based Systems									
Androulaki et al. 2008	C	$\{0,1\}$	S	$\mathbb{Z}$	○	G	●	○	Sum
Kerschbaum 2009	C	$\{0,1\}$	S	$[0,1]$	●	G	●	●	Beta reputation
Schiffner et al. 2009	C	$\{-1,1\}$	S	$\mathbb{Z}$	●	G	●	●	Sum
Schiffner et al. 2011	C	$\{-,+\}$	S	$\mathbb{R}$	●	G	●	●	Open
Zhang et al. 2014	H		S	$\mathbb{R}$	●	G	●	●	Open
Busom et al. 2017	C	Text	S		●	G	●	●	Union
Proxy-based Systems									
Ries et al. 2011	C	$\{0,1\}$	M	$[0,1]$	●	L	○	●	Beta reputation
Petrlic et al. 2014	C	Vector, $\{0,1\}$	S	$\mathbb{Z}$	●	G	●		Sum
Mousa et al. 2017	C	$\{-1,0,1\}$, $[0,1]$	S	$[0,1]$	●	G	●	●	Bounded sum
Ma et al. 2018	C	$[0,1]$	M	$[0,1]$	●	G	●	●	Weighted mean
Signature-based Systems									
Bethencourt et al. 2010	H	$\{0,1\}$	S	$\mathbb{Z}$	●	G	●	○	Sum
Guo et al. 2013	C	$\{-1,1\}$	S	$\mathbb{Z}$	●	G	●	●	Sum
Lajoie-Mazenc et al. 2015	H	$\{-,+\}$, $\mathbb{Z}$	S	$\mathbb{R}$	●	G	●	●	Open
Chen et al. 2016	C		S	$\{0,1,\ldots,m\}$		G	●		Time discount function
Transitory Pseudonym-based Systems									
Miranda and Rodrigues 2006	C		S		●	G	●	●	Open
Steinbrecher 2006	C		S		●	G	●	●	Open
Anceaume et al. 2013	D	$[0,1]$	S	$[0,1]$	●	G		●	Beta reputation
Christin et al. 2013	C		S		●	G	●	●	Open
Other Systems									
Kinateder and Pearson 2003	D	$[0,1]$	S	$\mathbb{R}$	●	L	○	●	Open
Bo et al. 2007	H		S		●	G	●	●	Open

Legend

C – D – H	Centralized – Decentralized – Hybrid	●	Property satisfied
S – M	Single – Multiple	○	Property not satisfied
G – L	Global – Local		Property not specified or not applicable

Table 2: User Anonymity Oriented Systems – Security Fundamentals and Building Blocks.

System	Adversarial Model	Collusion Resistance	Reputation Binding	Trust Model	Building Blocks
Blockchain-based Systems					
Schaub et al. 2016	M	●	P	Trustless	Okamoto / Chaum blind signatures, PoS blockchain
Bazin et al. 2017	M	●	P	A-k, TTP	Merkle trees, blind signatures, non-interactive zero-knowledge proofs, blockchain
Dou et al. 2018	SH, M	●	P	A-k, TTP	Additive homomorphic encryption, verifiable secret sharing, blockchain for feedback storage
Lu et al. 2018	SH, M	◒	I	TTP	Merkle trees, digital certificates, blockchain
Owiyo et al. 2018	SH		P		SMPC, blind signatures, blockchain
Jo and Choi 2019	SH, M	◒	I	TTP	Group signatures, blind signatures, blockchain, smart contracts
Liu et al. 2019	M	●	I	A-k, TTP	PS signature, bulletproof system, non-interactive zero-knowledge proofs, PoS blockchain, smart contracts
Token-based Systems					
Androulaki et al. 2008	SH, M	●	I	A-k, TTP	E-cash, anonymous credential system, blind signatures
Schiffner et al. 2009	SH, M	●	I	A-k, TTP	E-cash, cryptographic signatures, one-show credentials
Schiffner et al. 2011	SH, M	●	I	A-k, TTP	Symmetric key encryption, homomorphic encryption, DC-Net, Diffie-Hellman key exchange
Zhang et al. 2014	SH, M	●	I	TTP	Bilinear maps, Camenisch and Lysyanskaya (CL) signatures, Pedersen commitment, non-interactive zero-knowledge proofs
Busom et al. 2017	SH, M	●	I	TTP	Chaum-Pedersen zero-knowledge proofs, Chaum-Pedersen blind signatures, verifiable secret sharing, oblivious transfer
Proxy-based Systems					
Petrlic et al. 2014	SH, M	◒	I	TTP	Paillier additive homomorphic encryption, zero-knowledge proofs
Mousa et al. 2017	SH, M	◒	I	TTP	Digital certificates
Signature-based Systems					
Bethencourt et al. 2010	SH, M	◒	I	TTP	Homomorphic encryption, selective-tag weakly CCA-secure encryption, zero-knowledge proofs, one-time signatures
Guo et al. 2013	SH, M	◒	I	TTP	Boneh-Boyen signature scheme, homomorphic encryption, selective-tag encryption, Groth-Sahai non-interactive proofs
Lajoie-Mazenc et al. 2015	SH, M	●	I	A-k, TTP	Verifiable secret sharing, non-interactive zero-knowledge proofs, anonymous proxy signatures, SXDH commitments
Chen et al. 2016	SH, M	◒	P	TTP	Boneh-Boyen-Shacham (BBS) short group signature scheme
Transitory Pseudonym-based Systems					
Miranda and Rodrigues 2006	SH, M	◒	I	TTP	Cryptographic signatures, blind signatures
Steinbrecher 2006	SH, M	◒	I	TTP	Identity management, cryptographic credentials, cryptographic signatures
Anceaume et al. 2013	M	●	I	A-k, TTP	Overlay network, Distributed Hash Tables (DHTs), cryptographic commitments
Christin et al. 2013	SH, M	◒	I	TTP	Cryptographic signatures, blind signatures
Other Systems					
Kinateder and Pearson 2003	SH, M	◒	I	TTP	Trusted Platform Module (TPM), cryptographic signatures
Bo et al. 2007	SH, M	◒	I	TTP	Smart cards, cryptographic signatures, hash chain, zero-knowledge proof of possession

Legend

SH – M	Semi-Honest – Malicious
I – P	Identity – Pseudonym
A-k – C-k – TTP	Arbitrary k – Chosen k – Trusted Third Party
●	Strong resistance to collusion
◒	Partial resistance to collusion
○	Weak or no resistance to collusion
	Collusion resistance not specified or not applicable

Table 3: Feedback Confidentiality Oriented Systems – Security Fundamentals and Building Blocks.

System	Adversarial Model	Collusion Resistance	Reputation Binding	Trust Model	Building Blocks
Blockchain-based Systems					
Azad et al. 2018	SH, M	●	P	A-k	Homomorphic encryption, non-interactive zero-knowledge proofs, public bulletin board (may be implemented by a blockchain)
Bag et al. 2018	M	●	P	A-k	SMPC, homomorphic encryption, zero-knowledge proofs, Schnorr signature protocol, public bulletin board (may be implemented by a blockchain)
Schiedermeier et al. 2019	M	●	P	A-k	SMPC, secret sharing, homomorphic encryption, blockchain
Zhao et al. 2019	SH, M	●	P	TTP	SMPC, additive secret sharing, blockchain, smart contracts
Azad et al. 2020	M	●	P	A-k	SMPC, homomorphic encryption, zero-knowledge proofs, public bulletin board (may be implemented by a blockchain)
SMPC-based Systems					
Pavlov et al. 2004	M	●	P	A-k	SMPC, Pederson verifiable secret sharing scheme, discrete-log commitment, zero-knowledge proofs
Gudes et al. 2009	SH	◒	P	A-k	SMPC
Nithyanand and Raman 2009	SH	◒	P	A-k	SMPC, Paillier additive homomorphic encryption
Gal-Oz et al. 2010	SH	●	P	A-k	SMPC, semantically-secure public-key encryption, homomorphic encryption
Hasan et al. 2013	M	●	P	C-k	SMPC, Paillier additive homomorphic encryption, non-interactive zero-knowledge proofs
Dimitriou and Michalas 2014	M	●	P	A-k	SMPC, Paillier additive homomorphic encryption, non-interactive zero-knowledge proofs
Dolev et al. 2014	M	●	P	A-k	SMPC, Paillier additive homomorphic encryption, Polhig-Hellman commutative encryption, ElGamal encryption
Clark et al. 2016	SH	●	P	C-k	SMPC, secret sharing, digital signatures
Token-based Systems					
Kerschbaum 2009	SH, M	◒	I	A-k, TTP	Homomorphic encryption, cryptographic pairings, zero-knowledge proofs
Proxy-based Systems					
Ries et al. 2011	SH, M	◒	P	TTP	Homomorphic encryption, zero-knowledge proofs
Ma et al. 2018	SH	◒	P	TTP	Somewhat-homomorphic encryption, cloud

Legend

SH – M	Semi-Honest – Malicious
I – P	Identity – Pseudonym
A-k – C-k – TTP	Arbitrary k – Chosen k – Trusted Third Party
●	Strong resistance to collusion
◒	Partial resistance to collusion
○	Weak or no resistance to collusion
	Collusion resistance not specified or not applicable

Table 4: User Anonymity Oriented Systems – Security Objectives.

