IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 26, 2019, accepted May 21, 2019, date of publication May 27, 2019, date of current version June 11, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2919322

A Survey on Measuring Anonymity in Anonymous Communication Systems

TIANBO LU[1,2], ZEYU DU[1,2], AND Z. JANE WANG[3]

[1]School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
[2]Key Laboratory of Trustworthy Distributed Computing and Service, Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, China
[3]Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada

Corresponding author: Tianbo Lu (lutb@bupt.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61170273, in part by the China Scholarship Council under Grant [2013]3050, and in part by the Beijing Natural Science Foundation under Grant 4194086.

ABSTRACT The popularity of Internet applications has made communication privacy an increasingly important security requirement. As an important aspect of privacy, anonymity ensures that a subject may use resources or services without disclosing user identity or corresponding relationships. Since the seminal work by Chaum for anonymous communication, many different anonymous communication systems, and anonymous protocols have been developed and investigated extensively. In recent years, anonymous communication systems have evolved from academic tools used by specialists to mass-market software used by millions of ordinary people. How to evaluate and quantify the anonymity that different anonymous communication systems can offer has been a new focus. Though some efforts have been made on anonymity metrics and measuring techniques, systematic research on measuring anonymity of anonymous communication systems is still needed. In this paper, we give a comprehensive overview of state-of-the-art research on measuring anonymity. We first summarize the anonymous mechanisms applied by different anonymous systems. We then introduce the formalization of the notion of anonymity for measuring the anonymity degree. Further, metrics based on various theories and approaches are reviewed. We particularly elaborate on the metrics based on information theory as a separate section because of its popularity and multiple branches. The metrics based on information entropy is probably of the greatest interest.

INDEX TERMS Anonymous systems, formalization, metrics, information theory.

I. INTRODUCTION

In recent years, the Internet has significantly changed people's way of living and communicating with each other. Before the early 90s, Internet was mainly a scientific research network with little consideration on security and privacy. Then thanks to the appearance of World Wide Web, millions of users can surf the Internet. The Internet becomes an information sharing platform for hundreds of millions of users. While people are enjoying the great convenience brought by the Internet, they have to face the accompanying privacy threat.

A network that considers personal privacy should allow its users to decide which information can be made public or released. Such information may include the reading habit, shopping habit, family address, social relations, browsing interests, identities, username and passwords, and so on. At present, the content of the information could be protected from being read with the help of relatively mature data encryption algorithms, but the message header information, such as source address, destination address, and message length, which is needed by the TCP/IP protocol, is difficult to be hidden by encryption. Attackers could obtain valuable personal information of the two communication parties through traffic analysis or other attack methods. For example, attackers could figure out whom an email is sent to, what service is used by a certain IP address, between which users there are regular email communications. The leakage of such information can bring great threat to users' privacy. Therefore, it is not enough to protect users' privacy just through data encryption. Anonymity has become a basic requirement for protecting privacy in electronic voting, auction, commerce, medical treatment and information interaction of military and intelligence departments. Anonymous communication technology,

The associate editor coordinating the review of this manuscript and approving it for publication was Zheng Yan.

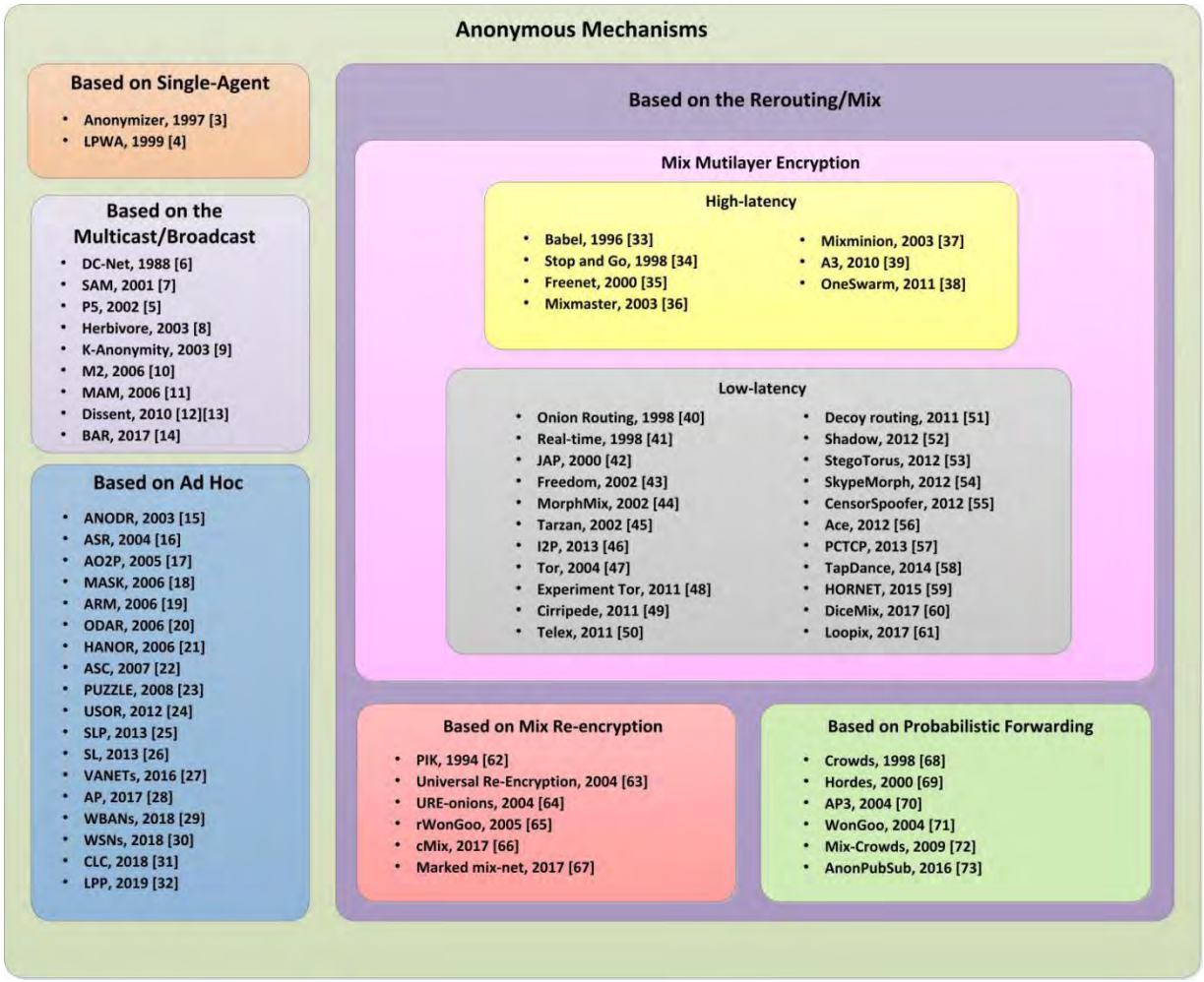

FIGURE 1. **Different types of anonymous systems based on different anonymous mechanisms.**

which mainly focuses on how to hide the identities or address information of one side or both sides in communications, emerges as a critical topic.

Anonymous communication aims to preserve communication privacy within the shared public network environment [1]. Since Chaum [2] proposed the notion of Mix and the idea that hiding the email communication between the user and the server by using several relay servers called Mix, many anonymous communication systems have been proposed, based on different anonymous mechanisms, as shown in Figure 1. We now elaborate different anonymous mechanisms as follows.

Anonymous systems based on single-agent, such as Anonymizer [3] and LPWA [4], can forward messages between the client and the server through one agent, which can make some modifications or hide the address information. This type of anonymous systems is simple and easy to use. For example, the Anonymizer [3] can protect users' privacy when surfing the Internet, by setting up a third-party website (http://www.anonymizer.com) as a middleman between the user and the site to be visited. However, it can provide little anonymity and is vulnerable to attacks. Once the agent is compromised, all information would be exposed, including the address information.

Anonymous systems based on multicast/broadcast can achieve anonymity through one-to-many communications among hosts. For example, P5, which is designed for providing scalable anonymity, is an anonymous protocol based on broadcast [5]. P5 creates a broadcast hierarchy, in which different levels of hierarchy provide different levels of anonymity at the cost of communication bandwidth and reliability. In P5, all messages sent to a certain receiver are from a single upstream node, thus the receiver doesn't know the original message sender, and the sender also doesn't know where the receiver is or which host or address the receiver is using. With the use of the multicast or broadcast technology, senders send messages to a group of recipients, which looks the same. The larger the number of group members is, the less possible that an attacker could guess who the real receiver is. Compared to the systems based on single-agent, this type of

systems could obtain more anonymity. Anonymous systems based on multicast/broadcast include DC-Net [6], SAM [7], P5 [5], Herbivore [8], K-Anonymity [9], M2 [10], MAM [11], Dissent [12], [13], and BAR [14].

An Ad Hoc network is composed of a group of wireless mobile nodes. In such a network, all nodes are equal and able to join or quit the network at any time. Due to the nature of ad hoc networks, privacy of the users is at a greater risk than in traditional networks. Different from the cable network structure, the anonymous communication protocol for cable networks can't be applied to ad hoc networks. In the literature, several anonymous systems based on ad hoc networks were developed, including ANODR [15], ASR [16], AO2P [17], MASK [18], ARM [19], ODAR [20], and so on [21]–[24]. In recent years, several systems and protocols based on Wireless Sensor Networks (WSN) have been developed [25]–[32]. However, most of these systems try to provide mutual anonymity and achieve anonymity via a predetermined path, which may suffer from unreliable delivery and high processing overheads.

Anonymous systems based on Mix/Rerouting can be divided into three sub-types: mix multi-layer encryption, mix re-encryption and probabilistic forwarding, in which the systems based on multi-layer encryption can be further divided into high-latency and low-latency. Anonymous systems based on high-latency mix multilayer encryption include Babel [33], Stop and Go [34], Freenet [35], Mixmaster [36], Mixminion [37], OneSwarm [38] and A3 [39]. Anonymous systems based on low-latency multi-layer encryption include Onion Routing, JAP, I2P, Tor, SkypeMorph, Cirripede, TapDance, CensorSpoofer, and so on [40]–[61]. Anonymous systems based on mix re-encryption include PIK [62], Universal Re-encryption [63], URE-onions [64], rWonGoo [65], cMix [66], and Marked mix-net [67]. The above two types of anonymous systems mainly obtain anonymity through one or several intermediate nodes called Mix. A Mix is a message pool to store messages from former nodes and then send the messages in a confusing order. In this way, attackers couldn't detect the corresponding relationships between senders and receivers. For example, Tor network can provide users with low-latency anonymous communication, by building circuits with publicly listed relays to anonymously reach the destinations. Later, Tor envisions the possibility of unlisted entry points to Tor network, called bridges, since the publicly listed relays can be easily blocked. However, bridges can still be detected by powerful censors by observing the communications between bridges and user nodes. Anonymous systems based on probabilistic forwarding include Crowds [68], Hordes [69], AP3 [70], WanGoo [71], Mix-Crowds [72], and AnonPubSub [73]. In this type of anonymous systems, there are a group of senders and receivers cooperating with each other in forwarding messages. The intermediate node works not to store messages, but to immediately make choice of sending the message to a next intermediate node randomly or the real receiver with certain probability.

Different anonymous systems may provide different anonymity strength, while how to measure the anonymity is a new research focus. Though there has been some research on evaluating and measuring anonymity of anonymous systems [74]–[79], [108], [180], a systematic research on measuring anonymity is still lacking. Therefore, in this paper, we give a thorough formalization of the notion of anonymity, and review and examine the various approaches and metrics for measuring anonymity. Particularly, the approaches based on information theory for measuring anonymity are discussed and reviewed from several perspectives.

The rest of this paper will be organized as follows. Section II introduces some methods and definitions for formalizing the anonymity with respect to process calculi, epistemic logic, function views, UC framework, differential privacy, probabilistic automata, and I/O automata. Section III elaborates and analyzes different approaches and metrics for measuring anonymity. The metrics based on information theory are further elaborated in Section IV, since this direction is very important for the research of anonymity measurement and consists of various branches. Section IV presents a comprehensive survey on efforts for measuring anonymity based on the information theory from various perspectives. Finally, a conclusion is given.

II. FORMALIZATIONS OF ANONYMITY

In 1998, Reiter and Rubin [68] defined the degree of anonymity as the probability that an intruder can assign to a player of being the original sender of a message. To ensure anonymity, it requires an appropriate set of other subjects with the same attributes potentially. Thus, anonymity can be viewed as un-identifiability of a subject within a set of subjects (the anonymity set). To possibly quantify anonymity, Pfitzmann and Hansen gave a more precise and complicated definition of anonymity in paper [80]. They proposed the definition from the perspective of attackers: anonymity of a subject means that attackers cannot sufficiently identify the subject within a certain set of subjects, the anonymity set. The realization of anonymity of a certain entity needs the action of other entities to hide his/her own action, and thus it is essential for him/her to blend into a set of entities. The strength of anonymity of a subject may change when the surrounding environment is changed. For example, the subject uses the network differently or uses a modified communication network. The anonymity delta can be defined as the difference between the anonymity of the subject taking into account the attacker's observations and that when given the attacker's prior knowledge only [80]. The strength of anonymity of a subject can be quantified in a concrete network, so is the anonymity delta.

The efforts on the formalization of anonymity can be classified into methodologies based on process calculi, epistemic logic, function views, UC framework, differential privacy, probabilistic automata, and I/O automata, as summarized in Figure 2. We now elaborate them in the following sub-sections.

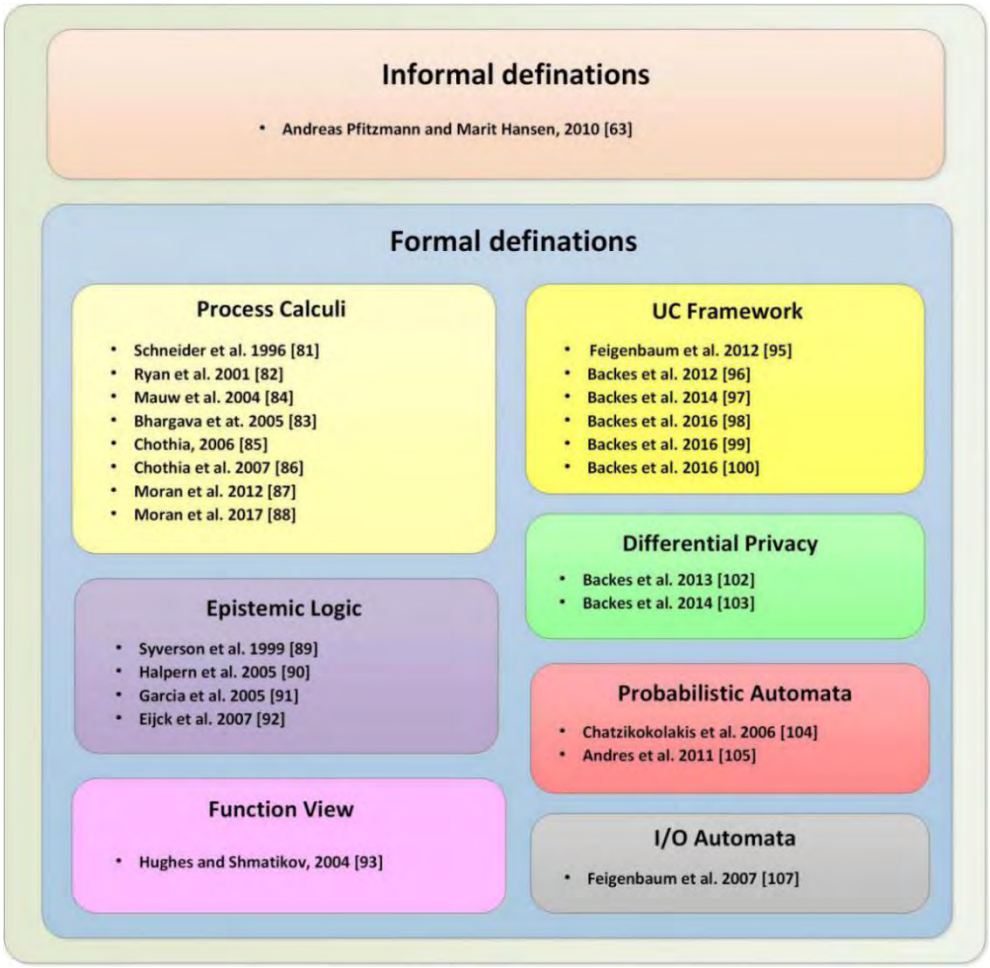

FIGURE 2. Different formalizations of anonymity.