	Privacy							Integrity				
System	**Multiple Pseudonyms**	**User-Pseudo Unlinkability**	**Pseudo-Pseudo Unlinkability**	**Rater Anonymity**	**Ratee Anonymity**	**Inquirer Anonymity**	**Reputation Transfer**	**Unforgeability**	**Distinctness**	**Accountability**	**Authorizability**	**Verifiability**
Blockchain-based Systems												
Schaub et al. 2016	●	●	●	●	○	●	○	●	○	○	●	●
Bazin et al. 2017	●	●	●	●	○	●	○	●	○	○	●	●
Dou et al. 2018	○	●		●	○	○	○	●	○	○	●	○
Lu et al. 2018	●	●	●	●	●		○	●	○	●	○	○
Owiyo et al. 2018	●	●	●	●	○	○	○	●	○		●	●
Jo and Choi 2019	○	●	○	●	○	○	○	●	○	●	○	●
Liu et al. 2019	○	●		●	○			●		●	●	●
Token-based Systems												
Androulaki et al. 2008	●	●	●	●	●	●	●	●	○	●	○	○
Schiffner et al. 2009	●	●	●	●	●	●	●	●	○	●	●	○
Schiffner et al. 2011	●	●	●	●	●	●	●	●	○		●	○
Zhang et al. 2014	●	●	●	●	○	○		●		●	●	
Busom et al. 2017	●	●	●	●	○		○	●			●	
Proxy-based Systems												
Petrlic et al. 2014	●	●	●	●	○	●	○	●			●	●
Mousa et al. 2017	●	●	●		●	○	●	●		○	●	○
Signature-based Systems												
Bethencourt et al. 2010	●	●	●	●	●		●	●	●			
Guo et al. 2013	●	●	●	●	●			●				
Lajoie-Mazenc et al. 2015	●	●	●	●	●	●	●	●	●		●	●
Chen et al. 2016	●	●	●	●	○		○	●		●	○	○
Transitory Pseudonym-based Systems												
Miranda and Rodrigues 2006	●	●	●	○	●	○	●	●	○	○	○	○
Steinbrecher 2006	●	●		○	●	○		●	○		●	
Anceaume et al. 2013	●	●	●	●	○	○		●	○	○	●	●
Christin et al. 2013	●	●	●	○	●	○	●	●			●	
Other Systems												
Kinateder and Pearson 2003	●	●	●	●	●	○		●	○	●	○	○
Bo et al. 2007	●	●	●	●	○	○	○		○		○	○

Legend

●	Property satisfied
◒	Property partially satisfied
○	Property not satisfied
	Property not specified or not applicable

Table 5: Feedback Confidentiality Oriented Systems – Security Objectives.

	Privacy			Integrity			
System	**Confidentiality (Intermediate Info)**	**Confidentiality (Public Info)**	**Privacy of Relationships**	**Correct Range**	**Correct Computation**	**Authorizability**	**Verifiability**
Blockchain-based Systems							
Azad et al. 2018	●	◒		●	●	●	●
Bag et al. 2018	●	◒	●	●	●	○	●
Schiedermeier et al. 2019	●	○		◒	●	◒	●
Zhao et al. 2019	●	●		◒	●	●	
Azad et al. 2020	●	○	○	●	●	○	●
SMPC-based Systems							
Pavlov et al. 2004	●	○		●	●	○	○
Gudes et al. 2009	●	○	◒	●	●	○	○
Nithyanand and Raman 2009	●	○		●	●	○	○
Gal-Oz et al. 2010	●	○	◒	●	●	○	○
Hasan et al. 2013	●	◒	○	●	●	○	○
Dimitriou and Michalas 2014	●	◒		●	●	○	○
Dolev et al. 2014	●	○		●	◒	○	○
Clark et al. 2016	●	○	○	●	●	○	○
Token-based Systems							
Kerschbaum 2009	●	◒		●	●	●	●
Proxy-based Systems							
Ries et al. 2011	●	◒		●	●	○	○
Ma et al. 2018	●	●		◒	●	●	○

Legend

●	Property satisfied
◒	Property partially satisfied
○	Property not satisfied
	Property not specified or not applicable

Table 6: Measures Against Challenges.

System	Sybil Attack	Ballot Stuffing	Slandering	Whitewashing	Oscillation	Random Ratings	Free Riding
Blockchain-based Systems							
Schaub et al. 2016	●	◒	◒	◒	○	○	●
Bazin et al. 2017	●	◒	○	●	◒	○	○
Azad et al. 2018	◒	◒	◒	◒	○	○	○
Bag et al. 2018	◒	◒	○	○	○	○	○
Dou et al. 2018	◒	○	○	◒	◒	○	○
Lu et al. 2018	●	◒	●	●	●	●	
Owiyo et al. 2018	○	◒	◒	○	○	◒	○
Jo and Choi 2019	●	●	◒	●	○		
Liu et al. 2019	●	◒	○	○	○	○	○
Schiedermeier et al. 2019	◒	◒	○	○	○	○	○
Zhao et al. 2019	○	◒	◒	○	○	○	●
Azad et al. 2020	○	○	○	◒	○	○	○
SMPC-Based Systems							
Pavlov et al. 2004	○	◒	◒	○	○	○	○
Gudes et al. 2009	◒	◒	◒	○	○	○	○
Nithyanand and Raman 2009	◒	◒	◒	○	○	○	○
Gal-Oz et al. 2010	◒	◒	◒	○	○	○	○
Hasan et al. 2013	○	○	○	○	○	○	○
Dimitriou and Michalas 2014	○	○	○	○	○	○	○
Dolev et al. 2014	○	◒	◒	○	○	◒	○
Clark et al. 2016	○	○	○	○	○	○	○
Token-based Systems							
Androulaki et al. 2008	●	○	◒	◒	○	●	○
Kerschbaum 2009	●	◒	◒	●	○	○	●
Schiffner et al. 2009	●	○	◒	●	○	●	○
Schiffner et al. 2011	●	◒	◒	●	○	◒	○
Zhang et al. 2014	●		●	●	○	◒	○
Busom et al. 2017	●	●	◒	◒	○	◒	○
Proxy-based Systems							
Ries et al. 2011	◒	◒	◒	○	○	◒	○
Petrlic et al. 2014	●	◒	◒	●	●	◒	○
Mousa et al. 2017	●	●	●	●	◒	●	●
Ma et al. 2018	○	●	●	○	●	●	●
Signature-based Systems							
Bethencourt et al. 2010	●	●	●	◒	◒	●	●
Guo et al. 2013	●		◒	◒	◒		
Lajoie-Mazenc et al. 2015	●	◒	◒	●	○	○	○
Chen et al. 2016		◒	○	◒	◒	○	○
Transitory Pseudonym-based Systems							
Miranda and Rodrigues 2006	●	○	○	◒	○	○	
Steinbrecher 2006		○	○		○	○	
Anceaume et al. 2013	●	●	◒	●	◒	◒	○
Christin et al. 2013	●	●	●	●	○	●	
Other Systems							
Kinateder and Pearson 2003	◒	◒	○	◒	○	○	○
Bo et al. 2007	●	○	◒	○	○	○	○

Legend

●	Strong or explicit measures
◒	Partial or implicit measures
○	Weak or no measures
	Measures not specified or not applicable

8.1 Schaub et al. 2016

Schaub et al. [4] design a reputation system for real-world e-commerce applications. It is therefore assumed that a customer c's real identity will be disclosed to the service provider SP during a transaction. Instead of complete anonymity, the system emphasizes user anonymity specifically for the feedback submission stage. The system requires unlinkability of the user to the rating, unlinkability of the rating to the transaction, and unlinkability of the rating to other ratings by the same user. These properties ensure that c can submit a rating without identification by the SP, and thus achieve user anonymity for feedback submission.

In order to receive a rating from a customer, the service provider SP is required to spend a certain amount of coins of the native cryptocurrency of the system. This approach is advantageous in a number of ways. It discourages the ballot stuffing attack, since the SP will need to spend coins proportional to the number of artificial ratings. Moreover, the cryptocurrency allows the system to incentivize mining its blockchain by rewarding the creation of new blocks with coins. The service providers can either mine the coins themselves or they may acquire the coins on open market from other miners. The system thus ensures the continuity of the blockchain through incentivized mining, which in turn also ensures the trustlessness property of the system.

A customer c can compute the reputation of a service provider SP by aggregating the ratings about the SP available in the public blockchain of the system. The ratings are aggregation function agnostic, therefore any aggregation function of the customer's choosing can be used for computing the reputation. Moreover, the user can consult text reviews submitted along with the numerical ratings. If the reputation is acceptable, c generates a one time private/public key pair specifically for the transaction with SP.

After the transaction has taken place, c asks SP for a blinded token authenticating the transaction. The SP can issue a token to c if the SP has at least n coins available on his address on the blockchain. The n coins are necessary, since this amount will be deducted from the SP upon submission of a rating by the customer. c then verifies the token and unblinds it, breaking the link between himself and the transaction. When c wishes to rate SP, he broadcasts a message containing the SP's address, the unblinded token, and his rating. A miner of the blockchain who creates a new block then verifies and includes this rating in the block, which is eventually appended to the blockchain.

In addition to ballot stuffing, the system also offers resistance against bad mouthing. In order to submit a feedback about SP, a real transaction needs to take place and its cost needs to be paid to the service provider. It is therefore not possible for an adversary to submit frivolous negative feedback about the service provider without incurring a cost. A Sybil attack is not feasible for either the customer or the service provider since owning multiple addresses in the system does not provide any apparent adversarial advantage. The system is also fairly immune to free riding because (other than potentially generating some network traffic) consulting the blockchain for computing the reputation of a

service provider does not directly draw any resources from the raters or the ratee. Moreover, the system is robust against out of range feedback since feedback is public and is verified by miners before integration into the blockchain.

8.2 Bag et al. 2018

Bag et al. [3] present PrivRep, a privacy aware decentralized and personalized reputation system for electronic marketplaces. The system computes a personalized reputation score of a business entity by taking into account only the trust scores from a set of personally trusted users. This is done so without disclosing neither the identities of participants in the trusted set nor their trust scores.

The architecture of PrivRep comprises of the raters, the marketplace, and a Public Bulletin Board (PBB). Although, not explicitly stated by the authors, the public bulletin board described in the paper lends itself well to implementation by a blockchain. In a more recent paper [11] by the same authors, they do describe a blockchain as "essentially a public bulletin board with distributed data storage and computing power", which "hence can be used in our system to realize the PBB".

The feedback providers homomorphically encrypt their rating scores and publish them on the public bulletin board. The feedback providers also publish non-interactive zero-knowledge proofs to demonstrate that the encrypted rating scores lie within the correct range. The reputation engine, which is operated by the owner of the marketplace, runs a secure multi-party computation protocol to compute personalized reputation scores. The reputation engine considers feedback from only personally trusted sources. The feedback providers do not learn whether their submitted scores are included or discarded in the computation of a particular reputation score. The set of trusted participants is constituted by the reputation engine.

The system is shown to be secure under the malicious adversarial model. The adversary may collude with up to $\Delta - 2$ users, where Δ is the number of trusted feedback providers in the protocol. Δ is less than n, which is the size of the set of all feedback providers in the protocol. Privacy is guaranteed if there are at least 2 honest users who provide different feedback. The trust model in this system is arbitrary k. The Δ users in a protocol are selected by the reputation engine. The privacy of the users depends on that set of Δ users. The system provides partial resistance to Sybil attacks and ballot stuffing since the reputation engine is able to select trusted feedback providers for the computation of reputation.