A. FORMALIZATION FROM THE VIEW OF PROCESS CALCULI

Process calculi is a diverse family of related approaches for formally modeling concurrent systems. The framework and techniques of process calculi have been used extensively in the area of security, to formally define security properties and to verify cryptographic protocols. Recently, process calculi were also used for the formalization of anonymity [81]–[88].

They are simple and can be used to specify systems and system properties. With the use of calculi, such as communicating sequential processes (CSP), pi-calculus and so on, a system can be modularized. Model-checker can be used to verify properties of all the modular as well as the entire system.

In 1996, Schneider and Sidiropoulos [81] proposed a definition of anonymity with the use of CSP notation. The ideal anonymous property can ensure that a message from the true sender could be considered equally with the other members in the anonymous system. In CSP, the message of the form $i.x$ is equal to the form $j.x$. The set USERS consist of all users who are collaborating with each other in the system to provide anonymity. Then the set of messages used to make confusion for a given piece of information x can be regarded as set A:

$$A = \{i.x | i \in USERS\}.$$

A process P is strongly anonymous on an alphabet A

$$if : f_A^{-1}(f_A(P)) = P$$

where equality is with respect to traces, and

$$f_A(x) = \alpha \ if \ x \in A$$
$$f_A(x) = x \ if \ x \notin A$$

where $\alpha \notin \Sigma$, and Σ is the set of all possible events.

These equations state that the process $f_A(P)$ is identical to the original process P, which means that whenever any event from the set A happens, it is equally well considered to be any other event in the attacker set.

In 2005, Bhargava and Palamidessi [83] described the notion of anonymity by combining both probability and nondeterminism, and the general situation in which both the system and users can have both probabilistic and nondeterministic behaviors. They assume that nothing may be known about the relative frequency by which the users perform

the anonymous action, and build a model of computation to express both probabilistic and nondeterministic choices. Mauw *et al.* proposed a formal definition of the anonymity in presence of an observing intruder [84]. They apply an information theoretic method to measure and analyze the degree of anonymity of the anonymous communication systems such as the onion routing, and validate their definition of anonymity with a probabilistic analysis of the onion routing protocol in a process algebraic framework.

In 2006, Chothia [85] used the pi-calculus to analyze the MUTE system for anonymous file-sharing. They build pi-calculus models of a node that is innocent of sharing files, and a node that is guilty of file-sharing and of the network environment. In 2012, Moran *et al.* [87] proposed formal definitions of anonymity for voting protocols using the process algebra CSP. With using the process algebra CSP, they analyze a number of anonymity definitions, and give formal definitions for strong and weak anonymity by highlighting the differences between these definitions. For analyzing voting systems, they show that the strong anonymity definition they proposed is too strong and the weak anonymity definition they proposed is proved to be a suitable property.

Later, in 2017, Moran and Heather [88] proposed a novel intruder model for automated reasoning about anonymity and secrecy properties of voting systems. With using CSP and FDR model checker, they adapt a lazy spy which not only observes a protocol run, but also interacts with the protocol. This model also avoids the eagerness of pre-computation of unnecessary deductions.

B. FORMALIZATION FROM THE VIEW OF EPISTEMIC LOGIC

Epistemic modal logic is a subfield of modal logic that is concerned with reasoning about knowledge [89]. In 1999, Syverson and Stubblebine [90] set out S5 axioms and rules for propositional and epistemic logic related to anonymous systems. They propose that the anonymity of a system is the natural ability to hide certain facts about a principal or a set of principals from the adversary. They provide precise notion of the anonymity properties with the use of epistemic expression.

In 2005, Halpern and O'Neill [91] provided several definitions of anonymity with respect to agents, actions and observers in multi-agent systems, which consist of some agents in the local state at a given point in time and whose global state includes the local state of each agent and the state of environment. Their research focuses on providing an appropriate semantic framework to consider anonymity. While the former one pays more attention to formalizing various actions and facts, including sending, receiving, encrypting and so on.

In 2005, Garcia *et al.* [92] proposed a formal framework for the analysis of information hiding properties of anonymous communication protocols in terms of epistemic logic. They illustrate the approach by providing sender anonymity and unlinkability for two anonymizing protocols, Onion Routing and Crowds. In their approach, they express various information hiding properties in the language of epistemic logic, which makes it possible to reason about the messages in a run and about the knowledge agents gain from these messages.

C. FORMALIZATION FROM THE VIEW OF FUNCTION VIEW

In 2004, Hughes and Shmatikov [93] introduced a modular framework for expressing information hiding properties with the use of function and function knowledge, which represents an attacker's partial knowledge of a function and includes three elements, namely graph, image and kernel. They describe system behaviors with the use of a series of functions. Anonymity in their paper is the degree of a function being opaque to the attackers. If an attacker could connect the function of an action with the agent that performed it, the anonymity is lost.

D. FORMALIZATION FROM THE VIEW OF UC FRAMEWORK

The universally composable framework (UC framework) is a general framework for representing cryptographic protocols and analyzing their security, which was proposed by Canettiin 2001 [94]. The framework allows specifying the security requirements of any cryptographic task in a unified and systematic way. The UC framework is often used as a standard way to express the function and security properties of cryptographic protocols. Feigenbaum *et al.* [95] employed the UC framework, which abstracts essential properties of the onion routing in the presence of an active adversary, for probabilistic analysis of onion routing. They build a black-box model of anonymous communication in the UC framework, analyze the probabilistic user action and protocol operations, and analyze relationship anonymity in the onion routing model built in the UC framework.

In 2012, Backes *et al.* [96] proposed a security definition for the onion routing methodology in the universal composability (UC) framework. The definition in the UC framework, as an ideal functionality in their paper, gives appropriate considerations to the goals of various system entities.

Backes *et al.* [97] in 2014 extended the UC framework by proposing the TUC framework (Time-sensitive Universal Composability) which incorporates a notion of time while preserving universal composability, for time-sensitive and modular analysis of the Tor anonymous systems. They argue that the previously proposed frameworks for strong anonymity guarantees of the Onion Routing protocol that underlies Tor anonymous systems are not capable of modeling the class of traffic-related timing attacks against Tor (e.g., the website fingerprinting and the traffic correlation). Due to the lack of the communication model that enables a composable security analysis of complex protocols against time-sensitive adversaries, it is hard to include timing-sensitive attacks into rigorous analysis. The authors argue that the previous frameworks allow unrestricted activation orders: e.g., it might happen that a message that was sent in the past (over a direct connection) arrives after a time-out mechanism already closed a port, only because the

sending party was not activated early enough. In their work, they show that the TUC framework, which incorporates a comprehensive notion of time in an asynchronous communication model with sequential activation, allows for rigorously proving strong anonymity guarantees in the presence of time-sensitive adversaries who mount traffic-related timing attacks.

In 2016, Shirazi *et al.* [98] proposed a classification to centralize the unique routing characteristics, deployability, and performance of all commonly considered approaches (including Mixnets, DC-nets, onion routing, and DHT-based protocols), and to survey previous research on designing, developing, and deploying systems for anonymous communication. This includes the topology of the underlying network; the routing information that must be made available to the initiator of the conversation; the underlying communication model; and performance-related indicators. The taxonomy and comparative assessment provide important insights into the differences between existing anonymous communication protocol categories, and it helps to clarify the relationship between the routing characteristics of these protocols.

In 2016, Backes *et al.* [99] defined the concept of Anonymous RAM (Anon-RAM) and provide a provable security construction. Anon-RAM is a novel multi-user storage primitive that provides strong privacy and integrity guarantees. Anon-RAM combines privacy features of anonymous communication with oblivious RAM (ORAM) schemes, allowing it to simultaneously protect the privacy of content, access patterns and user's identity. Anon-RAM further protects integrity by preventing malicious users from corrupting other users' data. They proposed two secure Anon-RAM schemes with different design and time-complexity. The first scheme is simple in design; like efficient ORAM schemes, its time-complexity is poly-logarithmic in the number of cells (per user), but it is linear in the number of users. The second scheme reduces the overall complexity to poly-logarithmic in the total number of cells (of all users), at the cost of requiring two (non-colluding) servers.

Also, in 2016, Backes *et al.* [100] proposed a relative linkability measure to rank identities within social media sites by their anonymity. They analyze whether anonymity in a single social media site can protect a user from being across-site linked. They show that anonymity alone is not enough to assess linkability risks by evaluating this measure on their data set. They then mitigate this insufficiency and propose the absolute linkability measure that uses information about matching identities.

E. FORMALIZATION FROM THE VIEW OF DIFFERENTIAL PRIVACY

In cryptography, differential privacy aims to provide means to maximize the accuracy of queries from statistical database while minimizing the chances of identifying its records. In 2008, Dwork introduced differential privacy which can intuitively capture the increased risk to one's privacy incurred by participating in a database and can achieve any level of privacy [101]. For privacy preserving computations, the notion of differential privacy is a standard for quantifying privacy. Informally, differential privacy of a mechanism guarantees that the mechanism doesn't leak any information of a user to even an adversary who has auxiliary information about the rest of the user base. It has also been generalized to protocols against computational bounded adversaries, which has led to the notion of computational differential privacy.

In 2013, Backes *et al.* [102] proposed a framework for analyzing anonymous communication protocols, AnoA, in which a formal definition of anonymity was given based on a novel generalization of differential privacy. The security notion in the framework is based on interacting Turing Machines. Since differential privacy doesn't allow for leakage of data, it cannot be directly used in anonymous communication protocols which inherently leak to the recipient the data sent from a sender to the recipient. They generalized the original computational differential privacy to allow more fine-grained notion of adjacency, considering arbitrary protocols in contrast to incorruptible and monolithic mechanism, and to grant the adversary the possibility of compromising parties in the mechanism to accurately model the adversary. Through the modeling and formalization of anonymity of anonymous communication protocols based on the generalization of differential privacy, complex communication systems such as Tor and their different anonymity properties can be analyzed in a unified manner in the framework.

F. FORMALIZATION FROM THE VIEW OF PROBABILISTIC AUTOMATA

The growing concern about anonymity on the Internet, results in lots of work on formalization and verification of anonymity, in particular, the probabilistic aspects of anonymity [104].

It is well known that the raising of nondeterminism, due to the possible interleaving and interactions of parallel components, can cause unintended information leaks. In 2011, Andres *et al.* [105] studied the anonymity in probabilistic concurrent systems form the view of probabilistic automata, by the presence of randomization and concurrency. They try to solve the above problem of untended information leaks by fixing the strategy of scheduler beforehand. In their work, they define the notion of strong probabilistic anonymity under various notions of observables, and propose a sufficient technique to prove strong probabilistic anonymity based on automorphisms.

In the purely nondeterministic setting, the strong anonymity of a system is often defined and proved as follows: Take two users A and B and a trace in which user A is the culprit. Find a trace that looks the same to the adversary, but in which user B is the culprit. This trace can be easily obtained by switching the behaviors of A and B.

G. FORMALIZATION FROM THE VIEW OF I/O AUTOMATA

The input/output automaton model [106], developed by Lynch and Tuttle, is a labeled transition system model for

components in asynchronous concurrent systems. The actions of an I/O automaton are classified as input, output and internal actions, where input actions are required to be always enabled. An I/O automaton has ''tasks''; in a fair execution of an I/O automaton, all tasks are required to get turns infinitely often. The behavior of an I/O automaton is describable in terms of traces, or alternatively in terms of fair traces. Both types of behavior notions are compositional.

In 2007, Feigenbaum *et al.* employed an I/O automata, which provides asynchronous computation and communication, in their work [107] for modeling the famous anonymity protocol Onion Routing with provable anonymity. In their work, they analyze the anonymity of Onion Routing in their I/O automata model, and define anonymity and unlinkability with respect to an adversary in the model.

H. SUMMARY AND COMPARISONS

This section provides an overview of different formalizations of the notion of anonymity in anonymous communication systems from various views. The formalizations are reviewed and compared in Table 1, which describes the year, whether based on passive observer or active attacker, whether applied on real anonymous systems, and main features and comments on each formalization.

We can see that, with the use of CSP and ACP, formalization of the notion of anonymity is well studied and applied on different anonymous systems, however this type of formalizations don't consider active attackers in their formal definitions. In [93], the notion of function view is proposed to represent a mathematical abstraction of partial knowledge of a function in formalizing anonymity. It demonstrates how any formalism for system specification that provides an equivalence relation on system configurations induces function views, and how information hiding properties could be stated naturally in terms of opaqueness of these views. However, the framework and formalization proposed is still a theoretic approach and lacks a practical verification. In the view of epistemic logic, the notions of probabilistic anonymity and conditional anonymity are proposed and formalized, and modal logic is used to axiomatize anonymity. Also, the use of process algebras makes it easy for specifying systems as well as system properties, and the properties of the system described can be easily verified by model-checkers.

In addition, since many efforts on formalizing the notion of anonymity, like formalizations from the views of process calculi [81]–[88], epistemic logic [89]–[92], function view [93], probabilistic automata [104], [105] and so on, have only been applied to simple protocols such as DC-net, it is not clear if these frameworks or methodologies can capture an adversary with auxiliary information in a more complex protocol in real application environments. For these formalizations which are only applied to the protocol DC-net, one question is whether they can also be applied to the DISSENT system since DISSENT is an implementation of DC-net. It seems difficult to model complex protocols like the onion routing and their traffic analysis attack, and only the formalizations from the views of the UC framework and differential privacy are applied to the Tor network, the most successful anonymous system. However, these formalizations don't consider the adaptively corrupting adversaries and active attacks on Tor such as selective denial-of-service attacks.

III. METRICS FOR MEASURING ANONYMITY

The design of anonymous communication systems is a relatively new field, but the desire to quantify anonymity that the anonymous systems offer has been an important challenge since the beginning [108]. It is of critical importance that not only can we quantifying the anonymity these anonymous systems can offer, but that the metrics used to represent realistic expectations can be expressed clearly and the implementations actually offer the anonymity they promise. In this section, we investigate various metrics for measuring anonymity of anonymous systems, based on different theories and techniques, as summarized in Figure 3.

A. METRICS BASED ON SET THEORY

Chaum [6] introduced the notion of an anonymity set as the set of participants who are likely to be the sender or the recipient of a particular message. The anonymity set is used to hide the real sender or recipient. As the size of the set increases, so does the degree of anonymity.

Reiter and Rubin [68] define the anonymity of one particular user as $1-p$, where p is the probability of the user being the original sender. For calculating the degree of anonymity of the whole system, the user number needs to be considered.