8.3 Jo and Choi 2019

Jo and Choi [5] present BPRF, a blockchain-based privacy preserving reputation framework for participatory sensing systems. The system has two concurrent goals: 1) protecting the privacy of users who submit sensing data; and 2) ensuring data trustworthiness by managing the reputation of users in the context of the reliability of the data submitted. A participating user is able to submit a

sensing report anonymously and in an unlinkable manner. However, fellow users (e.g., those in the same location) can independently observe the environment and can then submit feedback about the veracity of the sensing report. The architecture of BPRF comprises of a smart contract on a blockchain that manages the reputation of a participant user based on their sensing data and the corresponding feedback. A reliable sensing report earns the participating user a reward token, whereas a disputed one earns a penalty token. Reputation values of users are transparently managed by the smart contract and the blockchain and are thus publicly auditable.

Although, the reputation is managed by a smart contract on a decentralized blockchain, the system overall has a hybrid architecture due to the inclusion of centralized trusted parties, such as the application servers and a trace server. An application server employs a group signature algorithm to maintain groups corresponding to different reputation levels. Membership of a user in a group represents association with the reputation of that group. Group signatures are used for a group member to send sensing reports without revealing identity, yet demonstrating reputation. Reputation is not transferable between members of different groups.

Reputation is identity bound because users are authenticated using a PKI. If they exit and re-enter the system, they can be recognized and re-assigned their existing reputation. This mechanism provides strong resistance against Sybil attacks and whitewashing. BPRF considers the users to operate under the malicious adversarial model. However, the majority of users is considered to be honest. Moreover, the relaxed semi-honest model is assumed for the servers and they are required not to collude with each other. An application server and the trace server may collaborate to reveal the identity of a misbehaving user, thus providing accountability. The system provides protection against out of range feedback since a trusted application server receives feedback directly.

8.4 Liu et al. 2019

Liu et al. [1] propose a reputation system that preserves user anonymity in a retail marketing environment. The architecture of the system comprises of: retailers whose reputation is managed by the system; consumers who transact with the retailers and provide rating scores; an Identity Management entity (IDM) that issues unique identities and credentials to the retailers and the consumers; and a Proof of Stake (PoS) blockchain.

The design goals of the system include: 1) Bounded confidentiality – Even though a rating score provided by a consumer is kept private, the consumer is unable to submit a rating score that falls out of a predefined range. 2) Conditional anonymity – The anonymity of a consumer is guaranteed for operations such as providing a rating score. However, the IDM is able to retrieve the true identity of a consumer in case of misbehavior. 3) Unforgeability – Consumers are unable to forge credentials issued by the IDM and rating tokens issued by retailers. 4) Confined unlinkability – An adversary cannot observe whether two valid rating scores for two different retailers come from the same consumer.

Yet, the rating scores can be linked to the consumer in case he submits multiple scores for the same transaction. 5) Transparency – Rating score submission and reputation computation is transparent and publicly verifiable.

The system operates as follows: Retailers and consumers must register themselves with the IDM using their true identity. The IDM issues anonymous identity credentials to consumers upon registration. A consumer can then transact with a retailer using their anonymous credential and an anonymous payment channel. After the transaction, the retailer issues an anonymous rating token to the consumer. The IDM constitutes a committee of retailers for the rating generation and verification process. The consumer chooses a rating and encrypts it using the public keys of the committee members. The consumer then constructs a zero-knowledge proof of correctness of the rating score. Additionally, the consumer constructs zero-knowledge proofs of possession of a valid credential and a valid rating token. The committee of retailers receives the encrypted rating score and the corresponding zero-knowledge proofs. After verifying its correctness, the committee is able to aggregate the newly submitted rating with the reputation score of the target retailer, while maintaining its confidentiality. The system also enables the committee to detect repeat ratings (ballot stuffing). The committee notifies the IDM in case of misbehavior, which in turn can reveal the identity of the misbehaving consumer.

The rating generation, verification, and aggregation operations take place on the PoS blockchain through smart contracts. This allows the system to provide transparency and public auditability. In order to breach confidentiality, either all committee members or the slot leader (the participant who creates a block on the chain for a given time slot) must collude. A user needs to trust the committee of retailers therefore the arbitrary k trust model applies. Additionally, the IDM is a centralized trusted third party. The system is secure under the malicious adversarial model.

8.5 Schiedermeier et al. 2019

Schiedermeier et al. [45] describe a protocol for holding referendums in trustless networks. The protocol is a secure multi-party computation protocol assisted by a blockchain that serves as a public communication channel among the participants. A referendum protocol can serve as a reputation protocol where the subject of the referendum is considered to be the ratee and the voters are considered to be the raters. The key objectives of the protocol are as follows: 1) confidentiality of the votes; 2) transparency, that is, maintaining a public trace of all operations performed and the information exchanged among the participants; 3) outcome verifiability, that is, any participant is able to autonomously verify the correctness of the outcome of the referendum; and 4) immutability of proceedings, that is, all published information regarding the execution of an instance of the protocol is persisted and accessible permanently.

The participants of the protocol comprise of: 1) an initiator who initiates a referendum and defines its parameters such as the referendum subject and the list of voters (identified by their public keys); 2) voters, who submit their votes;

and 3) workers, who perform intermediate computations for the execution of the protocol. In order to vote, a voter generates n secret shares of his vote, which are homomorphically encrypted with the public keys of the n workers respectively. The shares are published on a blockchain to be retrieved by the workers. After the expiration of the voting phase, each worker aggregates the shares encrypted with her key. The worker does not gain access to the private votes because she does not have access to sufficient number of decrypted shares of any voter. The intermediate results are also placed on the blockchain by the workers. Any querier can then aggregate the intermediate results to determine the final result.

The protocol is analyzed to be secure against a number of threats posed by malicious adversaries. Considering that the protocol uses a t out of n secret sharing scheme, collusion would be possible between up to $t-1$ workers. The authors discuss some heuristics for minimizing the risk of collusion. The initiator of the protocol authorizes the pseudonymous users that participate in the referendum. The protocol therefore provides partial resistance to Sybil attacks and ballot stuffing. An arbitrary set of workers need to be trusted by a voter therefore the arbitrary k trust model applies. Other aspects of the protocol (such as, information storage on the blockchain) are trustless.

9 Privacy Preserving Reputation Systems in the Literature

In this section, we discuss in detail one or two systems in the literature for each of the categories identified in Section 6. Blockchain-based systems are described in the previous section.

9.1 SMPC-based Systems

9.1.1 Hasan et al. 2013

Hasan et al. [33] present Malicious-k-shares, a decentralized additive privacy preserving reputation protocol based on SMPC, which is secure under the malicious adversarial model. The adversarial agents in this model aim to learn private information as well as to disrupt the protocol. This paper introduces the chosen k trust model instead of the prior arbitrary k trust model for privacy preservation.

In the Malicious-k-shares protocol, an agent is required to partially trust on only k fellow feedback providers in order to preserve its privacy. Experimental results in the paper obtained using real trust graphs demonstrate that a high majority of agents can find k sufficiently trustworthy agents in a set of $n-1$ fellow feedback providers such that k is small compared to $n-1$. This idea leads to a protocol that requires only $O(n + log\ N)$ messages, where n and N are the number of agents in the protocol and the environment respectively. This approach improves on prior approaches (as proposed by Gudes et al. [34]

and Pavlov et al. [18] for decentralized additive privacy preserving reputation protocols) where an agent is required to partially trust on all arbitrary $n-1$ fellow feedback providers to preserve its privacy, which results in a high communication complexity of $O(n^3+N)$ messages. Moreover, before submitting their feedback, agents in the Malicious-k-shares protocol can quantify the risk to their privacy. The agents can thus abstain if the risk to their privacy is undesirably high.

The Malicious-k-shares protocol includes mechanisms for preventing malicious agents from taking two particularly disruptive actions. This is done so in a decentralized manner without relying on trusted third parties. The two actions are as follows: 1) Taking advantage of private feedback, a malicious agent can be tempted to submit a value that is outside the valid interval for feedback. 2) A malicious agent can disrupt the protocol by making erroneous computations and reporting false results. The above challenges are addressed through constructions based on set-membership and plain-text equality non-interactive zero-knowledge proofs and an additive homomorphic cryptosystem.

In order to compute the reputation of a target agent, the querier first obtains the set of the source agents who can provide feedback about that target agent. The set is obtained from the target agent's *source managers*, who are assigned and located using a Distributed Hash Table (DHT). The querier then initiates the protocol by sending the set of agents to all the agents in the set.

Each source agent a selects k agents from the set based on his own subjective knowledge of their trustworthiness in the context of preserving privacy. The agent then sends each of them an additive share of his private feedback value, encrypted with the recipient agent's public key using an additive homomorphic cryptosystem. Each source agent is required to prove that it has generated correct shares. Correctness implies that the sum of all shares is a value that lies in the correct interval for feedback. The querying agent q serves as a relay for the shares.

Agent a also submits to q each of the shares encrypted with his own public key. Additionally, the agent submits a set-membership zero-knowledge proof, which shows that the sum of these shares belongs to the correct interval. The querying agent verifies the correctness by using the additive homomorphic property. This is done by adding the set of shares encrypted with agent a's key and then by verifying the proof. Moreover, agent a sends a plaintext-equality zero-knowledge proof for each share, which demonstrates that the same plaintext is contained in a share encrypted with the recipient's public key and a share encrypted with the sender's public key. The querying agent can verify the equality of all pairs of shares and be assured that agent a indeed sent correct shares.

In addition to proving that a source agent a sent correct shares, the agent must also prove that it has correctly computed the sum of the shares. The agent a computes the sum and sends it to q encrypted with q's public key. Agent a also sends a plaintext-equality zero-knowledge proof. The querying agent q computes the encrypted sum of the shares from the encrypted shares that it had relayed to an agent a. The proof sent by agent a shows that the encrypted sum that is computed independently by q and the encrypted sum

sent by agent a contain the same plaintext. Verifying this proof satisfies q that agent a correctly computed the sum of the shares.

When the querier has received the sums of shares and the proofs from all agents in the set, it can compute the reputation of the target agent as the total sum, as well as verify the correctness of each of the shares and the computed sums.

9.1.2 Dimitriou and Michalas 2014

Dimitriou and Michalas [57] present StR^M (Splitting the Random values in the Malicious model), an SMPC-based protocol for privacy preserving trust computation in decentralized environments in the presence of malicious adversaries. A key characteristic of the StR^M protocol is that their is no single node in a protocol instance that serves as a relay. Moreover, the authors formally prove that the protocol is resistant to collusion against as many as $n-1$ corrupted nodes.