Berthold *et al.* [42] defines the degree of anonymity as $A = \log_2 N$, where N is the number of all possible senders of a message in an anonymous system. It is clear from the definition that the degree of anonymity merely depends on the number of the users. Obviously, the condition that an attacker may acquire more information through traffic analysis attack, flooding attack, collusion attack, timing attack and so on and then make a more precise guess about the sender is ignored. Under this condition, the uncertainty who the sender is declines, so does the degree of anonymity of the system. But this formula can't reflect this situation. Therefore, this quantitative method can only be applied to situations without attacks.

In 2001, Berthold *et al.* [79] evaluated the degree of anonymity in analyzing the disadvantages of free MIX routers. Free MIX router is a widely used method to provide anonymity service using MIXes, in which each participant can choose his/her own MIX router freely. In their work, they show that some attacks are possible in networks with freely chosen MIX routers and estimate their impact on the achievable degree of anonymity. The degree of anonymity can be defined by the size of the group of participants, i.e., the number of possible senders or receivers. The anonymity can be measured as $A = \log_2 N$, where N represents the number of possible senders.

In 2015, Chen *et al.* [109] proposed the concept of set-theoretic conditional anonymity by considering the threat

TABLE 1. Summary and comparisons of different formalizations of anonymity.

	year	Formal Definitions	Passive observer	Active attacker	Applications	Comments
Process Calculi	1996 [81]	strong anonymity	√	×	DC-Net	CSP, not apply to real-world anonymity protocols
	2004 [84]	anonymity group	√	×	Onion Routing	ACP, qualitative analysis methodology
	2005 [83]	anonymity system	√	×	DC-Net	Pi-calculus, combine both probability and nondeterminism
	2007 [86]	Choice/player anonymity degree	√	×	DC-Net and FOO voting scheme	ACP, automatically check the anonymity with μ-CRL.
	2012 [87]	Strong/weak anonymity	√	×	CVS and Prêt â Voter	CPS, state exploration problem
	2017 [88]	anonymity	√	√	Voting System	CPS, automated reasoning about anonymity
Epistemic logic	1999 [89]	anonymity	√	√	Anonymizer	give an epistemic characterization of anonymity
	2005 [90]	probabilistic/conditional Anonymity	√	×	DC-Net	formalize the definitions of probabilistic anonymity
	2005 [91]	Sender anonymity, Unlinkability	√	×	Crowds and Onion Routing	observational equivalence
Function View	2004 [93]	Anonymity, Privacy	√	√	NO	formalize and classify a range of anonymity and privacy properties
UC framework	2012 [95]	relationship anonymity	×	√	Onion Routing/Tor	formally model relationship anonymity in Tor.
	2012 [96]	security definition of onion routing	√	×	Orion Routing/Tor	define a provably secure Onion Routing protocol
	2016	taxonomy of	√	×	Mixnets, DC-nets,	differences between classes of anonymous communication

from non-probabilistic adversary. They also proposed a metric for set-theoretic conditional anonymity and a variant of an existing metric for probabilistic conditional anonymity to evaluate system's degree of anonymity quantitatively. They show that when adversary obtains more observable outputs from the system, the system loses more anonymity,

TABLE 1. *(Continued.)* **Summary and comparisons of different formalizations of anonymity.**

	[98]	anonymous protocols			onion routing	protocols
	2016 [99]	definition of Anonymous RAM	√	×	RAM	define the concept of Anon-RAM
	2016 [100]	rank identities by anonymity	√	×	social media system	propose a relative linkability measure to rank identities by anonymity
Differential Privacy	2013 [102]	Sender/relationship anonymity, unlinkability	√	×	Tor	present a generic framework for defining, analyzing, and quantifying anonymity properties
	2014 [103]	Sender/Recipient/Relationship Anonymity	√	×	Tor	present a framework for rigorously assessing the degree of anonymity in the Tor network
Probabilistic automata	2006 [104]	anonymity system	√	×	Crowds	propose a formalization of probable innocence
	2011 [105]	Strong anonymity	√	×	DC-Net	prove probabilistic strong anonymity based on automorphisms
I/O automata	2007 [107]	Sender/receiver anonymity, unlinkability	×	√	Onion Routing	give an asynchronous model using IO automata

confirming the intuition of observing reveal sensitive information.

B. METRICS BASED ON INFORMAL CONTINUUM

In 1998, Reiter and Rubin *et al.* [68] introduced the Crowds anonymous communication system, which achieves anonymity by grouping users into a large and geographically diverse group that collectively issues requests on behalf of its members. In their paper, they introduce the notion of degree of anonymity for describing and proving anonymity properties of Crowds.

They argue that the degree of anonymity against an attacker can be viewed as a continuum with six levels, ranging from no anonymity (provably exposed), where the attackers can prove the sender, receiver and their relationship, to complete anonymity with some intermediate points: beyond suspicion, probable innocence, possible innocence, and exposed. The definition and proofs in their work can also be used for proving anonymity property of other anonymous systems and approaches.

C. METRICS BASED ON EVIDENCE THEORY

In the 1960s, Dempste [110] proposed to use probabilistic upper limit and lower limit to express uncertainty in reality. Later in 1976, Shafer developed this into an imprecise reasoning theory, called Dempster-Shafer (D-S) evidence theory [111], [112]. This theory uses the prior probability assignment function to find posteriori evidence interval, which is expressed by reliable function and the likelihood function, and helps to better grasp the uncertainty and unpredictability of a proposition. It provides a way to make evidence fusion.

The D-S evidence theory is based on a nonempty set $\boldsymbol{\Theta}$, which is composed of mutual and exhaustive elements. Any proposition being discussed belongs to power set 2^{Θ}. $M(A)$, is be a basic probability assignment function defined in 2^{Θ}. $M_1, M_2, \ldots, M_n$ are n probability assignment functions, and their fusion is $M = M_1 \oplus M_2 \ldots \oplus M_n$.

In 2006, Dijiang [113] introduced how to measure the anonymity of wireless mobile ad-hoc networks from the views of adversary, using evidence theory based on generalized information theory. In evidence theory, evidence is a concrete measurement of the real work. In his approach, it is the number of detected packets within a given time period. Huang makes an assumption that adversaries can detect, capture and monitor the traffic and locate the signal source, while cannot decrypt the content of captured frames. A captured packet is an evidence of communications between mobile nodes. According to the numbers of captured packets, an adversary could make all possible mappings of the mobile nodes. Then the adversary would continue the monitoring and capturing

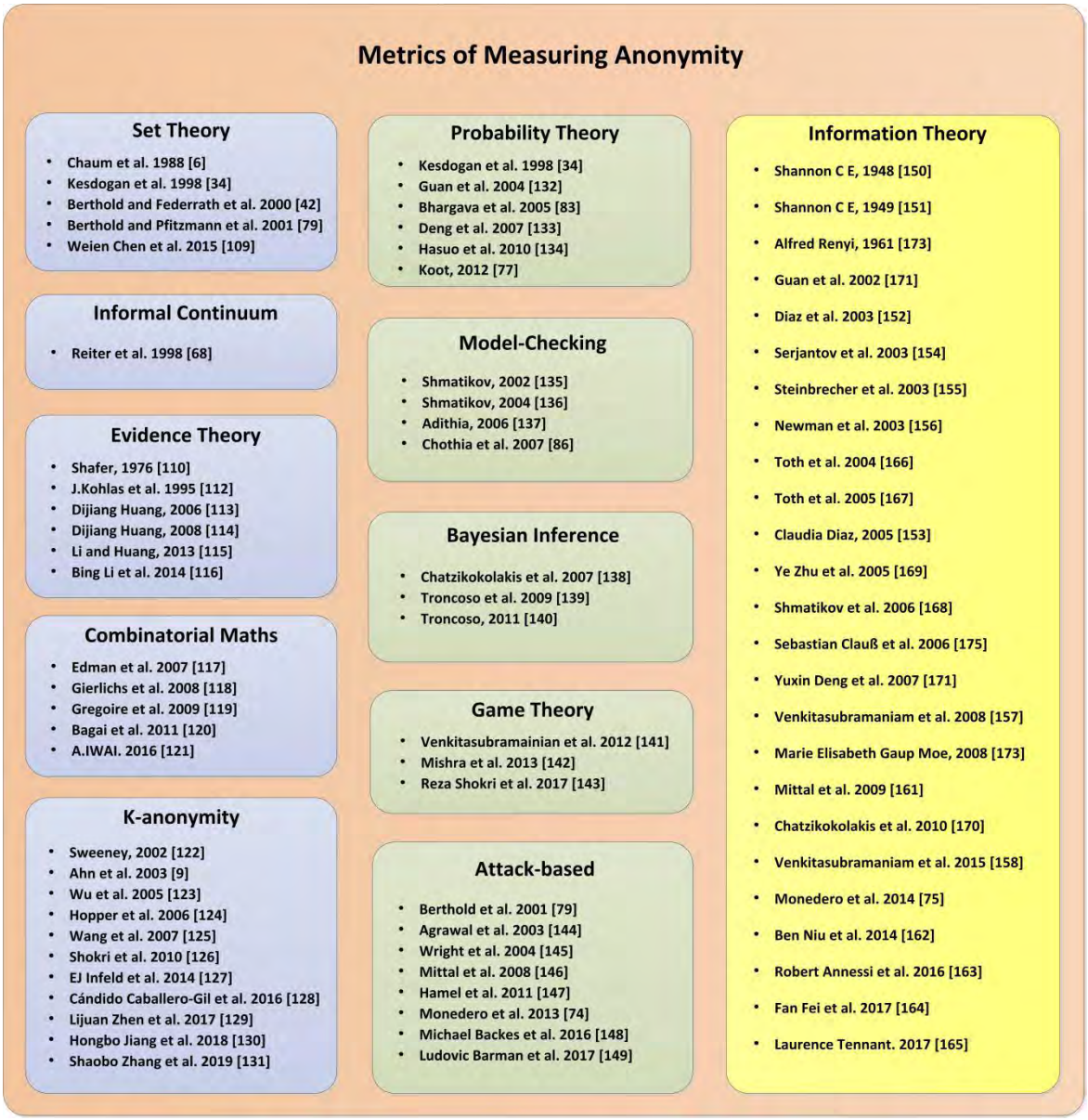

FIGURE 3. Summary of different types of metrics for measuring anonymity.

for some certain time. Different numbers of packets captured in different places can be linked to different sources and support different mappings. This method is a kickoff initiative for measuring the anonymity of wireless mobile ad-hoc networks. Later, in 2008 Dijiang [114] proposed a two-step unlinkability measuring approach for MANET, i.e., evidence collection using statistical packet-counting traffic analysis and evidence theory-based unlinkability measure. However, it is a theoretic approach based on evidence theory, and localization errors and scalability issues are not considered in the system assumption. In 2013, Li and Huang [115] further developed this theoretic approach to incorporate localization errors in anonymity analysis and propose the notion of super-nodes to model group based on mobility. In 2014, Li *et al.* [116], based on the previous approach, developed a comprehensive evidence based on method to handle the information that a monitoring system can acquire in realistic models and the corresponding analysis approach to process the various evidences from multiple sources.

D. METRICS BASED ON COMBINATORIAL MATHS

Many works on measuring the anonymity of anonymous communication system focus on the level of anonymity from the perspective of a single user or message. However, it is not clear how to generalize such a metric to clearly express a system-wide anonymity level. Edman *et al.* [117] first proposed a new system-wide metric, based on the permanent of a matrix, which measures the amount of additional information

needed by an observer to reveal the whole communication pattern between senders and recipients in a mix-based anonymity communication system as a whole. Instead of computing the size of the anonymity set for a given message, they consider simultaneously all incoming and outgoing messages in an anonymous communication system. The metric based on the permanent of a matrix can be used alongside the matrices which typically measure the anonymity from the perspective of a single user or message, as a complementary tool to represent a reasonable and intuitive combination of the individual level of anonymity and a more complete system-wide perspective.

They use the permanent of a (0,1)-matrix of size $n \times n$ which captures a scenario where one can model the feasibility between inputs and outputs as a (0, 1) relation, as the basis of the permanent-based metric. Consider an anonymous communication system, in which there exists an observer who is able to observe some or all the messages entering or exiting the anonymity system, and an input or output of the anonymous system may be a message or flow. Given a doubly stochastic matrix P representing the probabilities of input-output relationships in an anonymous system, the degree of anonymity defined by Edman *et al.* [117] can be expressed as:

$$d(P) = \begin{cases} 0 & n = 1 \\ \dfrac{\log(per(P))}{\log(\frac{n!}{n^n})} & n > 1 \end{cases}$$

where $n!/n^n$ is the minimum value of the permanent of a $n \times n$ doubly stochastic matrix, conjecturing by the Vander Waerden.

Later, after Edman et al. proposed their combinatorial approach for measuring the system anonymity level, Gierlichs et al. proved that the metric proposed by Edman fails to capture the anonymity loss caused by subjects sending or receiving more than one message [118]. They redefine the metric and generalize the system's anonymity level in scenarios where user relations can be modeled as yes/no relations to cases where subjects send or receive arbitrary number of messages. The key difference between the two matrices is that Gierlichs et al. consider the relationships between the senders and recipients rather than between individual input and output messages. And their observation can also be applied to anonymity metrics that measure the size of anonymity sets.

Similarly, in 2009, Grégoire and Hamel [119] extended the combinational approach proposed by Edman *et al.* [117] to include the calculation information about how many messages were sent or received by a user and they define a new metric that can be computed exactly and efficiently using classical and elegant techniques from combinational enumeration. In 2011, Bagai *et al.* [120] analyzed the system-wide anonymity metric of Edman and showed that while the metric for narrower class of infeasibility attacks is sound, the metric for probabilistic attacks has shortcomings. They then proposed a more accurate metric for both infeasible attacks and probabilistic attacks.

In 2016, Iwai [121] proposed a prototype system review, particularly from the viewpoint of its usage of a combinatorial anonymity measure. The prototype can automatically prevent unintended information leakage using a framework that analyzes the input data to find elements that can lead to information leakage and a mechanism for correcting flaws by modifying the questionnaire design in the database.

E. METRICS BASED ON K-ANONYMITY

Sweeney and Anonymity in 2002 [122] first proposed the notion of k-anonymity, a formal protection model for protecting private data in information release. A release provides k-anonymity protection if the information for each person contained in the release cannot be distinguished from at least k-1 individuals whose information also appears in the release. The k-anonymity protection model is important because it forms the basis on which the real-world systems known as Datafly, μ -Argus and k-Similar provide guarantees of privacy protection.

In 2003, Ahn *et al.* [9] introduced the notion of sender and receiver k-anonymity. A communication protocol is sender k - anonymous if it guarantees that an adversary, trying to determine the sender of a particular message, can only narrow down his/her search to a set of k suspects. Receiver k-anonymity is defined similarly: an adversary, at best, can only narrow down the possible receivers to a set of size k.

An anonymous communication protocol for the message space M is a computation among n parties $P_1, \ldots, P_n$. Let $H \in \{P_1, \ldots, P_n\}$ denote the set of honest parties. A protocol P is sender k-anonymous if it induces a partition $\{V_1, \ldots, V_l\}$ of H such that:

1. $|V_S| \geq k$ for all $1 \leq s \leq l$; and
2. For every $1 \leq s \leq l$, for all $P_i, P_j \in V_s$ for every $(msg, p) \in M \times [n] \cup \{(nil, nil)\}$, $P(P_i(msg, p), *)$ and $P(P_j(msg, p), *)$ are computationally indistinguishable.