The StR^M protocol is based on a simpler protocol called StR presented in the same paper, which is secure under the semi-honest adversarial model. In the StR protocol, the querier creates the set of the n voters, orders them in a circle, and sends each of them the identity of its successor in the circle. Each node chooses a random blinding factor r, then splits it into n shares, and sends a share to each node in the circle. After receiving the $n-1$ shares from the other nodes, a participant voter node computes its blinded vote. It subtracts n shares from the value of its vote. These shares comprise of the $n-1$ shares received and its own n^{th} share. It also adds its blinding factor r. The blinded vote is then sent to the querier who can compute the trust as the mean of the votes by summing up the received blinded votes.

In the StR^M protocol, each node encrypts the shares with its own public key in a homomorphic cryptosystem as well as the receiver's public key. The node also generates a zero-knowledge proof of plaintext equality to demonstrate that the two encryptions contain the same share. Moreover, the node creates a zero-knowledge proof of set membership for its vote. This is to prove that the encrypted vote lies in the expected interval. The zero-knowledge proofs, the encrypted shares, and the encrypted vote are broadcast to the network. Each node receives the $n-1$ shares encrypted with its public key. Each node decrypts and subtracts the received shares from its vote. All nodes can verify the computation performed by every other node in the network due to the messages being broadcast and due to the homomorphic property of the cryptosystem being used.

After the computation, each node encrypts its result with the querier's public key. The node also generates a zero-knowledge proof of equality. This proof is used to demonstrate that the node's blinded vote (which was encrypted by its own public key, and could be verified by other nodes) and the encrypted vote being sent to the querier, contain the same value. When the querier receives all the values encrypted with its public key, it can decrypt them, and compute the sum and then the trust as the mean. Since all the blinded votes and the

zero-knowledge proofs were broadcast, the querier can also verify that the nodes generated the blinded votes correctly.

9.2 Token-based Systems

9.2.1 Androulaki et al. 2008

The contributions of the work by Androulaki et al. [37] include formal definition and realization of a secure identity bound reputation system. The user is allowed to switch to new pseudonyms, however, the reputation of the user is maintained, while keeping their identity anonymous. The reputation system has a hybrid architecture, with both centralized and decentralized elements. The users interact with each other in a peer-to-peer manner. Whereas, reputation management operations are centralized with an entity called the bank. The building blocks employed include e-cash, anonymous credential systems, blind signatures, and anonymous communication networks.

Two different adversarial models are considered for different sets of objectives of the system. The malicious adversarial model is considered for the privacy related objectives, such as unlinkability and anonymity. Whereas, the semi-honest model is considered for the integrity related objectives of the system, which include reputation correctness, reputation unforgeability, etc.

The reputation score of a user is computed as the sum of the *repcoins* that he has received. However, the true reputation score is not revealed to the inquirer. Instead, the user demonstrates his membership to one of the reputation groups that are defined by the bank. A user U requests the bank for a credential called *cred* for the group G. If the user U has sufficient *repcoins* for membership in G, the bank issues to U the *cred* that U can use to demonstrate that it belongs to the group G. A brief overview of the protocol is presented below.

Reputation granting process. 1) A user U withdraws a wallet W of *repcoins* from the bank. Each *repcoin* is of the form (S, π), where S is a serial number and π is a proof of the validity of the coin. 2) User U, via a pseudonym P_U, awards a *repcoin* (S, π) to P_M, which is the pseudonym of a user M. 3) User M, via P_M, deposits the *repcoin* (S, π) to the bank. 4) Upon successful deposit of the *repcoin*, the bank issues a blind permission σ to P_M. 5) User M deposits σ to the bank, who increases M's reputation score by one.

Reputation demonstration process. 1) User M requests a credential for the group G. 2) If M has enough reputation points for G, the bank issues an anonymous credential *cred* to M. 3) M proves his membership in the group G to P_U by using *cred* via his pseudonym P_M.

The system has a strong defense against Sybil attacks in the semi-honest model. A user can create multiple pseudonyms, however, his number of *repcoins* and reputation points remain unaffected. A user cannot create multiple

identities because the account is associated with his unique public key. However, if the malicious adversarial model is considered for the bank, the system does not have a defense against Sybil attacks since the bank can allow multiple accounts for the same user.

9.3 Proxy-based Systems

9.3.1 Ries et al. 2011

In this system by Ries et al. [65], a user U, who wishes to interact with a service provider SP, chooses a set of other users, $S = \{S_1, ..., S_n\}$, who will provide their feedback about SP. A TTP Z then provides each S_i with a public key from a homomorphic cryptosystem, with Z the only one knowing the corresponding secret key. When S_i answers U's query for feedback, she also sends to Z her encrypted feedback score, obfuscated by adding a random number shared with a partner, and the trust score of U in S_i.

Upon receiving all the feedback scores from the different users in S, U aggregates the reputation score using the properties of the homomorphic cryptosystem. U then sends the encrypted score to Z for verification and decryption. Z computes the same aggregation and verifies its equality with the one computed by U. If the two match, Z decrypts the sum and sends it to U, who can then decide if she wishes to engage in a transaction with SP. S_i's feedback confidentiality is thus guaranteed by Z.

Once the transaction is complete, the system proposes a way for U to update the value of her trust in S_i. User U compares her own feedback about SP to the encrypted feedback of S_i, and updates her trust in S_i as a consequence.

9.4 Signature-based Systems

9.4.1 Lajoie-Mazenc et al. 2015

Lajoie-Mazenc et al. [70] propose a decentralized privacy preserving reputation system for the client / Service Provider (SP) model. The system offers non-monotonic reputation. This is in contrast to earlier systems in the signature-based systems category, such as the one by Bethencourt [67], which only provide monotonic reputation. The protocol aims to be secure against malicious adversaries. The protocol comprises of the following three main steps.

Commitment. This step takes place before the transaction between the client and the service provider. They start by choosing n nodes that will constitute the share carriers (SC). The client and the SP then follow a protocol that enables them to compute an invariant of the tuple $(client, SP)$, denoted by *inv*. The SP first computes a *pre_inv* and sends it to the client along with a zero-knowledge proof of correctness. The client and the SP also exchange signatures on their certificates (issued by a Certification Authority) and zero-knowledge proofs of registration. The client can compute *masked_inv* upon receiving *pre_inv*. Using *pre_inv* and *masked_inv*, the SP

can determine *inv*. However, at this step, the SP can also deviate from the protocol if the SP knows that the client has assigned him unfavorable feedback before. To prevent this, the SP shouldn't learn the value of *masked_inv* before the transaction takes place. Therefore, the client commits to *masked_inv* by splitting this secret among the n share carriers, in such a way that at least $t < n$ of them must collude in order to retrieve it. The authors suggest $t = \lceil n/3 \rceil$ as an optimal value.

Feedback submission. After the transaction has been completed, the client chooses a rating ρ for the service provider. The client then commits to the rating (for example, by hashing it with a random nonce) and then transmits *masked_inv* to the SP. The SP can now compute *inv*. The SP sends it to the client, along with its identity, which can be divulged at this stage. The SP also reveals the nonce that was used to compute *pre_inv* and opens the commitment on its pseudonym. If the client does not send ρ and *masked_inv* after a certain delay, the SP can contact the share carriers. They in turn can convoke the client. If the client still does not respond, they can retrieve *masked_inv* from the shares of the secret. After receiving the *masked_inv*, the SP is able to compute *inv* and issue the feedback report. If the service provider doesn't respond to the client, the client follows a similar protocol in order to retrieve the identifier of the SP and compute *inv* in order to issue the feedback report.

Reputation computation. At the end of each round, the share carriers send the reports they collected since the last round to the accredited signers, which are entities that constitute a distributed trusted authority. The rounds correspond to some fixed time duration or some fixed number of transactions. The accredited signers verify the proofs and the signatures. If they are valid, they keep the rating ρ, *inv*, the identifier of the SP and the identifier of the transaction. The accredited signers then update the reputation of the SP, sign it and send it to the SP.

Rater anonymity is achieved using pseudonyms that are commitments to the identities. The commitment can be verified by the accredited signers but not by the service providers. Linkability is achieved through the computation of *inv*, an invariant unique to the tuple $(client, SP)$, but which does not reveal the client's identity. Robustness, in case of protocol abortion, is achieved through the escrow mechanism for *masked_inv*. After a transaction is performed, the client is assured that it will be able to issue a rating, and the service provider is assured that it will receive a rating. Due to the signatures and zero-knowledge proofs, the integrity of the transmitted information is ensured. Registration of new users is restricted by the certification authority. The authors suggest a fee for new users in order to prevent Sybil attacks. The protocol is secure against whitewashing if the certification authority delivers only one certificate for each identity. In this case, dishonest service providers would not be able to leave the network and re-enter with a different identity.

9.5 Transitory Pseudonym-based Systems

9.5.1 Christin et al. 2013

IncogniSense, a reputation system proposed by Christin et al. [74], is based on the the RuP (Reputation using Pseudonyms) system by Miranda and Rodrigues [71].

We first give a brief description of the RuP system. In this system, a user is uniquely identified by a certified pseudonym that is valid only for a predefined time slot. A TTP, called pseudonym certification authority (PCA), is assumed in the architecture, which issues certified pseudonyms. The TTP is not trusted to learn the link between the real identity of the user and the pseudonym or their multiple pseudonyms. The system therefore offers a scheme based on blind signatures for a user to anonymously transfer reputation between pseudonyms.

In the RuP system, a user U receives a pseudonym valid for a time slot T certified by the PCA. The PCA uses a probabilistic blinding method to assess the validity of this pseudonym. U sends n different pseudonyms and then the PCA randomly checks $n-1$ pseudonyms. If they are all correct, it declares the remaining pseudonym correct and blindly signs it, thus certifying its validity for the next time slot. The blind signature of the pseudonym by the certification authority enforces the unlinkability between U and its ratings.

IncogniSense is similar to RuP in the sense that users in IncogniSense also use pseudonyms verified by a TTP, that are allocated for a certain time slot T. However, Christin et al. propose an updated approach of transferring reputation from one pseudonym to another, which further emphasizes unlinkability. They introduce periodic signature keys and different transfer keys for reputation tokens. This approach decouples the time interval of validity of the pseudonym and the value of the reputation to transfer. The TTP signs a blinded reputation token RT from a pseudonym used during the previous time slot. RT contains a certain amount of reputation that U can then transfer to its pseudonym for the new time slot. Giving RT back to the TTP links it to his new pseudonym, without linking it to the previous one. IncogniSense proposes three ways to do the reputation transfer: 1) flooring the reputation value; 2) dividing the reputation value in several tokens and randomly discarding one; or 3) dividing the reputation value in several tokens and randomly linking some of them to the new pseudonym.

9.6 Other Systems

9.6.1 Kinateder and Pearson 2003

The decentralized reputation system proposed by Kinateder and Pearson [19] requires a Trusted Platform Module (TPM) chip at each node. The TPM enables a node to demonstrate that it is a valid node and a legitimate member of the reputation system without disclosing its true identity. This permits the node to provide feedback anonymously.

A node in the system can take up one of following three roles at any given time: *recommender*, *requester*, and *accumulator*. A recommender node has interacted directly with other nodes and has feedback about them. An accumulator node stores feedback about other nodes that it has received.

An attacker is unable to provide false feedback on an honest user's behalf since each feedback is digitally signed by the recommender. A requester node can also verify through the recommender's TPM that it has not been compromised by the adversary. An attacker is unable to access an honest user's private database and modify data such as feedback, reputation, etc. This is achieved due to the protected data storage functionality of the TPM. Therefore, a requester can be certain that the given feedback is not false.

An attacker does not learn the true identity of a feedback provider since only pseudonyms are used. The pseudonym is protected by the TPM and the Certification Authority (CA) of the user. Moreover, the use of MIX cascades is suggested to prevent the attacker from correlating the pseudonym with the IP address of the user. In case of legal justification, the CA of a user can reveal his true identity.

10 Discussion

The fine-grained analysis and comparison of privacy preserving reputation systems carried out in this survey, according to the proposed analysis frameworks, reveal a number of insights into this field of research.

Our first observation concerns the utilization of blockchain by privacy preserving reputation systems. We note from the literature studied that the advent of the blockchain technology has provided a fresh impetus to research on privacy preserving reputation systems. A majority of the systems published since 2016 that are listed in this survey utilize blockchain as one of the building blocks. We looked at 12 privacy preserving reputation systems since 2016 that are blockchain-based. In contrast, we discovered only 5 systems that do not utilize blockchain. The reasons for the adoption of blockchain are evident. For example, in the case of Schaub et al.'s [4] system, using blockchain enables the system to provide the property of trustlessness, which was not offered by any prior systems. Another example is the system by Schiedermeier et al. [45], which is able to guarantee transparency and immutability by employing a blockchain. These properties are mostly absent in pre-blockchain systems.

Despite the successful application of blockchain, we do note that the development of non-blockchain-based privacy preserving reputation systems still holds importance. We can cite a couple of reasons. Firstly, blockchain can be an expensive building block to rely on in terms of the resources consumed. The computing cycles and the network bandwidth spent, and more worryingly the carbon footprint of popular blockchain-based systems such as Bitcoin, remain a significant concern [97]. Secondly, certain applications do not benefit as much as others from the decentralization and the trustlessness that blockchain offers. One such application is mobile participatory or crowd sensing. We note that

two (Ma et al. [6] and Mousa et al. [7]) of the five non-blockchain-based privacy preserving reputation systems since 2016 listed are for this application area. Moreover, they both employ a centralized architecture due to the nature of the application, which collects reports from mobile users and centralizes the data for subsequent analysis and exploitation. We acknowledge that two (Jo and Choi [5] and Zhao et al. [8]) of the blockchain-based systems included in the survey also target the participatory sensing application area. These two systems benefit from the smart contract functionality of blockchain technology to transparently manage the reputation of participants. However, we can observe that both systems employ centralized TTPs in their architecture and thus do not take full advantage of the decentralization and trustlessness properties of blockchain.

Our above observations lead us to another notable and perhaps undesirable trend. Fully decentralized systems have existed since before blockchain. A key advantage that blockchain is able to offer in addition to decentralization is trustlessness. However, we remark that among all the blockchain-based systems studied, only one system (Schaub et al. [4]) benefits from this novel trust model to propose a fully trustless privacy preserving reputation system. Other blockchain-based systems do benefit in part from the trustlessness of blockchain but end up proposing hybrid trust models that include arbitrary k trusted users, chosen k trusted users, or even TTPs. We believe that one of the future directions in this area of research is to exploit the blockchain technology to its full potential and build truly trustless systems.

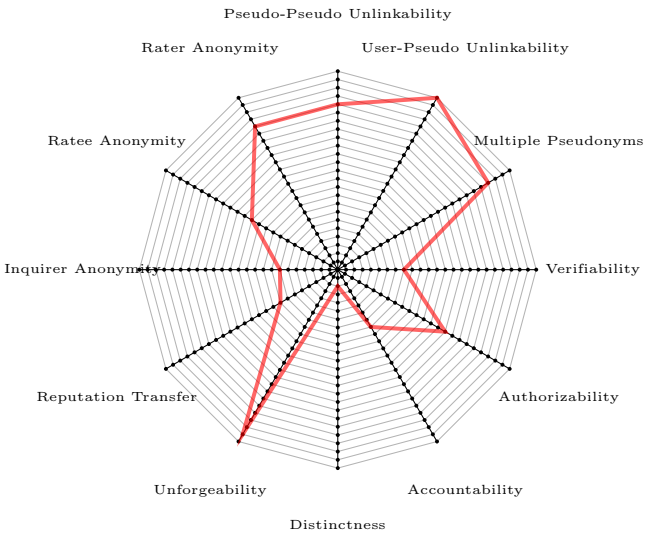

Figure 4: The number of user anonymity oriented systems (out of a total of 24) that fully satisfy the given security objectives. Note: there are no systems that partially satisfy the objectives.

Next, we look at the success of the surveyed systems in guaranteeing the security of users. As discussed earlier in Section 4, the objectives of security include privacy and integrity. We first address user anonymity oriented systems. Figure 4 illustrates the 12 identified individual security objectives of user anonymity oriented systems and the portion of the 24 systems included in the literature that fulfill each of these objectives.

In terms of privacy properties, we can observe that all of the systems guarantee user-pseudo unlinkability (24 systems). This is to be expected since this is a vital goal of user anonymity oriented systems. Moreover, a high majority of the systems enable multiple pseudonyms (21 systems), pseudo-pseudo unlinkability (20 systems), and rater anonymity (20 systems). This is another positive sign indicating success of the systems toward providing strong privacy to the users. On the other hand, we can note that much fewer systems aim for guaranteeing ratee anonymity (12 systems) and inquirer anonymity (7 systems). These properties have been ignored by most of the systems even though these are important properties for the privacy of roles other than the raters. We can identify inclusion of these objectives in future privacy preserving reputation systems as another direction of research. Reputation transfer and aggregation is another property that is offered by some systems but not provided by most others. We believe that this is an important property for long term sustainable privacy in the system and should thus be given priority as well.

Moving to the properties of integrity, we are pleased to observe that almost all systems (23) enforce unforgeability, an essential property for the correct functioning of the user anonymity oriented systems. Unfortunately, the assessment is not as bright for the rest of the integrity properties. There are 8 or less systems implementing the properties of either distinctness, accountability, or verifiability. The property of authorizability is offered by only 15 of the systems that we have studied. This is a worrisome figure since we believe that authorizability must be a critical feature of any anonymity oriented privacy preserving reputation system. Absence of this property can allow an adversary to take unfair advantage of anonymity and mount attacks such as ballot stuffing and slandering. The encouraging news is that if we consider only the subset of systems since 2016, we can observe that 7 out of the 10 systems offer authorizability. Thus, the trend is moving favorably in the direction toward including authorizability.

We now discuss the feedback confidentiality oriented systems and their success in enforcing the listed security objectives. Figure 5 shows the 3 privacy objectives and the 4 integrity objectives of feedback confidentiality oriented systems and the fraction of the 16 systems that satisfy those objectives.

Considering the privacy objectives, we observe that all systems ensure that feedback confidentiality is maintained even if the adversary has access to intermediate information revealed during the execution of the protocols. This is the primary privacy objective of feedback confidentiality systems therefore this property is the minimum expectation from any system. In contrast, we observe that only half of the systems can guarantee to some degree that an adversary will be unable to derive the feedback values from publicly available information, which includes the computed reputation scores. However, this issue is generally

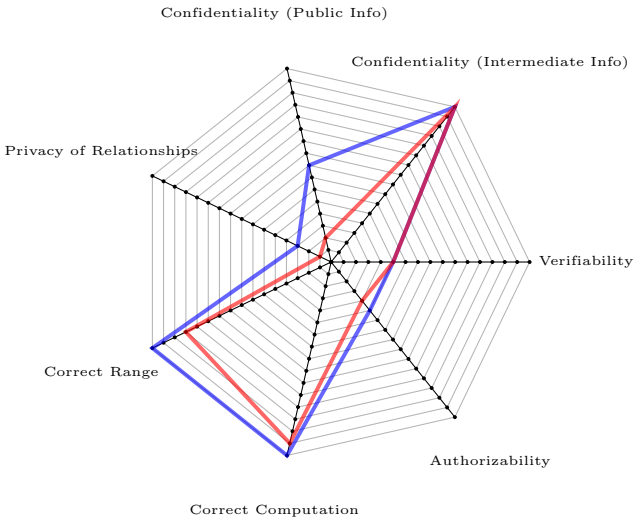

Figure 5: The number of feedback confidentiality oriented systems (out of a total of 16) that satisfy the given security objectives. Blue line: fully or partially satisfied. Red line: fully satisfied.

of concern when the number of participants is low. Therefore, even if systems do not ensure this property, they should be able to take measures to either warn users when their privacy is at risk or prevent execution of protocol instances with few participants. The third privacy related property, that is privacy of relationships, concerns a subset of the systems that rely on relationships between users for privacy preservation. We observe that only 3 systems are able to satisfy this property to some extent.

Looking at the integrity objectives, we appreciate that almost all systems fully enforce correct computation as well as guarantee that submitted feedback will respect the correct range. This is a reassuring trend since these two properties imply that systems are able to produce correct reputation scores despite the confidentiality of the feedback values. Regrettably, similar to anonymity oriented systems, the feedback confidentiality oriented systems also largely ignore the properties of authorizability (5 systems) and verifiability (5 systems). Even if we consider recent feedback confidentiality oriented systems since 2016, we remark that only 3 out of the 7 systems fully satisfy this property. As we argued earlier, authorizability is an important property, therefore future work on feedback confidentiality oriented privacy preserving reputation systems should focus on its inclusion.