Each honest party's message are indistinguishable from those sent by at least (k-1) other honest parties. A protocol P is receiver k-anonymous if it induces a partition $\{V_1, \ldots, V_l\}$ of H such that:

1. $|V_S| \geq k$ for all $1 \leq s \leq l$; and
2. For every $1 \leq s \leq l$, for all $P_i, P_j \in V_s$, for every $P' \in H, msg \in M : P(P'(msg, P_i), *)$ and $P(P'(msg, P_j), *)$ are computationally indistinguishable.

Please refer to [9] for more details. In their work, they show that there exist simple and efficient protocols that are k-anonymous for both the sender and the receiver in a model where a polynomial time adversary can see all traffic in the network and can control up to a constant fraction of the participants.

In 2005, Wu and Bertino [123] proposed and investigated a zone-based k-anonymity routing base protocol, to achieve destination anonymity in positioning routing algorithms. They argue that initially setting anonymity zone large can help to meet the destination anonymity requirement for longer

time at a relatively low control overhead. The effectiveness of k-anonymity against traffic analysis and surveillance is analyzed in [124].

In 2007, Wang *et al.* [125] introduced a novel approach by taking advantages of the hierarchical ring structure and the mix technique. They prove that the protocol can provide k-anonymous for both the sender and the recipient, even if a polynomial time adversary can eavesdrop all network traffic and control a fraction of participants. In 2010, Shokri *et al.* [126] analyzed the effectiveness of k-anonymity approaches for protecting location privacy in the presence of various types of adversaries. They argue that the k-anonymity approaches, which construct the cloaking region based on the users' location privacy, do not reliably relate to location privacy and may even be detrimental to users' location privacy.

In 2014, Infeld *et al.* [127] analyzed how the sparsity of a typical aggregate social relation impacts the network overhead of online communication systems designed to provide k-anonymity. It is called symmetric disclosure that, once users are grouped into anonymity sets, there will likely be few related pairs of users between any two particular sets, and thus the sets need to be large in order to provide cover traffic between them. They argue that the associated overhead can be reduced by having both parties specify both the origin and the target sets of the communication.

In 2016, Caballero-Gil *et al.* [128] proposed a novel revocation scheme which is able to track and revoke specific malicious users only after receiving some complaints while otherwise guaranteeing node's k-anonymity. With widely evaluated with NS-2 simulator and an analytical model validated with scripts, the results show that their work can increase privacy protection while allowing revocation with little extra costs.

In 2017, Zhen *et al.* [129] proposed a clustering algorithm to realize the establishment of anonymous group in the anonymous model. The algorithm is based on the k-anonymity location privacy preserving model to eliminate outliers.

In 2018, Jiang *et al.* [130] proposed a robust location privacy preserving algorithm named RobLoP against LDA in continuous LBS queries. The key insight of RobLoP is to theoretically derive the constraints of both MMB and MAB in a uniform way. It provides a necessary condition of the pairwise user to be safely cloaked against LDA. RobLoP first identifies those candidate users who can be cloaked with the requesting user. RobLoP then searches for a set of so-called strict point set that include candidate set and other auxiliary points, as a sufficient condition under which RobLoP can finally generate the cloaked region successfully.

In 2019, Zhang *et al.* [131] proposed an enhanced user privacy scheme through caching and spatial K-anonymity (CSKA) in continuous LBSs; it adopts multilevel caching to reduce the risk of exposure of users' information to untrusted LSPs. In continuous LBS queries, the scheme first utilizes the Markov model to predict the next query location based on the user mobility. They then designed an algorithm for forming spatial K-anonymity to improve the user's cache hit rate and enhance the user location privacy according to the predicted location, cell's cache contribution rate and data freshness. The security analysis and simulation results show that CSKA scheme can provide higher privacy protection than a few previous methods and can minimize the overhead of the LBS server.

F. METRICS BASED ON PROBABILITY THEORY

Probabilistic anonymity is based on publicly known security parameter, which determines the security of anonymous protocols [34]. For probabilistic anonymity the insecurity, expressed as the probability of having only one honest participant, approaches 0 exponentially as the security parameter is changed linearly.

Given an attacker model E and a finite set of all users T, let R be a role for user (sender or recipient) in respect to a message M. If, for an attacker according to model E, the a-posteriori probability p that a user u has the role R in respect to M is non-zero ($p > 0$), then u is an element of the anonymity set $U \subseteq T$.

A technique (method) provides an R anonymity set of size n if the cardinality of U is n. An algorithm provides deterministic anonymity if n is always greater than 1. Though the technique presented provides deterministic anonymity, the user identities have to be verified to be secure. The notion of probabilistic anonymity comes from the technique above, with some changes to the information theoretic property.

Given an attacker model E, let AL be am algorithm providing anonymity with a complexity parameter p. We say that AL is probabilistically secure against the attacker model E if AL, for a distinct message M, can be broken with probability α and if

1. The a-posteriori probability of insecurity after any possible attack within the attacker model E is the same as the a-priori probability before an attack occurs, i.e. α remains constant for a given p, and
2. the probability of insecurity approaches 0 at an exponential rate as p is increased linearly.

If AL is probabilistically secure, it provides probabilistic anonymity.

In 2004, Guan *et al.* [132] quantitatively analyzed the anonymous communication systems, using a probabilistic method. They study the probability that the true identity of a sender can be discovered in an anonymous communication system, assuming a passive adversary model where some nodes may be compromised by the adversary. Several insightful results are obtained based on the quantitative analysis, among which are some interesting ones. Contrary to the intuition, the probability that the true identity of a sender can be discovered might not always decrease as the path length increases. Also, the complexity of the path topology does not have an important impact on this probability. While paths with complicated topology perform better than the simple ones, the difference is relatively small.

In 2005, Bhargavaand Palamidessi [83] proposed a notion of anonymity which combines both probability

and nondeterminism. It is suitable for describing the most general situation in which both the systems and users can have both probabilistic and nondeterministic behaviors. Later, Deng *et al.* [133] proposed a notion of weak probabilistic anonymity, considering the fact that some amount of probabilistic information may be revealed by the anonymous protocol, that may be gathered and used by a passive observer to infer the linkage relation. In their work, they study the degree of anonymity that the protocol can still ensure, despite the information leakage.

In 2010, Ichiro Hasuo and Kawabe [134] presented a simulation-based proof method for the above notion of probabilistic anonymity [83]. In particular, they obtain an appropriate notion of probabilistic simulation as an instantiation of generic definition, for which soundness theorem comes for free.

G. METRICS BASED ON MODEL-CHECKING

Model checking has been extensively used to analyze the properties such as security, authentication, and anonymity. Existing techniques for formal analysis of anonymity in a non-deterministic protocol model include traditional process formalisms such as CSP.

In 2002, Shmatikov [135] used probabilistic model checking to analyze the anonymity properties of the Crowds, which depends on the probabilistic behavior of protocol participants and can be formally expressed only in terms of relative probabilities of certain observations by the adversary.

Probabilistic model checking requires two inputs: a description of the system to be analyzed, typically given in some high-level modeling language; a formal specification of quantitative properties of the system to be analyzed.

They use discrete-time Markov chains to formally model the behavior of the group members in Crowds, specify anonymity properties of the system as temporal logic formulas, and use a probabilistic model checker to verify them. Using the probabilistic model checker PRISM, they show that, as the size of the group increases, the confidence of the corrupt members to detect the correct sender increases and thus results in degradation of the anonymity level of the system. Later, in 2004, Shmatikov [136] again used the probabilistic model checker PRISM to analyze the crowd system for web browsing and demonstrated how probabilistic model checking techniques can be used to formally analyze security properties of a peer-to-peer group communication system based on random message routing among members.

As a mathematical abstract of stochastic systems, Discrete Time Markov Chain (DTMC), Continuous Time Markov chain (CTMC) and Markov Decision Process MDP are three basic models for system description model checking. If there is only probability choice, DTMC is needed; if there are both probability choice and non-deterministic choice, MDP is needed; if there is no non-deterministic choice and modeling is for continuous time, CTMC is needed.

PRISM is probabilistic model-checker and can model and analyze probabilistic behavior. The input of it is the description (modeled by the discrete-time Markov chain) of the system to be analyzed and the anonymous properties to be checked. PRISM consists of several relatively independent modules with a series of variables. Different variables determine different anonymous states of the system. Accordingly, the output of it is another anonymous state of the system with new variables of each module after being attacked by traffic analysis and so on [137].

In 2006, Adithia [137] carried on a probabilistic analysis of the anonymity of Crowds, Adithia, Onion Routing and Tarzan with the use of PRISM and maked a comparison of different anonymous systems.

In 2005, Bhargava and Palamidessi [83] proposed a new way to research anonymity of anonymous systems. They formulate the notions of anonymity in terms of observables for processes and develop a model-checker for the probabilisticpi-calculus.

In 2007, Chothia *et al.* [86] used μ CRL tools to make a protocol specification and measure anonymity of an anonymous protocol automatically.

H. METRICS BASED ON BAYESIAN INFERENCE

The inference of the information concealed by the anonymous communication protocols can be regarded as a hypothesis-testing problem. In 2007, Chatzikokolakis *et al.* [138] considered the Bayesian approach to the above problem, and investigated the probability of error associated with the inference when the MAP (Maximum Aposterior Probability) decision rule is adopted.

By using the Bayesian inference, they study the probability of error (i.e., the probability of making the wrong guess). Assume a channel as a tuple $< A, O, p(\cdot|\cdot) >$, where A, O are the sets of input and output values respectively and p(o | a) is the conditional probability of observing output $o \in O$ when $a \in A$ is the input. The a priori and the posteriori probabilities of a are related by Bayes' theorem:

$$p(a \mid o) = \frac{p(o|a)p(a)}{p(o)}$$

Let $\theta_f \rightarrow [0, 1]$ be the function that associates to each $a \in A$ the probability that f gives the wrong input fact when $a \in A$ has occurred. Then, the probability of error for f is then obtained as the sum of the probability of error for each possible input, averaged on the probability of the input:

$$P_f = \sum\nolimits_a p(a)\theta_n(a).$$

Please refer to [KCP 2007] for more details. Also, they apply the methodology to the Crowds protocol and show how to compute the bounds on the probability that an adversary could break anonymity.

In 2009, Troncoso and Danezis [139] performed traffic analysis of anonymous systems, in particular the mix networks, in the context of Bayesian inference. They use a generative probabilistic model of mix network to build a Markov Chain Monte Carlo inference engine, that calculates the anonymity given an observation of network traces.

Troncoso in his PhD thesis [140] also use Bayesian inference in analyzing the anonymity degree of an anonymous system. He shows that Bayesian inference and the associated Markov Chain Monte Carlo sampling techniques form an appropriate framework to evaluate the resistance of anonymous communication systems to traffic analysis.

I. METRICS BASED ON GAME THEORY

The package transmission times of nodes in a network can reveal significant information about the source-destination pairs and the routes of traffic flow in the network. Such information may be gathered by adversaries. While anonymous networking is the act of communicating over a network without revealing the identities of source-destinations or path of flow of packets. In 2012, Venkitasubramaniam and Tong [141] investigated the problem of maximizing anonymity of an anonymous network from a game-theoretic perspective.

They argue that the typical design of anonymous networking protocols models adversaries as omniscient and capable of monitoring every single transmission in the network perfectly, which may not be practical. They study the anonymity in networks under a more general adversary model, in which an unknown subset of the nodes are monitored by the adversary.

Quantifying anonymity using conditional entropy of the routes given the adversary's observations, the problem of optimizing anonymity is formulated as a two players zero-sum game between the network designer and the adversary. The task of the adversary is to choose a subset of nodes to monitor so that anonymity of routes is minimum, whereas the task of the network designer is to maximize anonymity by choosing a subset of nodes to evade flow detection by generating independent transmission schedules. If the network designer is aware of which nodes of the network are being monitored by the adversary, the optimal set of nodes can be chosen such that minimum information is revealed through the monitored nodes. However, if the adversary is aware of the set of nodes that the network designer has chosen to protect, then he can alter his choice of nodes to monitor so that maximum information about the network routes is retrieved. Parv Venkitasubramaniam *et al.* mainly study the interplay of network designer and adversary using a game-theoretic approach.

They prove that in the two-player game, there is a unique saddle point equilibrium for a general category of finite networks. Applying the game-theoretic model, they consider a general class of parallel relay networks and introduce asymmetry into the properties of the relay rate and information model. They also show that the game-theoretic approach can be used to study large parallel relay networks.

In 2013, Mishra and Venkitasubramaniam [142] proposed a detection theoretic framework to study the optimization of mixing strategy of anonymous systems under the constraints on network resources such as memory and bandwidth. In their work, they propose a general game-theoretic model to study the mixing strategies when an adversary is capable of capturing a fraction of incoming packets.

In 2017, Shokri *et al.* [143] proposed a methodology that can design optimal user-centric location obfuscation mechanisms respecting each user's quality of service requirements, while maximizing the expected error that the optimal adversary incurs in reconstructing the user's actual trace. The methodology is based on the mutual optimization of user/adversary objectives (maximizing location privacy vs. minimizing localization error) formalized as a Stackelberg Bayesian game. This formalization makes the solution robust against any location inference attack.

J. METRICS BASED ON ATTACKS

Anonymity services hide user identity at the network or the address level, but are vulnerable to attacks involving repeated observations of the user. Agrawal and Kesdogan [144] argue that quantifying the number of observations required for an attack is a useful measure of anonymity. In their work, they analyze the disclosure attack, measuring its effects against user communications protected by an anonymity technique, and estimate how many times an attacker must observe a user's anonymous communication acts to find the user's all communication partners.

In 2001, Berthold *et al.* [79] showed that some additional attacks are possible in networks with freely chosen MIX routes. They estimated these attacks' impact on the achievable degree of anonymity and then evaluated the relevance of these attacks with respect to existing systems like e.g. Mixmaster, Crowds, and Freedom.

In 2004, Wright *et al.* [145] investigated the attacks by corrupt group members that degrade the anonymity of the anonymous protocols including Crowds, Onion Routing, Hordes, Web Mixes, and DC-net. Their results show that fully-connected DC-Net is most resilient to these attacks, but it suffers from the scalability issue.

In 2008, Mittal and Borisov [146] showed how the information leaks in the lookup mechanisms of structured peer-to-peer (P2P) anonymous communication systems can result in the degradation of anonymity.

Hamel *et al.* [147] proposed an attack-based approach for measuring the anonymity that the Tor network can provide in face of a partial network adversary, who observes and/or operates some of the network. They argue that the entropy metric for measuring anonymity has a number of shortcomings. The most important one is that the entropy measure is based on the holistic properties of the system, rather than the attacker. In their work, from the perspective of an adversary with a bandwidth budget, they derive theoretical and numerical results that illustrate path compromising in the Tor network. They propose that an attack-based measure would be superior for measuring anonymity in Tor network and to this end, they calculate the probability of path compromising given various bandwidth budgets available to the adversary, under three different Tor path selection algorithms.

In 2013, Monedero *et al.* [74] proposed a theoretical framework for privacy-preserving systems. The framework has a general definition of privacy according to the estimation error incurred by an attacker who aims to disclose the private information that the system is designed to conceal. They show that the framework permits interpreting and comparing a number of well-known metrics under a common perspective.