Lastly, we discuss the systems in terms of their measures against challenges other than privacy as analyzed in Table 6. Figure 6 illustrates the number of the 40 systems that propose defenses to the 7 listed challenges. We observe that the number of systems implementing measures against these challenges is fairly

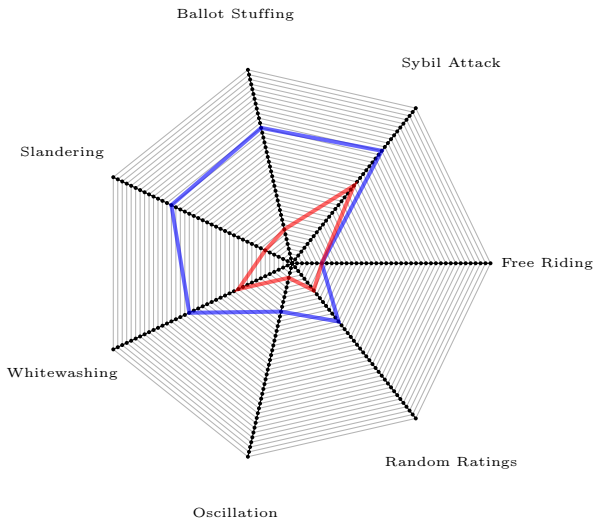

Figure 6: The number of systems (out of a total of 40) that propose measures for the listed challenges. Blue line: strong or partial measures. Red line: strong measures.

low all across the board. This is particularly true for systems that propose strong measures. A majority of the systems shows some level of resistance to the Sybil attack (29 systems), ballot stuffing (28 systems), slandering (27 systems), and whitewashing (23 systems). Defenses against other attacks are mostly overlooked: oscillation (10 systems), random ratings (15 systems), and free riding (6 systems). The figures are starkly lower when we consider only systems that offer strong measures. For example, no more than 7 systems implement strong measures against any of the following attacks: ballot stuffing, slandering, oscillation, random ratings, and free riding.

Moreover, Table 6 reveals that only two systems (Mousa et al. [7] and Benthencourt et al. [67]), out of the 40 systems analyzed, provide somewhat comprehensive resistance to the challenges. However, both these systems employ TTPs in their architecture. None of the systems with a fully decentralized architecture or with less intrusive trust models offers resistance to the full range of challenges. Table 6 further shows that their is no noticeable improvement in recent systems toward offering better resistance to these challenges.

There is clearly more work that needs to be done in the field of privacy preserving reputation systems in terms of defenses against attacks other than breach of privacy. Privacy preserving reputation systems are fundamentally reputation systems and their overall success thus relies on countering their basic challenges as well. One possible reason for the non-inclusion of robust protection against these challenges is that anonymity and privacy add further obstacles to preventing attacks such as ballot stuffing, slandering, random ratings, free riding, and others. An adversary may exploit the anonymity and privacy offered

by a system to mount these attacks while simultaneously foregoing accountability. From these observations, an evident direction for future research in the area that can be pointed out is conceiving systems that provide comprehensive protection against the broad range of challenges faced by reputations systems.

11 Related Work

Bellini et al. [98] author a survey on blockchain-based distributed trust and reputation management systems. The survey defines uniform taxonomies for blockchain and for systems aimed at managing trust and reputation. Additionally, the survey employs the Formal Concept Analysis (FCA) technique to analyze the literature. The authors provide recommendations for the utilization of blockchain in the context of trust and reputation management. In contrast to the work by Bellini et al., our survey focuses specifically on privacy preserving reputation systems based on blockchain as well as other cryptographic building blocks.

The survey by Michalas et al. [99] is one of the closest to our work as it addresses privacy in decentralized additive reputation systems. Michalas et al. identify and analyze the vulnerabilities of privacy preserving reputation systems in the semi-honest and the malicious adversarial models. The survey covers three sets of decentralized additive reputation systems (from Pavlov et al. [18], Hasan et al. [33], and Dolev et al. [60]). In comparison, our survey aims to provide a broader perspective of the field of privacy preserving reputation systems.

Schiffner et al. [25] present an analysis of some privacy preserving reputation systems in the literature as part of their paper that describes a novel system that preserves privacy as well as maintains liveliness. Their analysis compares their own system with two others, namely those by Androulaki et al. [37] and Voss [79]. On the other hand, our work establishes an analysis framework that covers a wide array of privacy preserving reputation systems. Moreover, we analyze and compare several privacy preserving reputation systems belonging to the two different categories of user anonymity and feedback confidentiality.

A survey by Chang et al. [100] studies approaches for promoting honest feedback in reputation systems, which include protecting the privacy of the feedback providers as well as providing them incentives. The work is focused in large part on the latter category, that is, providing incentives. However, four privacy oriented systems (Pavlov et al. [18], Hasan et al. [33], Gudes et al. [34], and Kinateder and Pearson [19]) are also analyzed and compared.

Hasan et al. [101] author a book chapter on privacy preserving reputation management in the context of social networks that describes in detail some privacy preserving reputation systems. However, this work does not establish analysis frameworks as extensive as the current survey and discusses a much smaller subset of the systems in the literature. Moreover, blockchain-based systems are not covered.

Tran et al.'s position paper [102] on the challenges and opportunities of privacy preserving reputation management in fully decentralized systems includes

a summary of the systems in this category.

Mousa et al.'s survey [103] describes trust management and reputation systems for mobile participatory sensing applications. Several reputation systems specific to this application area including some that respect privacy are discussed. Moreover, the survey identifies participant privacy preservation as one of the future research directions for reputation systems that serve mobile participatory sensing applications.

Koutrouli and Tsalgatidou [104] present a survey describing the taxonomy of attacks and defense mechanisms in peer-to-peer reputation systems. The conflict between privacy and trust is discussed as part of this survey. The authors describe that estimating reputation and trust requires users to sacrifice their privacy in terms of information regarding their transactions and their opinions. The survey offers guidelines for building resilient peer-to-peer reputation systems including some recommendations for balancing the tension between privacy and trust.

Hoffman et al. [24] present a survey of attack and defense techniques for reputation systems. The survey describes a number of challenges that reputation systems face and techniques that can resolve those challenges. However, their survey does not address the issue of privacy in reputation systems. A survey by Mármol and Pérez [105] also analyzes threat scenarios for reputation systems. Their survey does not cover privacy preserving reputation systems either.

12 Conclusion

In this survey, we presented an in-depth analysis of a broad range of privacy preserving reputation systems. To the best of our knowledge, this is the first survey to have covered privacy preserving reputation systems in an extensive manner.

The survey identified the various dimensions of privacy preserving reputation systems. An analysis framework that allows for the decomposition and comparison of privacy preserving reputation systems in a normalized manner is proposed. As a first step, we presented an analysis framework that covers the fundamental elements that are common to all reputation systems and not just those that preserve privacy. We identified the following elements for this initial framework: the architecture of the system, the properties of the feedback, the properties of the reputation, the feedback aggregation model, the challenges addressed, and the reputation query costs. We then presented the analysis framework that specifically addresses privacy preserving reputation systems and decomposes them according to the following dimensions: the nature of the adversary, reputation binding, the trust model, the security objectives of the system, and the building blocks utilized.

Additionally, we identified the security requirements of privacy preserving reputation systems that cut across multiple types of such systems. It is observed that there are two main types of privacy preserving reputation systems: 1) systems that preserve the anonymity of the users, and 2) systems that don't

necessarily preserve the anonymity of the users but preserve the confidentiality of their feedback. We noted that the security-related requirements can be further subdivided into privacy requirements and integrity requirements.

The survey listed the building blocks of current privacy preserving reputation systems. We observed that the various strategies and associated building blocks offer individual advantages and disadvantages. For example, the E-Cash strategy can be used to preserve the anonymity of users, however, it requires a centralized entity which makes it unsuitable for decentralized networks.

We presented a fine-grained analysis and comparison of 40 privacy preserving reputation systems using our analysis frameworks. We established several categories of systems according to their security mechanisms and classified the privacy preserving reputation systems according to these categories. Our detailed comparison of privacy preserving reputation systems in a normalized manner using our analysis frameworks reveals the differences between the systems in the literature as well as their chronological evolution.

The survey included detailed descriptions of a number of important and representative systems from each of the security mechanism-based categories that we defined. We discussed the details of their protocols and security approaches as well as highlighted their individual strengths and other salient features. We placed an emphasis on blockchain-based systems as they are a recent significant development in the area of privacy preserving reputation systems. We discussed five individual instances of these systems including the first trustless decentralized system by Schaub et al. [4] as well as more recent systems.

Our fine-grained analysis, comparison, and discussion led to the identification of a number of insights into this field of research. We observed that the advent of the blockchain technology has provided a fresh impetus to research on privacy preserving reputation systems. A majority of the systems published since 2016 that are listed in this survey utilize blockchain as one of the building blocks. However, we also noted that one of the future directions is to exploit the blockchain technology to its full potential and build truly trustless systems. We looked at the success of the surveyed systems in guaranteeing the security of users. It was observed that a high majority of both anonymity oriented and feedback confidentiality oriented systems are able to guarantee their respective essential privacy and integrity properties. However, there also exist many properties that have been mostly ignored. We identified authorizability as one of the important such properties that needs to be addressed by systems in the future. Lastly, analyzing the systems in terms of their measures against challenges other than privacy, we remarked that conceiving systems that provide comprehensive protection against a broad range of challenges is an evident direction for future research in the area.

Acknowledgments

The first author would like to thank Rémi Canillas for valuable suggestions, including those regarding the categorization of the systems according to their

security mechanisms. Moreover, the first author would like to thank Dr. Sonia Ben Mokhtar for helpful comments on an initial version of the manuscript.

References

[1] D. Liu, A. Alahmadi, J. Ni, X. Lin, X. Shen, Anonymous reputation system for iiot-enabled retail marketing atop pos blockchain, IEEE Transactions on Industrial Informatics 15 (6) (2019) 3527–3537.

[2] M. A. Azad, S. Bag, F. Hao, Privbox: Verifiable decentralized reputation system for online marketplaces, Future Generation Computer Systems 89 (2018) 44–57.

[3] S. Bag, M. A. Azad, F. Hao, A privacy-aware decentralized and personalized reputation system, Computers & Security 77 (2018) 514–530.

[4] A. Schaub, R. Bazin, O. Hasan, L. Brunie, A trustless privacy-preserving reputation system, in: IFIP International Conference on ICT Systems Security and Privacy Protection, Springer, 2016, pp. 398–411.