In 2016, Backes *et al.* [148] proposed a rigorous methodology to quantify the anonymity provided by Tor against various structural attacks, that is, adversaries that corrupt Tor nodes and perform eavesdropping attacks to deanonymize Tor users. First, they provide an algorithmic approach for computing the anonymity impact of structural attacks against Tor. The algorithm is parametric in the path selection algorithm and is capable of reasoning about variants of Tor and alternative path selection algorithms as well. Second, they present formalizations of various instantiations of structural attacks against Tor and show that the computed anonymity impact of each of these adversaries indeed constitutes a worst-case anonymity bound for the cryptographic realization of Tor. Third, they use the methodology to conduct a rigorous, large-scale evaluation of Tor's anonymity which establishes worst-case anonymity bounds against various structural attacks for Tor and for alternative path selection algorithms such as DistribuTor, SelekTOR, and LASTor.

In 2017, Barman *et al.* [149] proposed an anonymous communication protocol named PriFi for LANs provably secure against traffic-analysis attack which is application agnostic and has low latency. PriFi builds on Dining Cryptographers networks (DC-nets) and solves several of their limitations. For instance, the communication latency is reduced via a client/relay/server architecture tailored to LANs, where a set of servers assist the anonymization process without adding latency. Unlike mix networks and other DC-nets systems, a client's packets remain on their usual network path without additional hops. Moreover, PriFi protects clients from equivocation attacks without adding significant latency overhead or communication among clients.

In the next section, we will elaborate the metrics based on information theory separately, because this kind of measuring metrics has many branches and it is of great importance for the anonymity measurement research. A summary of all measuring metrics of anonymity will also be given at the end of the next section.

IV. METRICS BASED ON INFORMATION THEORY

In the middle of 19th century, a new science called information theory was established by Claude E. Shannon, with the publication of his two famous papers [150], [151]. Many approaches and techniques based on information theory are used to evaluate and measure the anonymity of anonymous communication systems. In this section, we will elaborate the metrics based on information theory from various perspectives: information entropy, min entropy/max entropy, mutual entropy, relative entropy, conditional entropy, Rényi entropy, and mutual information, as summarized in Figure 4.

A. INFORMATION ENTROPY

Information Entropy is defined as the emergence probability of a discrete random event. It measures the uncertainty of distribution and can be used to measure the degree of anonymity.

In 2002, Diaz *et al.* [152], [153] and Serjantov and Danezis [154] respectively quantified the anonymity with the use of information entropy. They showed that the size of an anonymity set is not inadequate for expressing anonymous systems, whose members are not equally likely to be the sender or receiver. The two papers are based on Shannon's information theory and take the probabilities of users being senders into account. Both suggested methods can be applied to evaluate the degree of anonymity that a system can provide under a particular attack and to compare that of different systems.

With a comparison of a pool mix and a traditional threshold mix, Serjantov proves that there are concerns in the method that only takes the anonymity set into consideration to calculate the anonymity degree of a system. There are always n messages in a pool mix whose initial n messages are dummy messages produced by the mix itself. When N new messages come into the mix, it selects n messages from the $n + N$ ones and forwards them to the next site. In this situation, the anonymity set should include the senders of all messages through this mix. While in a threshold mix, there are zero messages initially. If the number of new coming messages reaches N, all the N messages will be sent to the next site in a disrupted order. Serjantov also shows that traffic analysis on these systems with the maximum route length restriction is very powerful and it will reduce the degree of anonymity significantly. He defines the entropy of an anonymous system as:

$$S = -\sum_{u=1}^{n} p_u log_2(p_u)$$

where n is the number of users of the system and p_u is the probability that a user u acts as a role $r \in \{sender, recipient\}$ for a particular message.

In Claudia Diaz's method, there is a basic assumption that the number of senders is considered constant and the behavior of senders are regarded as an independent, identical, stationary stochastic Poisson process. The system model includes Senders, Recipients and Mixes. The attack model describes three properties of attackers: Internal-External, Passive-Active and Local-Global. Different combinations of the properties indicate different attack abilities. Then the corresponding attackers will assign probabilities of the users having sent a particular message after the attack and the distribution of probabilities. Therefore, it is not only the size of the anonymity set that determines the degree of anonymity of the attacked system. He defines the degree of anonymity as

$$d = \frac{S}{S_{max}}$$

where S is the current entropy of the system defined above and $S_{\max}$ is the maximum entropy of the system, that is $\log_2 n$ where n is the number of users in the anonymous system.

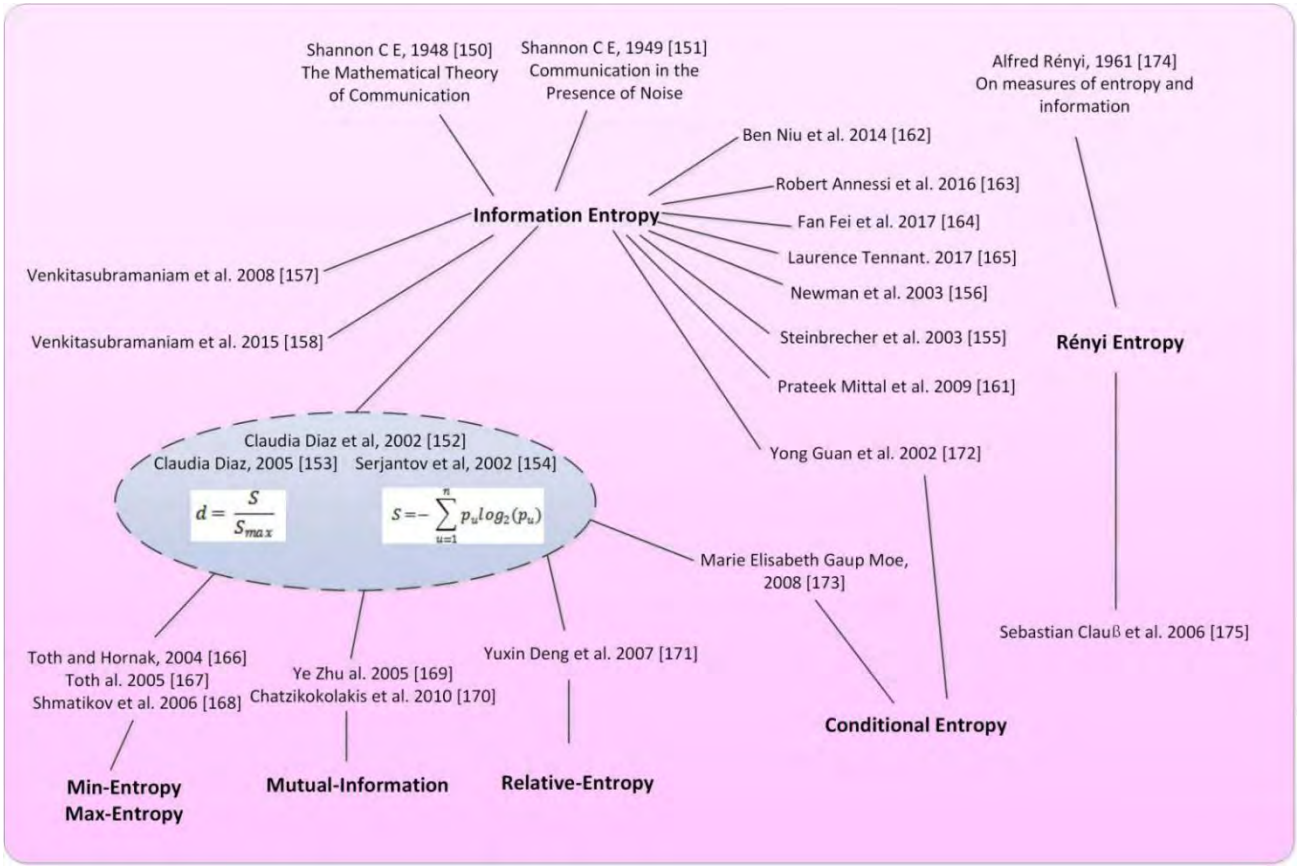

FIGURE 4. Metrics based on information theory.

Steinbrecher and Kopsell [155] generally and formally define the notion of unlinkability and study the impact of unlinkability on anonymity. They use the method of information theory to describe unlinkability probabilistically. Newman *et al.* [156] take an opposite approach to evaluate how much protection a Traffic Analysis Prevention (TAP) system can provide for its users. Its focus is on determining the overall amount of anonymity a TAP system can provide, or the amount it can provide for a given cost in padding and rerouting.

In 2008, Venkitasubramaniam and Tong [157] designed scheduling strategies for wireless nodes using directed signaling, with Shannon's equivocation as an information theoretic measure of anonymity. In 2015, Venkitasubramaniam and Mishra [158] studied the maximum anonymity achievable by the packet shuffling method when the mixes are memory limited, by developing an information theoretic framework. They use the Shannon information entropy of the posterior distribution of packet sources from an eavesdropper's perspective as the measure of anonymity and show that the maximum achievable anonymity of a single mix with buffer size b packets serving two independent Poisson sources and with equal arrival rates is shown to be $\log[2\cos\left(\frac{\pi}{b+3}\right)]$.

In addition, Chatzikokolakis [159], O'Connor *et al.* [160], Mittal *et al.* [161], and so on also make extending studies on measuring anonymity based on information entropy.

In 2014, Niu *et al.* [162] proposed a DummyLocation Selection (DLS) algorithm to achieve k-anonymity for users in LBS. Unlike existing approaches, the DLS algorithm carefully selects dummy locations considering that side information may be exploited by adversaries. They first select these dummy locations based on the entropy metric, and then propose an enhanced-DLS algorithm to ensure that the selected dummy locations are spread as far as possible. Evaluation results show that the DLS algorithm can significantly improve the privacy level in terms of entropy.

In 2016, Annessi and Schmiedecker [163] proposed and evaluated methods to measure and improve performance in the Tor network based on Shannon's information entropy theory. They use active Round-TripTime (RTT) to estimate the quality of circuits, and to validate the distribution of RTT values.

In 2017, Fei *et al.* [164] proposed a method for evaluating the degree of anonymity of LBS (Location Based Service) based on information entropy. This method reflects how the attacker identifies the user's real location in the anonymous candidate set, thereby further adopts the k-anonymity-based

privacy protection mode to provide lower-cost privacy protection.

In 2017, Tennant [165] applied the information entropy theory to the evaluation of the anonymity of the electronic ledger. He proposed that the behavior of each user in the cryptocurrency can affect the anonymity of other users in the system. As time goes by, attackers can get a large amount of user information, this reduces anonymity of the system greatly.

B. MIN-ENTROPY/MAX-ENTROPY

In 2004, Tóth and Hornák [166], [167] introduced the notion of source hiding and destination-hiding. They introduce the formal model of the PROB-channel and explore what conclusions a passive observer can draw by only knowing public parameters and the timing of events. They argue that, if an attacker can't assign a sender to any message with a probability greater than $\boldsymbol{\Theta}$, the system is source-hiding with $\boldsymbol{\Theta}$. Similarly, when it comes to the recipient with $\boldsymbol{\Omega}$, then the system is destination-hiding. They propose using min-entropy and max-entropy for measuring local anonymity.

Then, in 2006, Shmatikov and Wang [168] studied the relationship anonymity measurement in mix networks. They provide a formal definition and a calculation methodology for relationship anonymity, which is like the methodologies for sender anonymity because relationship anonymity is sensitive to the distribution of message destinations. Their methodology for measuring relationship anonymity cooperates the information-theoretic metrics of max-entropy and min-entropy. They also illustrate the methodology for measuring relationship anonymity in several simulated networks.

In 2014, Monedero *et al.* [75] found the optimal mix parametrization and characterizes the optimal trade-off between the contrasting aspects of anonymity and delay, for two information-theoretic measures of anonymity. Experimental results showed that mix optimization may lead to substantial delay reductions for a desirable level of anonymity.

C. MUTUAL INFORMATION

In 2005, Zhu and Bettati [169] proposed a quantitative metric based on mutual information, which takes into account possible heterogeneity. They model the effectiveness of a single mix or of mix networks in terms of information leakage and measure it in terms of covert channel capacity and describe the relationship between the anonymity degree and information leakage.

They argue that measures for anonymity in anonymous communication systems must be on one hand simple and concise, and on the other hand reflect the realities of real systems. In addition, implementation quality and topologies of the anonymity measures must be considered as well in the anonymity measure metrics. They propose an anonymity degree that generalizes the information theoretic definitions previously proposed to quantify anonymity, and also discuss the relationship between the anonymity metric proposed based on mutual entropy and other information theory-based metrics.

In their paper, they formulate the anonymity degree as a function of the attack effectiveness:

$$D = 1 - \frac{I([S,R]_a, [S,R]_s)}{\log(m \cdot n)}$$

where $[S,R]_a$ represent the random variable that describes the actual sender and receiver pair, $[S,R]_s$ represents the attack, i.e., the attacker's estimate of $[S,R]_a$ through observation of the system, and $\log(m \cdot n)$ is used to normalize the anonymity degree into the range of [0,1].

In their mutual-information based anonymity degree, the entropy-based degree is included by averaging according to the a priori probability of traffic between each pair. In comparison with other entropy-based definitions, the proposed definition can describe the anonymity provided by a network of mixes.

In 2010 Chatzikokolakis *et al.* [170] showed that measures of information leakage based on mutual information and capacity can be calculated, automatically, from trial runs of a system alone. They develop a tool to automatically perform this analysis, which is used to measure the loss of anonymity. They also show that for conditional mutual information, appropriate adaptation of the Blahut-Arimoto algorithm can provide approximations of the maximizing input distributions.

D. RELATIVE ENTROPY

In 2007, Deng *et al.* [171] proposed to use relative entropy as a distance of two discrete probability distributions for measuring the degree of anonymity guaranteed by protocols (which can be interpreted as a fully probabilistic automaton). They argue that measure metrics based on information entropy, which takes into account the probability distribution of the users performing certain actions, where the probabilities are assigned by an attacker after observing the system, don't consider the attacker's knowledge about the users before running a protocol. Thus they extend the approach by proposing using the relative entropy.

The method they proposed quantifies the amount of probabilistic information revealed by the protocol, i.e., how much information an attacker can obtain after observing the outcomes of the protocol, together with the information he has before the protocol running. For a protocol that contains both nondeterministic and probabilistic behaviors, their method can deal with two sets of probability distributions by using the Hausdorff distance.

In their work, relative entropy is used as a distance of two discrete probability distributions to measure anonymity.

Let θ *and* θ' be two discrete probability distributions on a set S. The relative entropy can be defined by

$$D\left(\theta, \theta'\right) = \sum_{s \in S} \theta(s) \cdot log \frac{\theta(s)}{\theta'(s)}$$

E. CONDITIONAL ENTROPY

In 2002, Guan *et al.* [172] proposed using the conditional entropy to measure the anonymity of an anonymous system. They show how to calculate the degree of anonymity respectively under the condition of fixed-length strategy and variable-length strategy and prove that variable path length strategies usually perform better than fixed path length strategies. However, when the expected path length is sufficiently large, the difference of anonymity degrees is relatively small.

In their work, they describe how to select routes to maximize the ability of the anonymous communication systems to protect anonymity and design an optimal route selection strategy that maximizes the anonymity degree of a system, assuming a passive adversary who can compromise one or more nodes in an anonymous system. In addition, they show that the path selection problem can be cast as an optimization problem, whose solution yields an optimal path length distribution that maximizes the anonymity degree.

In 2008, Moe in their work [173] for quantifying the anonymity for mobile ad hoc networks, proposed that information theoretical entropy can be used for quantification of the anonymity offered by a routing protocol as the adversary captures an increasing number of nodes in the network. The proposed measurement metric based on conditional entropy is applied to ANODR and ARM routing protocols.

They argue that Entropy may be used as a measure of how evenly the probabilities are distributed within each distribution, but two distributions with the same entropy could still have very different qualitative anonymity.

F. RÉNYI ENTROPY

In 1961, Rényi [174] proposed a more comprehensive entropy with the parameter α,

$$H_\alpha(P) = \frac{1}{1-\alpha} \log_2 \sum_{(X)} p_i^\alpha.$$

When $\alpha = 0$, it is very similar to Shannon's entropy.

In 2006, Clauß and Schiffner [175] used Rényi entropy to make a generalization of Shannon, min- and max-entropy. They present different models for anonymity metrics on the network layer and on the application layer, and propose a way to merge these models into a combined model aiming at providing metrics usable within a user-centric identity management system.

In their work, they show that the larger α in the formula, the closer Rényi entropy approaches to Min-Entropy. On the other hand, the closer α is to zero the closer H_α approaches to Max-Entropy. Furthermore, if α approaches one, Rényi entropy converges to Shannon-Entropy. They then use Rényi entropy as a framework to define anonymity metrics appropriate for different situations, with the generalization of Shannon-, Min-and Max-Entropy. They distinguish the user's and the service provider's point of view using the notions of local and global anonymity. Based on the comparisons of the anonymity metrics, we note that due to the fact that entropies always depend on the whole source, it seems that Rényi entropy is more adequate for global metrics, while quantiles depend only on a single entity and thus fit better for local metrics.

G. SUMMARY ON ANONYMITY MEASURING METRICS

In the above two sections, we have reviewed the main metrics for measuring the anonymity in anonymous communication systems, based on various theories and methodologies. In Table 2, we summarize the year, threat models, anonymous systems and protocols applied to, main features and contributions of all above anonymity metrics for measuring anonymity in anonymous communication systems.

We can see that the widely used measuring metric based on set theory, k-anonymity, combinatorial maths, and information entropy have been extensively studied and verified. They also form the basic building blocks for measuring anonymity in wireless networks. Also, the three metrics of bayesian inference, k-anonymity, and evidence theory anonymity will likely be the foundational anonymity metrics for today's and tomorrow's wireless anonymous networks.

Vulnerabilities and defenses are important features in the design and implementation of anonymous networking protocols, thus they also should be considered into the measuring metrics for anonymities. In addition, linking the anonymity metrics to specific existing and emerging anonymous systems can help to improve the ability to measure anonymity and quantitatively compare anonymity properties in various anonymous systems.

In addition, the information-theoretic metrics provide a practical and relatively lightweight approach to measure the level of anonymity that anonymizing systems provide in different environments and under different constraints, but they cannot be used to specify an anonymizing system or prove (predict) that it provides any anonymity property.

The information theory based on metrics could face difficulties in applications. For instance, the information entropy is inherently an average measure and thus not good at providing intuition about the worst case, and the information entropy is rather uninteresting to reason about the existence or absence of single interactions between a target actor and actions. So, it is imperative for a measure metric to provide good intuitions about the security of repeated uses of the channel.

Since metrics based on information theory are easier to compute than other metrics and also could compute the anonymity level ignoring other affections such as system size and so on. However, information-theoretic metrics also have some shortcomings such as the individual nodes' probabilities to be sender or receiver of some message, which information-theoretic metrics depend on, are hard to compute and also information-theoretic metrics could only compute the anonymity level at given point of time and could not predict the future anonymity level.

Measuring metrics based on information theory is widely studied and has a promising prospect for measuring

TABLE 2. Summary and comparisons of different anonymity measuring metrics.

Metrics	Year	Threat model						Application	Comments
		P	A	G	L	I	E		
Set Theory	1988 [6]	√		√		√	√	DC-Net	Anonymity set size
	1998 [34]	√		√		√		Stop-and-Go MIX	Anonymity set size
	2015 [109]	√	√	√		√			set-theoretic conditional anonymity
Informal Continuum	1998 [68]	√			√	√			Degrees of anonymity
Evidence Theory	2006 [113]	√	√						For wireless communications
	2014 [116]	√							For wireless network
Combinatorial Maths	2007 [117]	√			√			Threshold/Timed/Pool Mixes;	Permanent-based anonymity metric
	2016 [121]	√			√				combinatorial anonymity measure
K-anonymity	2003 [9]	√		√		√			k-anonymous message transmission
	2014 [127]	√							Symmetric disclosure
	2016 [128]		√						track and revoke specific malicious users
	2017 [129]	√							a clustering algorithm aiming at eliminating outliers
	2018 [130]					√			a robust location privacy preserving algorithm
	2019 [131]	√							cache and spatial k-anonymity
Probability theory	2004 [132]	√			√	√			for rerouting-based system
	2010 [134]	√						DC-Net	Define anonymity automata
Model-Checking	2004 [136]	√			√	√		Crowds	Use model checker PRISM
Bayesian inference	2009 [139]	√						Mix-Net	Bayesian Traffic Analysis

TABLE 2. ***(Continued.)*** **Summary and comparisons of different anonymity measuring metrics.**

Game-theory	2012 [141]	√	√			√	√	NO	Maximizing anonymity
	2017 [143]	√	√			√	√		Maximizing anonymity
Attack-based	2003 [144]	√						simulation	Quantifying the number of observations for an attack
	2001 [79]		√				√	Mix-Net	Estimate attacks' impact on the anonymity
	2004 [145]	√				√		Crowds, Onion Routing, Mix-Net, DC-Net	Define the predecessor attack
	2008 [146]	√	√			√		AP3	Analyze information leaks
	2011 [147]	√						Tor	Measuring path compromise in Tor
	2013 [74]		√				√		Framework for privacy-preserving systems
	2016 [148]		√					Tor	Quantifying the anonymity of Tor
	2017 [149]		√					DC-nets	against traffic-analysis attacks
Information theory	2003 [152]	√	√	√	√	√	√	Mix based email, Crowds, Onion Routing	Information Entropy
	2003 [154]	√	√	√	√	√	√	Mix-Net, the Pool Mix	Information Entropy
	2003 [156]	√		√					Information Entropy
	2014 [162]	√		√					Information Entropy
	2016 [163]	√		√				Tor	Information Entropy
	2017 [164]	√		√					Information Entropy

TABLE 2. *(Continued.)* **Summary and comparisons of different anonymity measuring metrics.**

	2017 [165]	√		√					Information Entropy
	2015 [157]	√		√				MIX-Net	Information Entropy, Source anonymity
	2005 [167]	√							MIX/MAX entropy
	2006 [168]	√		√			√	free-route/stratified network	relationship anonymity; min-entropy
	2014 [75]	√					√	Mix-Net	Min-entropy
	2005 [169]	√	√	√	√	√	√	Mix-Net	Mutual information
	2010 [170]	√						Mixminion	Mutual information
	2007 [171]	√						DC-Net	Relative entropy
	2002 [171]	√				√			Conditional Entropy
	2008 [173]	√		√	√		√	ANODR, ARM	Conditional Entropy
	2006 [175]	√		√	√				Rényi Entropy

anonymity in the future. However, they also have restraints as we mentioned above and need to be further tested and applied to real-world anonymous communication systems like Tor and I2P. Anonymity measuring metrics based on attacks is another promising direction for future research. The anonymity level of an anonymous communication system, to some extent, depends on the ability to defend against passive and active attacks. A system has higher anonymity, if the attacker needs more time and resources to comprise it. On the other hand, if an anonymous communication system is easy to compromise, the anonymity of the system is considered to be weak.

V. CONCLUSION AND FUTURE WORK

Several surveys were published in the field of anonymity and privacy protection. For example, in the field of location privacy, Bugra Gedik *et al.* [176] first raise the study for protecting location privacy in mobile systems. They provide a personalized k-anonymity model which consists of a unified privacy personalization framework to support k-anonymity and an efficient message perturbation engine. Leping Huang *et al.* [177] provided a formal model for wireless location privacy protection and tried to utilize former MIX research results to define anonymity and evaluate privacy level of location privacy protection protocols.

Later, John Krumm's survey paper [178] in 2009 concentrated on computational location privacy. This survey reviews the existing computation-based privacy mechanisms which treat location data as geometric information and studies of people's attitude about location privacy, and focuses on the computational threats to location privacy. In the survey, anonymity is discussed as one of the countermeasures against computational threats. However, they don't discuss the quantity and metric of anonymity.

George Danezis *et al.* [179] provided an overview of anonymous communication systems. In their paper, all

anonymous communication systems since the establishment in 1981 were categorized according to their underlying principles: trust and semi-trust relays, and MIX-based systems. It is similar with the first part of our paper, but we study the different kinds of anonymous communication systems more systematically and elaborately. Anonymous measuring metrics were mentioned in the introduction part of their paper, however, thorough analysis and review of anonymity measuring metrics is needed.

None of the above research papers could provide systematic overview or analysis on the formalization and measuring metrics of anonymous communication systems. A comprehensive and systematic research on measuring anonymity of different anonymous communication systems are still lacking and our paper aims to fill this gap.

This article provides the first comprehensive and systematic survey of state-of-art research on measuring anonymity in anonymous communication systems. The formalizations of the notion of anonymity are introduced from the aspects of process calculi, epistemic logic, function views, UC framework, differential privacy, probabilistic automata, and I/O automata. The main metrics for measuring anonymity based on various theories and approaches are reviewed and elaborated. However, it is worth mentioning that so far there is no a measure metric yet that is practical and precise enough for the anonymity measurement in anonymous communication systems and can also reflect the real system in the meantime.

From our survey, we can see that the formalization of the definition of anonymity is thoroughly studied, and that some areas such as measuring metrics based on evidence theory, combinatorial mathematics, model-checking and information theory are still in the primary stage and have wide prospects for the future research. The summary and comparisons in different areas in measuring anonymity, provided in the paper, can be used for further investigation.

In addition, anonymous censorship circumvention systems and techniques like SkyMorph, Freewave and so on have been advanced rapidly recently, however, few research works are focusing on measuring the anonymity of such systems. More research efforts need to be made in the future.

As we mentioned in section 2, the current formalizations of anonymity in anonymous communication systems are mainly applied and verified at simple protocols or systems such as DC-net. Though the formalizations from the views of differential privacy and UC framework are applied to the popular anonymous system Tor, they don't consider the adaptively corrupting adversaries and active attacks on Tor such as selective denial-of-service attacks. Therefore, the verifications of these formalizations on more complex protocols or systems such as Tor, Mix-Net and the DISSENT system are needed in the future. In particular, since till now the verification still stays in the abstraction and theoretical stage, the methodologies of the UC framework and differential privacy can be further studied in anonymity measuring. It is expected that the UC framework and differential privacy should be applied and verified in the real Tor system combining with its hidden services and the recently proposed congestion-aware path selection algorithm.

REFERENCES

[1] J. Ren and J. Wu, "Survey on anonymous communications in computer networks," *Comput. Commun.*, vol. 33, no. 4, pp. 420–431, 2010.
[2] D. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Commun. ACM*, vol. 24, no. 2, pp. 84–90, 1981.
[3] J. Boyan, "The anonymizer: Protecting user privacy on the Web," *Comput.-Mediated Commun. Mag.*, vol. 4, no. 9, 1997.
[4] E. Gabber, P. B. Gibbons, D. M. Kristol, Y. Matias, and A. Mayer, "Consistent, yet anonymous, Web access with LPWA," *Commun. ACM*, vol. 42, no. 2, pp. 42–47, 1999.
[5] R. Sherwood, B. Bhattacharjee, and A. Stinivasan, "P5: A protocol for scalable anonymous communication," in *Proc. IEEE Symp. Secur. Privacy*, Dec. 2002, pp. 58–70.
[6] D. Chaum, "The dining cryptographers problem: Unconditional sender and recipient untraceability," *J. Cryptol.*, vol. 1, no. 1, pp. 65–75, 1988.
[7] N. Weiler, "Secure anonymous group infrastructure for common and future Internet applications," In *Proc. 17th Annu. Comput. Secur. Appl. Conf. (ACSAC)*, Dec. 2001, pp. 401–410.
[8] S. Goel, M. Robson, M. Polte, and E. Sirer, "Herbivore: A scalable and efficient protocol for anonymous communication," Cornell Univ., Ithaca, NY, USA, Tech. Rep. 1890, 2003.
[9] L. van Ahn, A. Bortz, and N. J. Hopper, "K-anonymous message transmission," in *Proc. 10th ACM Conf. Comput. Commun. Secur.*, 2003, pp. 122–130.
[10] G. Perng, M. K. Reiter, and C. Wang, "M2: Multicasting mixes for efficient and anonymous communication," in *Proc. 26th IEEE Int. Conf. Distrib. Comput. Syst.*, Jul. 2006, p. 59.
[11] L. Xiao, X. Liu, W. Gu, D. Xuan, and Y. Liu, "A design of overlay anonymous multicast protocol," in *Proc. 20th IEEE Int. Symp. Parallel Distrib. Process. (IPDPS)*, Apr. 2006, p. 10.
[12] H. Corrigan-Gibbs and B. Ford, "Dissent: Accountable group anonymity," in *Proc. 17th ACM Conf. Comput. Commun. Secur.*, 2010, pp. 340–350.
[13] E. Syta, H. Corrigan-Gibbs, S.-C. Weng, D. Wolinsky, B. Ford, and A. Johnson, "Security analysis of accountable anonymity in dissent," *ACM Trans. Inf. Syst. Secur.*, vol. 17, no. 1, 2014, Art. no. 4.
[14] P. Kotzanikolaou, G. Chatzisofroniou, and M. Burmester, "Broadcast anonymous routing (BAR): Scalable real-time anonymous communication," *Int. J. Inf. Secur.*, vol. 16, no. 3, pp. 313–326, 2017.
[15] J. Kong and X. Hong, "ANODR: Anonymous on demand routing with untraceable routes for mobile ad-hoc networks," in *Proc. 4th ACM Int. Symp. Mobile Ad-hoc Netw. Comput. (MobiHoc)*, Annapolis, MD, USA, 2003, pp. 291–302.
[16] B. Zhu, Z. Wna, M. S. Kankanhalli, F. Bao, and R. H. Deng, "Anonymous secure routing in mobile ad-hoc networks," in *Proc. 29th Annu. IEEE Int. Conf. Local Comput. Netw.*, Nov. 2004, pp. 102–108.
[17] X. Wu and B. Bhargava, "AO2P: Ad hoc on-demand position-based private routing protocol," *IEEE Trans. Mobile Comput.*, vol. 4, no. 4, pp. 335–348, Jul. 2005.
[18] Y. Zhang, W. Liu, W. Lou, and Y. Fang, "MASK: Anonymous on-demand routing in mobile ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 9, pp. 2376–2385, Sep. 2006.
[19] S. Seys and B. Preneel, "ARM: Anonymous routing protocol for mobile ad hoc networks," in *Proc. 20th IEEE Int. Conf. Adv. Inf. Netw. Appl. (AINA)*, Apr. 2006, pp. 133–137.
[20] D. Sy, R. Chen, and L. Bao, "ODAR: On-demand anonymous routing in ad hoc networks," in *Proc. IEEE Int. Conf. Mobile Adhoc Sensor Syst. (MASS)*, Oct. 2006, pp. 267–276.
[21] J. Liu, X. Hong, J. Kong, Q. Zheng, N. Hu, and P. G. Bradford, "A hierarchical anonymous routing scheme for mobile ad-hoc networks," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Oct. 2006, pp. 1–7.
[22] J.-C. Kao and R. Marculescu, "Real-time anonymous routing for mobile ad hoc networks," in *Proc. IEEE Int. Conf. Local Comput. Netw.*, Mar. 2007, pp. 4139–4144.
[23] J. Han and Y. Liu, "Mutual anonymity for mobile P2P systems," *IEEE Trans. Parallel Distrib. Syst.*, vol. 19, no. 8, pp. 1009–1019, Aug. 2008.
[24] Z. Wan, K. Ren, and M. Gu, "USOR: An unobservable secure on-demand routing protocol for mobile ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 5, pp. 1922–1932, May 2012.