[5] H. J. Jo, W. Choi, Bprf: Blockchain-based privacy-preserving reputation framework for participatory sensing systems, Plos one 14 (12) (2019) e0225688.

[6] L. Ma, X. Liu, Q. Pei, Y. Xiang, Privacy-preserving reputation management for edge computing enhanced mobile crowdsensing, IEEE Transactions on Services Computing 12 (5) (2018) 786–799.

[7] H. Mousa, S. B. Mokhtar, O. Hasan, L. Brunie, O. Younes, M. Hadhoud, Privasense: Privacy-preserving and reputation-aware mobile participatory sensing, in: Proceedings of the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, 2017, pp. 38–47.

[8] K. Zhao, S. Tang, B. Zhao, Y. Wu, Dynamic and privacy-preserving reputation management for blockchain-based mobile crowdsensing, IEEE Access 7 (2019) 74694–74710.

[9] Z. Lu, W. Liu, Q. Wang, G. Qu, Z. Liu, A privacy-preserving trust model based on blockchain for vanets, IEEE Access 6 (2018) 45655–45664.

[10] L. Chen, Q. Li, K. M. Martin, S.-L. Ng, Private reputation retrieval in public–a privacy-aware announcement scheme for vanets, IET Information Security 11 (4) (2016) 204–210.

[11] M. A. Azad, S. Bag, F. Hao, A. Shalaginov, Decentralized self-enforcing trust management system for social internet of things, IEEE Internet of Things Journal 7 (4) (2020) 2690–2703.

[12] M. A. Azad, S. Bag, F. Hao, M2m-rep: Reputation of machines in the internet of things, in: Proceedings of the 12th international conference on availability, reliability and security, 2017, pp. 1–7.

[13] P. Resnick, R. Zeckhauser, Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system, The Economics of the Internet and E-Commerce. Michael R. Baye, editor. Volume 11 of Advances in Applied Microeconomics (2002) 127–157.

[14] N. Miller, P. Resnick, R. Zeckhauser, Eliciting informative feedback: The peer-prediction method, Management Science 51 (9) (2005) 1359–1373.

[15] T. Minkus, K. W. Ross, I know what you're buying: Privacy breaches on ebay, in: International Symposium on Privacy Enhancing Technologies Symposium, Springer, 2014, pp. 164–183.

[16] E. Ho, Why you should think twice before trusting airbnb reviews (May 2015).
URL \url{https://mashable.com/2015/05/18/airbnb-reviews/}

[17] M. Mulshine, After a disappointing airbnb stay, i realized there's a major flaw in the review system (June 2015).
URL \url{https://www.businessinsider.com/why-airbnb-reviews-are-a-problem-for-the-site-2015-6}

[18] E. Pavlov, J. S. Rosenschein, Z. Topol, Supporting privacy in decentralized additive reputation systems, in: Proceedings of the Second International Conference on Trust Management (iTrust 2004), Oxford, UK, 2004.

[19] M. Kinateder, S. Pearson, A privacy-enhanced peer-to-peer reputation system, in: Proceedings of the 4th International Conference on Electronic Commerce and Web Technologies, 2003.

[20] R. Dingledine, N. Mathewson, P. Syverson, Reputation in p2p anonymity systems, in: Proceedings of the Workshop on Economics of Peer-to-Peer Systems, 2003.

[21] Y. Dou, H. C. Chan, M. H. Au, A distributed trust evaluation protocol with privacy protection for intercloud, IEEE Transactions on Parallel and Distributed Systems 30 (6) (2018) 1208–1221.

[22] D. D. S. Braga, M. Niemann, B. Hellingrath, F. B. D. L. Neto, Survey on computational trust and reputation models, ACM Computing Surveys (CSUR) 51 (5) (2018) 1–40.

[23] F. Hendrikx, K. Bubendorfer, R. Chard, Reputation systems: A survey and taxonomy, Journal of Parallel and Distributed Computing 75 (2015) 184–197.

[24] K. Hoffman, D. Zage, C. Nita-Rotaru, A survey of attack and defense techniques for reputation systems, ACM Computing Surveys 41 (4) (December 2009).

[25] S. Schiffner, S. Clauß, S. Steinbrecher, Privacy and liveliness for reputation systems, in: Proceedings of the Sixth European Workshop on Public Key Infrastructures, Services and Applications (EuroPKI'09), 2009, pp. 209 – 224.

[26] A. Josang, R. Ismail, C. Boyd, A survey of trust and reputation systems for online service provision, Decision Support Systems 43 (2) (2007) 618 – 644.

[27] R. Levien, Attack-Resistant Trust Metrics (Chapter 5). Computing with Social Trust., Springer London, 2008.

[28] S. D. Kamvar, M. T. Schlosser, H. GarciaMolina, The eigentrust algorithm for reputation management in p2p networks, in: Proceedings of the 12th International Conference on World Wide Web (WWW 2003), Budapest, Hungary, 2003.

[29] R. Zhou, K. Hwang, Powertrust: A robust and scalable reputation system for trusted peer-to-peer computing, IEEE Transactions on Parallel and Distributed Systems 18 (4) (2007) 460–473.

[30] A. Whitby, A. Josang, J. Indulska, Filtering out unfair ratings in bayesian reputation systems, in: Proceedings of the Workshop on Trust in Agent Societies, at the Autonomous Agents and Multi Agent Systems Conference (AAMAS2004), New York, 2004.

[31] J. R. Douceur, The sybil attack, in: Proceedings of the First International Workshop on Peer-to-Peer Systems, 2002.

[32] O. Goldreich, The Foundations of Cryptography - Volume 2, Cambridge University Press, 2004.

[33] O. Hasan, L. Brunie, E. Bertino, N. Shang, A decentralized privacy preserving reputation protocol for the malicious adversarial model, IEEE Transactions on Information Forensics and Security 8 (6) (2013) 949–962.

[34] E. Gudes, N. Gal-Oz, A. Grubshtein, Methods for computing trust and reputation while preserving privacy, in: Proceedings of the 23rd Annual IFIP WG 11.3 Working Conference on Data and Applications Security, 2009.

[35] M. Anwar, J. Greer, Reputation management in privacy-enhanced e-learning, in: Proceedings of the 3rd Annual Scientific Conference of the LORNET Research Network (I2LOR-06), Montreal, Canada, 2006.

[36] M. Anwar, J. Greer, Enabling reputation-based trust in privacy-enhanced learning systems, in: Proceedings of the 9th International Conference on Intelligent Tutoring Systems, Montreal, Canada, 2008.

[37] E. Androulaki, S. G. Choi, S. M. Bellovin, T. Malkin, Reputation systems for anonymous networks, in: Proceedings of the 8th Privacy Enhancing Technologies Symposium (PETS 2008), 2008.

[38] F. Kerschbaum, A verifiable, centralized, coercion-free reputation system, in: Proceedings of the 8th ACM workshop on Privacy in the electronic society (WPES'09), ACM, New York, NY, USA, 2009.

[39] C. Clifton, M. Kantarcioglu, J. Vaidya, X. Lin, M. Y. Zhu, Tools for privacy preserving distributed data mining, SIGKDD Explorations 4 (2) (2003) 28–34.

[40] P. Paillier, Public-key cryptosystems based on composite degree residuosity classes, in: Proceedings of the 17th International Conference on Theory and Application of Cryptographic Techniques, 1999.

[41] T. ElGamal, A public-key cryptosystem and a signature scheme based on discrete logarithms, IEEE Transactions on Information Theory IT-31 (4) (1985) 469–472.

[42] S. Goldwasser, S. Micali, C. Rackoff, The knowledge complexity of interactive proof systems, SIAM Journal on Computing 18 (1) (1989) 186–208.

[43] D. Chaum, Blind signatures for untraceable payments, in: Proc. Advances in Cryptology (CRYPTO '82), 1982.

[44] A. Michalas, N. Komninos, The lord of the sense: A privacy preserving reputation system for participatory sensing applications, in: 2014 IEEE Symposium on Computers and Communications (ISCC), IEEE, 2014, pp. 1–6.

[45] M. Schiedermeier, O. Hasan, L. Brunie, T. Mayer, H. Kosch, A transparent referendum protocol with immutable proceedings and verifiable outcome for trustless networks, in: International Conference on Complex Networks and Their Applications, Springer, 2019, pp. 647–658.

[46] D. Chaum, Blind signature systems, in: Advances in Cryptology (CRYPTO'83), 1983.

[47] M. Belenkiy, M. Chase, C. C. Erway, J. Jannotti, A. Kupcu, A. Lysyanskaya, E. Rachlin, Making p2p accountable without losing privacy, in: Proceedings of the 2007 ACM Workshop on Privacy in Electronic Society, 2007.

[48] J. Camenisch, A. Lysyanskaya, M. Meyerovich, Endorsed e-cash, in: Proceedings of the IEEE Symposium on Security and Privacy, 2007.

[49] C. Mitchell (Ed.), Trusted computing, Institution of Electrical Engineers, 2005.

[50] S. Pearson, B. Balacheff (Eds.), Trusted Computing Platforms: TCPA Technology in Context, Prentice Hall, 2003.

[51] R. Bazin, A. Schaub, O. Hasan, L. Brunie, Self-reported verifiable reputation with rater privacy, in: IFIP International Conference on Trust Management, Springer, 2017, pp. 180–195.

[52] E. Owiyo, Y. Wang, E. Asamoah, D. Kamenyi, I. Obiri, Decentralized privacy preserving reputation system, in: 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC), IEEE, 2018, pp. 665–672.

[53] N. Gal-Oz, N. Gilboa, E. Gudes, Schemes for privately computing trust and reputation, in: IFIP International Conference on Trust Management, Springer, 2010, pp. 1–16.

[54] N. Gal-Oz, E. Gudes, D. Hendler, A robust and knot-aware trust-based reputation model, in: Proceedings of the Joint iTrust and PST Conferences on Privacy, Trust Management and Security (IFIPTM 2008), 2008.

[55] R. Nithyanand, K. Raman, Fuzzy privacy preserving peer-to-peer reputation management, Cryptology ePrint Archive, Report 2009/442 (2009).

[56] O. Hasan, L. Brunie, E. Bertino, k-shares: A privacy preserving reputation protocol for decentralized environments, in: Proceedings of the 25th IFIP International Information Security Conference (SEC 2010), Brisbane, Australia, 2010, pp. 253–264.