[25] M. Conti, J. Willemsen, and B. Crispo, "Providing source location privacy in wireless sensor networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1238–1280, 3rd Quart., 2013.
[26] B. Alomair, A. Clark, J. Cuellar, and R. Poovendran, "Toward a statistical framework for source anonymity in sensor networks," *IEEE Trans. Mobile Comput.*, vol. 12, no. 2, pp. 248–260, Feb. 2013.
[27] H. Artail and N. Abbani, "A pseudonym management system to achieve anonymity in vehicular ad hoc networks," *IEEE Trans. Depend. Sec. Comput.*, vol. 13, no. 1, pp. 106–119, Jan./Feb. 2016.
[28] P. Gope, J. Lee, and T. Q. S. Quek, "Resilience of dos attacks in designing anonymous user authentication protocol for wireless sensor networks," *IEEE Sensors J.*, vol. 17, no. 2, pp. 498–503, Jan. 2017.
[29] S. Ji, Z. Gui, T. Zhou, H. Yan, and J. Shen, "An efficient and certificateless conditional privacy-preserving authentication scheme for wireless body area networks big data services," *IEEE Access*, vol. 6, pp. 69603–69611, 2018.
[30] R. Amin, S. K. H. Islam, N. Kumar, and K. K. R. Choo, "An untraceable and anonymous password authentication protocol for heterogeneous wireless sensor networks," *J. Netw. Comput. Appl.*, vol. 104, pp. 133–144, Feb. 2018.
[31] F. Li, Y. Han, and C. Jin, "Cost-effective and anonymous access control for wireless body area networks," *IEEE Syst. J.*, vol. 12, no. 1, pp. 747–758, Mar. 2018.
[32] J. Jiang, G. Han, H. Wang, and M. Guizani, "A survey on location privacy protection in wireless sensor networks," *J. Netw. Comput. Appl.*, vol. 125, pp. 93–114, Jan. 2019.
[33] C. Gulcu and G. Tsudik, "Mixing E-mail with Babel," in *Proc. Int. Soc. Symp. Netw. Distrib. Syst. Secur.*, Feb. 1996, pp. 2–16.
[34] D. Kesdogan, J. Egner, and R. Büschkes, "Stop- and- Go-MIXes providing probabilistic anonymity in an open system," in *Proc. 2nd Int. Workshop Inf. Hiding*, in Lecture Notes in Computer Science, vol. 1525, 1998, pp. 83–98.
[35] I. Clarke, O. Sandberg, B. Wiley, and T. W. Hong, "Freenet: A distributed anonymous information storage and retrieval system," in *Proc. Int. Workshop Designing Privacy Enhancing Technol., Design Issues Anonymity Unobservability*, in Lecture Notes in Computer Science, vol. 2009, 2000, pp. 46–66.
[36] U. Müller, L. Cottrell, P. Palfrader, and L. Sassaman, "Mixmaster protocol-version 2," *Draft*, Jul. 2003.
[37] G. Danezis, R. Dingledine, and N. Mathewson, "Mixminion: Design of a type III anonymous remailer protocol," in *Proc. IEEE Symp. Secur. Privacy*, May 2003, pp. 2–15.
[38] S. Prusty, M. Liberatore, and B. N. Levine, "Forensic investigation of the oneswarm anonymous filesharing system," in *Proc. 18th ACM Conf. Comput. Commun. Secur. (CCS)*, 2011, pp. 201–214.
[39] M. Sherr, A. Mao, W. R. Marczak, W. Zhou, B. T. Loo, and M. A. Blaze, "A3: An extensible platform for application-aware anonymity," in *Proc. Netw. Distrib. Syst. Secur. Symp. (NDSS)*, San Diego, CA, USA, 2010.
[40] M. G. Reed, P. F. Syverson, and D. M. Goldschlag, "Anonymous connections and onion routing," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 4, pp. 482–494, May 1998.
[41] A. Jerichow, J. Müller, A. Pfitzmann, B. Pfitzmann, and M. Waidner, "Real-time mixes: A bandwidth-efficient anonymity protocol," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 4, pp. 495–509, May 1998.
[42] O. Berthold, H. Federrath, and S. Köpsell, "Web MIXes: A system for anonymous and unobservable Internet access," in *Proc. Int. Workshop Designing Privacy Enhancing Technol., Design Issues Anonymity Unobservability*, in Lecture Notes in Computer Science, vol. 2009, 2000, pp. 115–129.
[43] P. Boucher, A. Shostack, and I. Goldberg, "Freedom systems 2.0 architecture," Zero Knowledge Systems, Montreal, QC, Canada, White Paper, Dec. 2000.
[44] M. Rennhard and B. Plattner, "Introducing MorphMix: Peer-to-peer based anonymous Internet usage with collusion detection," in *Proc. Workshop Privacy Electron. Soc. (WPES)*, 2002, pp. 91–102.
[45] M. J. Freedman and R. Morris, "Tarzan: A peer-to-peer anonymizing network layer," in *Proc. 9th ACM Conf. Comput. Commun. Secur. (CCS)*, 2002, pp. 193–206.
[46] C. Egger, J. Schlumberger, C. Kruegel, and G. Vigna, "Practical attacks against the I2P network," in *Proc. 16th Int. Symp. Res. Attacks, Intrusions Defenses (RAID)*, in Lecture Notes in Computer Science, vol. 8145, 2013, pp. 432–451.
[47] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second-generation onion router," in *Proc. 13th USENIX Secur. Symp.*, 2004, p. 21.
[48] K. Bauer, M. Sherr, D. McCoy, and D. Grunwald, "ExperimenTor: A testbed for safe and realistic tor experimentation," in *Proc. USENIX Workshop Cyber Secur. Experimentation Test (CSET)*, Aug. 2011, p. 7.
[49] R. Jansen and N. J. Hopper, "Shadow: Running Tor in a box for accurate and efficient experimentation," in *Proc. Netw. Distrib. Syst. Secur. Symp.*, Feb. 2012.
[50] Z. Weinberg, J. Wang, V. Yegneswaran, L. Briesemeister, S. Cheung, F. Wang, and D. Boneh, "StegoTorus: A camouflage proxy for the Tor anonymity system," in *Proc. ACM Conf. Comput. Commun. Secur.*, 2012, pp. 109–120.
[51] H. M. Moghaddam, B. Li, M. Derakhshani, and I. Goldberg, "SkypeMorph: Protocol obfuscation for Tor bridges," in *Proc. ACM Conf. Comput. Commun. Secur.*, 2012, pp. 97–108.
[52] M. AlSabah and I. Goldberg, "PCTCP: Per-circuit TCP-over-IPsec transport for anonymous communication overlay networks," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2013, pp. 349–360.
[53] A. Houmansadr, G. T. K. Nguyen, M. Caesar, and N. Borisov, "Cirripede: Circumvention infrastructure using router redirection with plausible deniability," in *Proc. 18th ACM Conf. Comput. Commun. Secur.*, 2011, pp. 187–200.
[54] E. Wustrow, M. C. Swanson, and J. A. Halderman, "TapDance: End-to-middle anticensorship without flow blocking," in *Proc. 23rd USENIX Secur. Symp.*, 2014, pp. 159–174.
[55] E. Wustrow, S. Wolchok, I. Goldberg, and J. A. Halderman, "Telex: Anticensorship in the network infrastructure," in *Proc. 20th USENIX Secur. Symp.*, San Francisco, CA, USA, Aug. 2011, p. 30.
[56] Q. Wang, X. Gong, G. T. K. Nguyen, A. Houmansadr, and N. Borisov, "CensorSpoofer: Asymmetric communication with IP spoofing for censorship-resistant Web browsing," 2012, *arXiv:1203.1673*. [Online]. Available: https://arxiv.org/abs/1203.1673
[57] J. Karlin, D. Ellard, A. W. Jackson, C. E. Jones, G. Lauer, D. Mankins, and W. T. Strayer, "Decoy routing: Toward unblockable Internet communication," in *Proc. USENIX Workshop Free Open Commun. Internet*, San Francisco,CA, Aug. 2011.
[58] M. Backes, A. Kate, and E. Mohammadi, "ACE: An efficient key-exchange protocol for onion routing," in *Proc. ACM Workshop Privacy Electron. Soc.*, 2012, pp. 55–64.
[59] C. Chen, D. E. Asoni, D. Barrera, G. Danezis, and A. Perrig, "HORNET: High-speed onion routing at the network layer," in *Proc. 22nd ACM SIGSAC Conf. Comput. Commun. Secur.*, 2015, pp. 1441–1454.
[60] T. Ruffing, P. Moreno-Sanchez, and A. Kate, "P2P mixing and unlinkable Bitcoin transactions," in *Proc. NDSS*, 2017, pp. 1–15.
[61] A. M. Piotrowska, J. Hayes, T. Elahi, S. Meiser, and G. Danezis, "The Loopix anonymity system," in *Proc. 26th USENIX Secur. Symp. (USENIX Secur.)*, 2017, pp. 1199–1216.
[62] C. Park, K. Itoh, and K. Kurosawa, "Efficient anonymous channel and all/nothing election scheme," in *Proc. Adv. Cryptol. (EUROCRYPT)*, in Lecture Notes in Computer Science, vol. 765. Berlin, Germany: Springer-Verlag, 1993, pp. 248–259.
[63] M. Gomulkiewicz, M. Klonowski, and M. Kutylowski, "Onions based on universal re-encryption—Anonymous communication immune against repetitive attack," in *Proc. Workshop Inf. Secur. Appl. (WISA)*, in Lecture Notes in Computer Science, vol. 3325. Berlin, Germany: Springer, 2004, pp. 400–410.
[64] P. Golle, M. Jakobsson, A. Juels, and P. Syverson, "Universal re-encryption for mixnets," in *Proc. RSA-Conf., Cryptogr. Track*, in Lecture Notes in Computer Science, vol. 2964. Berlin, Germany: Springer, 2004, pp. 163–178.
[65] T. Lu, B. Fang, Y. Sun, and L. Guo, "Some remarks on universal re-encryption and a novel practical anonymous tunnel," in *Networking and Mobile Computing* (Lecture Notes in Computer Science), vol. 3619. Berlin, Germany: Springer, 2005, pp. 853–862.
[66] D. Chaum, D. Das, A. Kate, A. Krasnova, J. De Ruiter, A. T. Sherman, and F. Javani, "cMix: Mixing with minimal real-time asymmetric cryptographic operations," in *Proc. Int. Conf. Appl. Cryptogr. Netw. Secur.* Cham, Switzerland: Springer, 2017, pp. 557–578.
[67] O. Pereira and R. L. Rivest, "Marked mix-nets," in *Proc. Int. Conf. Financial Cryptogr. Data Secur.* Cham, Switzerland: Springer, 2017, pp. 353–369.
[68] M. K. Reiter and A. D. Rubin, "Crowds: Anonymity for Web transactions," *ACM Trans. Inf. Syst. Secur.*, vol. 1, no. 1, pp. 62–92, 1998.
[69] C. Shields and B. N. Levine, "A protocol for anonymous communication over the Internet," in *Proc. 7th ACM Conf. Comput. Commun. Secur.*, 2000, pp. 33–42.

[70] A. Mislove, G. Oberoi, A. Post, C. Reis, P. Druschel, and D. S. Wallach, "AP3: Cooperative, decentralized anonymous communication," in *Proc. 11th ACM SIGOPS Eur. Workshop*, Leuven, Belgium, 2004, p. 30.

[71] T. Lu, B. Fang, X. Cheng, and Y. Sun, "WonGoo: A peer-to-peer protocol for anonymous communication," in *Proc. Int. Conf. Parallel Distrib. Process. Techn. Appl.*, 2004, pp. 1102–1106.

[72] W. H. Tang and H. W. Chan, "MIX-crowds, an anonymity scheme for file retrieval systems," in *Proc. IEEE INFOCOM*, Apr. 2009, pp. 1170–1178.

[73] J. Daubert, M. Fischer, T. Grube, S. Schiffner, P. Kikiras, and M. Mühlhäuser, "AnonPubSub: Anonymous publish-subscribe overlays," *Comput. Commun.*, vol. 76, pp. 42–53, Feb. 2016.

[74] D. R. Monedero, J. P. Arnau, C. Diaz, and J. Forné, "On the measurement of privacy as an attacker's estimation error," *Int. J. Inf. Secur.*, vol. 12, no. 2, pp. 129–149, 2013.

[75] D. Rebollo-Monedero, J. Parra-Arnau, J. Forné, and C. Diaz, "Optimizing the design parameters of threshold pool mixes for anonymity and delay," *Comput. Netw.*, vol. 67, pp. 180–200, Jul. 2014.

[76] G. Danezis, "Measuring anonymity: A few thoughts and a differentially private bound," in *Proc. DIMACS Workshop Measuring Anonymity* May 2013, p. 26.

[77] M. R. Koot, "Measuring and predicting anonymity," Ph.D. dissertation, Univ. Amsterdam, Amsterdam, The Netherlands, 2012.

[78] D. Kelly, R. Raines, R. Baldwin, M. Grimaila, and B. Mullins, "Exploring extant and emerging issues in anonymous networks: A taxonomy and survey of protocols and metrics," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 2, pp. 579–606, 2nd Quart., 2012.

[79] O. Berthold, A. Pfitzmann, and R. Standtke, "The disadvantages of free MIX routes and how to overcome them," in *Proc. Designing Privacy Enhancing Technol.*, in Lecture Notes in Computer Science, vol. 2009. Berlin, Germany: Springer, 2001, pp. 30–45.

[80] A. Pfitzmann and M. Hansen. (Aug. 2010). *A Terminology for Talking About Privacy by Data Minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management, (Version v0.34*. [Online]. Available: https://dud.inf.tudresden.de/literatur/AnonTerminologyv0.34.pdf

[81] S. Schneider and A. Sidiropoulos, "CSP and anonymity," in *Proc. 4th Eur. Symp. Res. Comput. Secur. (ESORICS)*, in Lecture Notes in Computer Science, vol. 1146. Berlin, Germany: Springer, 1996, pp. 198–218.

[82] P. Ryan and S. Schneider, *Modeling and Analysis of Security Protocols*. Reading, MA, USA: Addison-Wesley, 2001.

[83] M. Bhargava and C. Palamidessi, "Probabilistic anonymity," in *CONCUR 2005—Concurrency Theory* (Lecture Notes in Computer Science), vol. 3653. Berlin, Germany: Springer, 2005, pp. 171–185.

[84] S. Mauw, J. H. S. Verschuren, and E. P. de Vink, "A formalization of anonymity and onion routing," in *Proc. 9th Eur. Symp. Res. Comput. Secur. (ESORICS)*, in Lecture Notes in Computer Science, vol. 3193, 2004, pp. 109–124.

[85] T. Chothia, "Analysing the MUTE anonymous file-sharing system using the pi-calculus," in *Proc. Formal Techn. Networked Distrib. Syst. (FORTE)*, in Lecture Notes in Computer Science, vol. 4229, 2006, pp. 115–130.