[57] T. Dimitriou, A. Michalas, Multi-party trust computation in decentralized environments in the presence of malicious adversaries, Ad Hoc Networks 15 (2014) 53–66.

[58] T. Dimitriou, A. Michalas, Multi-party trust computation in decentralized environments, in: 2012 5th International Conference on New Technologies, Mobility and Security (NTMS), IEEE, 2012, pp. 1–5.

[59] S. Dolev, N. Gilboa, M. Kopeetsky, Efficient private multi-party computations of trust in the presence of curious and malicious users, Journal of Trust Management 1 (1) (2014) 8.

[60] S. Dolev, N. Gilboa, M. Kopeetsky, Computing multi-party trust privately: in o (n) time units sending one (possibly large) message at a time, in: Proceedings of the 2010 ACM Symposium on Applied Computing, 2010, pp. 1460–1465.

[61] M. R. Clark, K. Stewart, K. M. Hopkinson, Dynamic, privacy-preserving decentralized reputation systems, IEEE Transactions on Mobile Computing 16 (9) (2016) 2506–2517.

[62] S. Schiffner, S. Clauß, S. Steinbrecher, Privacy, liveliness and fairness for reputation, in: International Conference on Current Trends in Theory and Practice of Computer Science, Springer, 2011, pp. 506–519.

[63] K. Zhang, Z. Li, Y. Yang, A reputation system preserving the privacy of feedback providers and resisting sybil attacks, International Journal of Multimedia and Ubiquitous Engineering 9 (2) (2014) 141–152.

[64] N. Busom, R. Petrlic, F. Sebé, C. Sorge, M. Valls, A privacy-preserving reputation system with user rewards, Journal of Network and Computer Applications 80 (2017) 58–66.

[65] S. Ries, M. Fischlin, L. A. Martucci, M. Muuhlhauser, Learning whom to trust in a privacy-friendly way, in: 2011IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications, IEEE, 2011, pp. 214–225.

[66] R. Petrlic, S. Lutters, C. Sorge, Privacy-preserving reputation management, in: Proceedings of the 29th Annual ACM Symposium on Applied Computing, 2014, pp. 1712–1718.

[67] J. Bethencourt, E. Shi, D. Song, Signatures of reputation: Towards trust without identity, in: Proceedings of the Fourteenth International Conference on Financial Cryptography and Data Security (FC '10), 2010, pp. 400 – 407.

[68] L. Guo, Y. Fang, L. Wei, Fine-grained privacy-preserving reputation system for online social networks, in: 2013 IEEE/CIC International Conference on Communications in China (ICCC), IEEE, 2013, pp. 230–235.

[69] E. Anceaume, G. Guette, P. Lajoie-Mazenc, T. Sirvent, V. Viet Triem Tong, Extending signatures of reputation, Privacy and Identity Management for Emerging Services and Technologies, IFIP Advances in Information and Communication 421 (2014) 165–176.

[70] P. Lajoie-Mazenc, E. Anceaume, G. Guette, T. Sirvent, V. V. T. Tong, Efficient distributed privacy-preserving reputation mechanism handling non-monotonic ratings, hal.archives-ouvertes.fr (2015).

[71] H. Miranda, L. Rodrigues, A framework to provide anonymity in reputation systems, in: Proceedings of the Third Annual International Conference on Mobile and Ubiquitous Systems: Networking & Services, 2006.

[72] S. Steinbrecher, Design options for privacy-respecting reputation systems within centralised internet communities, in: Security and Privacy in Dynamic Environments, 2006.

[73] E. Anceaume, G. Guette, P. Lajoie-Mazenc, N. Prigent, V. V. T. Tong, A privacy preserving distributed reputation mechanism, in: 2013 IEEE International Conference on Communications (ICC), IEEE, 2013, pp. 1951–1956.

[74] D. Christin, C. Roßkopf, M. Hollick, L. A. Martucci, S. S. Kanhere, Incognisense: An anonymity-preserving reputation framework for participatory sensing applications, Pervasive and mobile Computing 9 (3) (2013) 353–371.

[75] Y. Bo, Z. Min, L. Guohuan, A reputation system with privacy and incentive, in: Proceedings of the Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD'07), 2007.

[76] R. Ismail, C. Boyd, A. Josang, S. Russell, Private reputation schemes for p2p systems, in: Proceedings of the Second International Workshop on Security in Information Systems (WOSIS'04), 2004.

[77] R. Ismail, C. Boyd, A. Josang, S. Russell, Strong privacy in reputation systems, in: Proceedings of the 4th International Workshop on Information Security Applications (WISA'03), 2004.

[78] D. Cvrcek, V. M. Jr., A. Patel, Evidence processing and privacy issues in evidence-based reputation systems, Computer Standards & Interfaces 27 (2005) 533 – 545.

[79] M. Voss, A. Heinemann, M. Muhlhauser, A privacy preserving reputation system for mobile information dissemination networks, in: Proceedings of the First International Conference on Security and Privacy for Emerging Areas in Communications Networks (SECURECOMM), 2005.

[80] M. Kinateder, R. Terdic, K. Rothermel, Strong pseudonymous communication for peer-to-peer reputation systems, in: Proceedings of the 2005 ACM symposium on Applied computing, 2005.

[81] L. Hao, S. Lu, J. Tang, A. Zhang, A low cost and reliable anonymity scheme in p2p reputation systems with trusted third parties, in: IEEE GLOBECOM 2008-2008 IEEE Global Telecommunications Conference, IEEE, 2008, pp. 1–5.

[82] J. Nin, B. Carminati, E. Ferrari, V. Torra, Computing reputation for collaborative private networks, in: Proceedings of the 33rd Annual IEEE International Computer Software and Applications Conference, 2009.

[83] F. G. Mármol, J. Girao, G. M. Pérez, Trims, a privacy-aware trust and reputation model for identity management systems, Computer Networks 54 (16) (2010) 2899–2912.

[84] Z. Zhang, J. Liu, Y. Kadobayashi, Stars: a simple and efficient scheme for providing transparent traceability and anonymity to reputation systems, in: Data Privacy Management and Autonomous Spontaneous Security, Springer, 2010, pp. 170–187.

[85] B. Kellermann, S. Pötzsch, S. Steinbrecher, Privacy-respecting reputation for wiki users, in: IFIP International Conference on Trust Management, Springer, 2011, pp. 223–239.

[86] M. T. Goodrich, F. Kerschbaum, Privacy-enhanced reputation-feedback methods to reduce feedback extortion in online auctions, in: Proceedings of the first ACM conference on Data and application security and privacy, 2011, pp. 273–282.

[87] O. Hasan, L. Brunie, E. Damiani, A privacy preserving reputation protocol for web service provider selection, in: 2011 IEEE 35th Annual Computer Software and Applications Conference Workshops, IEEE, 2011, pp. 404–409.

[88] O. Hasan, L. Brunie, E. Bertino, Preserving privacy of feedback providers in decentralized reputation systems, Computers & Security 31 (7) (2012) 816–826.

[89] K. L. Huang, S. S. Kanhere, W. Hu, A privacy-preserving reputation system for participatory sensing, in: 37th Annual IEEE Conference on Local Computer Networks, IEEE, 2012, pp. 10–18.

[90] M. H. Au, A. Kapadia, Perm: Practical reputation-based blacklisting without ttps, in: Proceedings of the 2012 ACM conference on Computer and communications security, 2012, pp. 929–940.

[91] M. H. Au, W. Susilo, Y. Mu, Constant-size dynamic k-taa, in: International conference on security and cryptography for networks, Springer, 2006, pp. 111–125.

[92] X. Wang, W. Cheng, P. Mohapatra, T. Abdelzaher, Enabling reputation and trust in privacy-preserving mobile sensing, IEEE Transactions on Mobile Computing 13 (12) (2013) 2777–2790.

[93] S. Clauß, S. Schiffner, F. Kerschbaum, k-anonymous reputation, in: ASIA CCS 2013, ACM, 2013.

[94] A. Aldini, A. Bogliolo, C. B. Lafuente, J.-M. Seigneur, On the tradeoff among trust, privacy, and cost in incentive-based networks, in: IFIP International Conference on Trust Management, Springer, 2014, pp. 205–212.

[95] S. Brangewitz, A. Jungmann, R. Petrlic, M. C. Platenius, Towards a flexible and privacy-preserving reputation system for markets of composed services, in: Proceedings of the 6th International Conferences on Advanced Service Computing (SERVICE COMPUTATION), 2014.

[96] M. Zhang, Y. Xia, O. Yuan, K. Morozov, Privacy-friendly weighted-reputation aggregation protocols against malicious adversaries in cloud services, International Journal of Communication Systems 29 (12) (2016) 1863–1872.

[97] C. Stoll, L. Klaaßen, U. Gallersdörfer, The carbon footprint of bitcoin, Joule 3 (7) (2019) 1647–1661.

[98] E. Bellini, Y. Iraqi, E. Damiani, Blockchain-based distributed trust and reputation management systems: a survey, IEEE Access 8 (2020) 21127–21151.

[99] A. Michalas, T. Dimitriou, T. Giannetsos, N. Komninos, N. R. Prasad, Vulnerabilities of decentralized additive reputation systems regarding the privacy of individual votes, Wireless Personal Communications 66 (3) (2012) 559–575.

[100] J. Chang, L. Xiao, W. Xu, A survey of approaches for promoting honest recommendations in reputation systems, in: CCF National Conference on Computer Engineering and Technology, Springer, 2018, pp. 179–191.

[101] O. Hasan, L. Brunie, Privacy preserving reputation management in social networks, in: Security and Privacy Preserving in Social Networks, Springer, 2013, pp. 245–280.

[102] N. H. Tran, L. Bahri, B. Q. Nguyen, Privacy-preserving reputation management in fully decentralized systems: Challenges and opportunities, in: The Joint International Symposium on Artificial Intelligence and Natural Language Processing, Springer, 2017, pp. 207–215.

[103] H. Mousa, S. B. Mokhtar, O. Hasan, O. Younes, M. Hadhoud, L. Brunie, Trust management and reputation systems in mobile participatory sensing applications: A survey, Computer Networks 90 (2015) 49–73.

[104] E. Koutrouli, A. Tsalgatidou, Taxonomy of attacks and defense mechanisms in p2p reputation systems?lessons for reputation system designers, Computer Science Review 6 (2-3) (2012) 47–70.

[105] F. G. Mármol, G. M. Pérez, Security threats scenarios in trust and reputation models for distributed systems, computers & security 28 (7) (2009) 545–556.