[86] T. Chothia, S. Orzan, J. Pang, and M. T. Dashti, "A framework for automatically checking anonymity with μCRL," in *Proc. 2nd Int. Conf. Trustworthy Global Comput. (TGC)*, in Lecture Notes in Computer Science, vol. 4661, 2007, pp. 301–318.

[87] M. Moran, J. Heather, and S. Schneider, "Verifying anonymity in voting systems using CSP," *Formal Aspects Comput.*, vol. 26, no. 1, pp. 63–98, 2012.

[88] M. Moran and J. Heather, "Automated analysis of voting systems under an active intruder model in CSP," 2017, *arXiv:1705.00795*. [Online]. Available: https://arxiv.org/abs/1705.00795

[89] J. van Eijck and S. Orzan, "Epistemic verification of anonymity," *Electron. Notes Theor. Comput. Sci.*, vol. 168, pp. 159–174, Feb. 2007.

[90] P. F. Syverson and S. G. Stubblebine, "Group principals and the formalization of anonymity," in *Proc. World Congr. Formal Methods*, in Lecture Notes in Computer Science, vol. 1708. Berlin, Germany: Springer, 1999, pp. 814–833.

[91] J. Y. Halpern and K. R. O'Neill, "Anonymity and information hiding in multiagent systems," *J. Comput. Secur.*, vol. 13, no. 3, pp. 483–514, 2005.

[92] F. D. Garcia, I. Hasuo, W. Pieters, and P. van Rossum, "Provable anonymity," in *Proc. ACM Workshop Formal Methods Secur. Eng.*, 2005, pp. 63–72.

[93] D. Hughes and V. Shmatikov, "Information hiding, anonymity and privacy: A modular approach," *J. Comput. Secur.*, vol. 12, no. 1, pp. 3–36, 2004.

[94] R. Canetti, "Universally composable security: A new paradigm for cryptographic protocols," in *Proc. 42nd IEEE Symp. Found. Comput. Sci.*, Oct. 2001, pp. 136–145.

[95] J. Feigenbaum, A. Johnson, and P. Syverson, "Probabilistic analysis of onion routing in a black-box model," *ACM Trans. Inf. Syst. Secur.*, vol. 15, no. 3, 2012, Art. no. 14.

[96] M. Backes, I. Goldberg, A. Kate, and E. Mohammadi, "Provably secure and practical onion routing," in *Proc. IACR Conf. Workshops, Cryptol. ePrint Arch.*, 2012.

[97] M. Backes, P. Manoharan, and E. Mohammadi, "TUC: Time-sensitive and modular analysis of anonymous communication," in *Proc. IEEE 27th Comput. Secur. Found. Symp. (CSF)*, Jul. 2014, pp. 383–397.

[98] F. Shirazi, M. Simeonovski, M. R. Asghar, M. Backes, and C. Diaz, "A survey on routing in anonymous communication protocols," *ACM Comput. Surv.*, vol. 51, no. 3, 2018, Art. no. 51.

[99] M. Backes, A. Herzberg, A. Kate, and I. Pryvalov, "Anonymous RAM," in *Proc. Eur. Symp. Res. Comput. Secur.* Cham, Switzerland: Springer, Sep. 2016, pp. 344–362.

[100] M. Backes, P. Berrang, O. Goga, K. P. Gummadi, and P. Manoharan, "On profile linkability despite anonymity in social media systems," in *Proc. ACM Workshop Privacy Electron. Soc.*, Oct. 2016, pp. 25–35.

[101] C. Dwork, "Differential privacy: A survey of results," in *Proc. 5th Int. Conf. Theory Appl. Models Comput.*, vol. 4978, 2008, pp. 1–19.

[102] M. Backes, A. Kate, P. Manoharan, S. Meiser, and E. Mohammadi, "AnoA: A framework for analyzing anonymous communication protocols," in *Proc. 26th IEEE Comput. Secur. Found. Symp.*, Jun. 2013, pp. 163–178.

[103] M. Backes, A. Kate, S. Meiser, and E. Mohammadi, "(nothing else) MATor(s): Monitoring the anonymity of Tor's path selection," in *Proc. 21st Conf. Comput. Commun. Secur. (CCS)*, 2014, pp. 513–524.

[104] K. Chatzikokolakis and C. Palamidessi, "Probable innocence revisited," in *Proc. 3rd Int. Workshop Formal Aspects Secur. Trust*, in Lecture Notes in Computer Science, vol. 3866. 2006, pp. 142–157.

[105] M. E. Andrés, C. Palamidessi, A. Sokolova, and P. V. Rossum, "Information hiding in probabilistic concurrent systems," *J. Theor. Comput. Sci.*, vol. 412, no. 28, pp. 3072–3089, 2011.

[106] N. A. Lynch and M. R. Tuttle, "An introduction to input/output automata," Massachusetts Inst. Technol., Cambridge, MA, USA, Tech. Rep. 02139, 1988.

[107] J. Feigenbaum, A. Johnson, and P. F. Syverson, "A model of onion routing with provable anonymity," in *Proc. 11th Int. Conf. Financial Cryptogr. Data Secur.*, in Lecture Notes in Computer Science, vol. 4886. 2007, pp. 57–71.

[108] S. J. Murdoch, "Quantifying and measuring anonymity," in *Proc. Data Privacy Manage. Auto. Spontaneous Secur.*, in Lecture Notes in Computer Science, vol. 8247. 2014, pp. 3–13.

[109] W. Chen, Y. Cao, and H. Wang, "Conditional anonymity with non-probabilistic adversary," *Inf. Sci.*, vol. 324, pp. 32–43, Dec. 2015.

[110] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *Ann. Math. Statist.*, vol. 38, no. 2, pp. 325–339, 1967.

[111] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ, USA: Princeton Univ. Press, 1976, pp. 10–28.

[112] J. Kohlas and P. A. Monney, *A Mathematical Theory of Hints: An Approach to the Dempster-Shafer Theory of Evidence* (Lecture Notes in Economics and Mathematical Systems), vol. 425. New York, NY, USA: Springer-Verlag, 1995.

[113] H. Dijiang, "On measuring anonymity for wireless mobile ad-hoc networks," in *Proc. 31st IEEE Conf. Local Comput. Netw.*, Nov. 2006, pp. 779–786.

[114] H. Dijiang, "Unlinkability measure for IEEE 802.11 based MANETs," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 1025–1034, Mar. 2008.

[115] B. Li and D. Huang, "Modeling anonymous MANET communications using super-nodes," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Nov. 2013, pp. 125–130.

[116] B. Li, D. Huang, and Z. Wang, "Refining traffic information for analysis using evidence theory," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Oct. 2014, pp. 1181–1186.

[117] M. Edman, F. Sivrikaya, and B. Yener, "A combinatorial approach to measuring anonymity," in *Proc. IEEE Intell. Secur. Inform.*, May 2007, pp. 356–363.

[118] B. Gierlichs, C. Troncoso, C. Diaz, and B. Preneel, "Revisiting a combinatorial approach toward measuring anonymity," in *Proc. 7th ACM Workshop Privacy Electron. Soc.*, 2008, pp. 111–116.
[119] J. C. Grégoire and A. M. Hamel, "A combinatorial enumeration approach for measuring anonymity," 2009, *arXiv:0902.1663*. [Online]. Available: https://arxiv.org/abs/0902.1663
[120] R. Bagai, H. Lu, R. Li, and B. Tang, "An accurate system-wide anonymity metric for probabilistic attacks," in *Proc. Int. Symp. Privacy Enhancing Technol.*, in Lecture Notes in Computer Science, vol. 6794. 2011, pp. 117–133.
[121] A. Iwai, "Reviewing privacy-enhanced social survey system that employs combinatorial anonymity measure," in *Proc. Int. Multi Conf. Eng. Comput. Scientists*, vol. 1, 2016, pp. 1–6.
[122] L. Sweeney, "K-anonymity: A model for protecting privacy," *Int. J. Uncertainty, Fuzziness Knowl.-Based Syst.*, vol. 10, no. 5, pp. 557–570, 2002.
[123] X. Wu and E. Bertino, "Achieving k-anonymity in mobile ad hoc networks," in *Proc. 1st IEEE ICNP Workshop Secure Netw. Protocols*, Nov. 2005, pp. 37–42.
[124] N. Hopper and E. Y. Vasserman, "On the effectiveness of k-anonymity against traffic analysis and surveillance," in *Proc. 5th ACM Workshop Privacy Electron. Soc.*, 2006, pp. 9–18.
[125] P. Wang, P. Ning, and D. S. Reeves, "A k-anonymous communication protocol for overlay networks," in *Proc. 2nd ACM Symp. Inf., Comput. Commun. Secur.*, 2007, pp. 45–56.
[126] R. Shokri, C. Troncoso, C. Diaz, J. Freudiger, and J. P. Hubaux, "Unraveling an old cloak: K-anonymity for location privacy," in *Proc. 9th Annu. ACM Workshop Privacy Electron. Soc.*, 2010, pp. 115–118.
[127] E. J. Infeld, "Symmetric disclosure: A fresh look at k-anonymity," in *Proc. 4th USENIX Workshop Free Open Commun. Internet*, 2014, pp. 1–8.
[128] C. Caballero-Gil, J. Molina-Gil, J. Hernández-Serrano, O. León, and M. Soriano-Ibanez, "Providing k-anonymity and revocation in ubiquitous VANETs," *Ad Hoc Netw.*, vol. 36, no. P2, pp. 482–494, 2015.
[129] L. Zheng, H. Yue, Z. Li, X. Pan, M. Wu, and F. Yang, "K-anonymity location privacy algorithm based on clustering," *IEEE Access*, vol. 6, pp. 28328–28338, 2018.
[130] H. Jiang, P. Zhao, and C. Wang, "RobLoP: Towards robust privacy preserving against location dependent attacks in continuous LBS queries," *IEEE/ACM Trans. Netw.*, vol. 26, no. 2, pp. 1018–1032, Apr. 2018.
[131] S. Zhang, X. Li, Z. Tan, T. Peng, and G. Wang, "A caching and spatial K-anonymity driven privacy enhancement scheme in continuous location-based services," *Future Gener. Comput. Syst.*, vol. 94, pp. 40–50, May 2019.
[132] Y. Guan, X. Fu, R. Bettati, and W. Zhao, "A quantitative analysis of anonymous communications," *IEEE Trans. Rel.*, vol. 53, no. 1, pp. 103–115, Mar. 2004.
[133] Y. Deng, C. Palamidessi, and J. Pang, "Weak probabilistic anonymity," *Electron. Notes Theor. Comput. Sci.*, vol. 180, no. 1, pp. 55–76, 2007.
[134] I. Hasuo and Y. Kawabe, "Probabilistic anonymity via coalgebraic simulations," *Theor. Comput. Sci.*, vol. 411, no. 22, pp. 2239–2259, 2010.
[135] V. Shmatikov, "Probabilistic analysis of anonymity," in *Proc. 15th IEEE Comput. Secur. Found. Workshop*, Jun. 2002, pp. 119–128.
[136] V. Shmatikov, "Probabilistic model checking of an anonymity system," *J. Comput. Secur.*, vol. 12, no. 3, pp. 355–377, 2004.
[137] M. T. Adithia, "Probabilistic analysis of network anonymity using PRISM," M.S. thesis, Dept. Math. Comput. Sci., Eindhoven Univ. Technol., Eindhoven, The Netherlands, 2006.
[138] K. Chatzikokolakis, C. Palamidessi, and P. Panangaden, "Probability of error in information-hiding protocols," in *Proc. 20th IEEE Comput. Secur. Found. Symp.*, Jul. 2007, pp. 341–354.
[139] C. Troncoso and G. Danezis, "The Bayesian traffic analysis of mix networks," in *Proc. 16th ACM Conf. Comput. Commun. Secur.*, 2009, pp. 369–379.
[140] C. Troncoso, "Design and analysis methods for privacy technologies," Ph.D. dissertation, Dept. Elect. Eng., Katholieke Univ. Leuven, Leuven, Belgium, 2011.
[141] P. Venkitasubramaniam and L. Tong, "A game-theoretic approach to anonymous networking," *IEEE/ACM Trans. Netw.*, vol. 20, no. 3, pp. 892–905, Jun. 2012.
[142] A. Mishra and P. Venkitasubramaniam, "Admissible length study in anonymous networking: A detection theoretic perspective," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1957–1969, Sep. 2013.
[143] R. Shokri, G. Theodorakopoulos, and C. Troncoso, "Privacy games along location traces: A game-theoretic framework for optimizing location privacy," *ACM Trans. Privacy Secur.*, vol. 19, no. 4, p. 11, 2017.
[144] D. Agrawal and D. Kesdogan, "Measuring anonymity: The disclosure attack," *IEEE Secur. Privacy*, vol. 1, no. 6, pp. 27–34, Nov./Dec. 2003.
[145] M. K. Wright, M. Adler, B. N. Levine, and C. Shields, "The predecessor attack: An analysis of a threat to anonymous communications systems," *ACM Trans. Inf. Syst. Secur.*, vol. 7, no. 4, pp. 489–522, 2004.
[146] P. Mittal and N. Borisov, "Information leaks in structured peer-to-peer anonymous communication systems," in *Proc. 15th ACM Conf. Comput. Commun. Secur.*, 2008, pp. 267–278.
[147] A. Hamel, J. C. Grégoire, and I. Goldberg, "The Mis-entropists: New approaches to measures in Tor," Cheriton School Comput. Sci., Univ. Waterloo, Waterloo, ON, USA, Tech. Rep. 2011-18, 2011.
[148] M. Backes, S. Meiser, and M. Slowik, "Your choice MATor(s)," in *Proc. Privacy Enhancing Technol.*, no. 2, 2016, pp. 40–60.
[149] L. Barman, "PriFi: A low-latency local-area anonymous communication network," 2017, *arXiv:1710.10237*. [Online]. Available: https://arxiv.org/abs/1710.10237
[150] C. E. Shannon, "The mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–390, 1948.
[151] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949.
[152] C. Díaz, S. Seys, J. Claessens, and B. Preneel, "Towards measuring anonymity," in *Proc. Int. Workshop Privacy Enhancing Technol.*, in Lecture Notes in Computer Science, vol. 2482. Berlin, Germany: Springer, Apr. 2002, pp. 54–68.
[153] C. Diaz, "Anonymity and privacy in electronic services," Faculteit Ingenieurswetenschappen, Katholieke Univ. Leuven, Heverlee, Belgium, 2005, pp. 23–40.
[154] A. Serjantov and G. Danezis, "Towards an information theoretic metric for anonymity," in *Proc. Int. Workshop Privacy Enhancing Technol.*, in Lecture Notes in Computer Science, vol. 2482. Berlin, Germany: Springer, Apr. 2002, pp. 41–53.
[155] S. Steinbrecher and S. Köpsell, "Modelling unlinkability," in *Proc. Int. Workshop Privacy Enhancing Technol.* (Lecture Notes in Computer Science), vol. 2760. Berlin, Germany: Springer, 2003, pp. 32–47.
[156] R. E. Newman, I. S. Moskowitz, P. Syverson, and A. Serjantov, "Metrics for traffic analysis prevention," in *Proc. Int. Workshop Privacy Enhancing Technol.* (Lecture Notes in Computer Science), vol. 2760. Berlin, Germany: Springer, 2003, pp. 48–65.
[157] P. Venkitasubramaniam and L. Tong, "Throughput anonymity trade-off in wireless networks under latency constraints," in *Proc. 27th IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2008, pp. 807–815.
[158] P. Venkitasubramaniam and A. Mishra, "Anonymity of memory-limited Chaum mixes under timing analysis: An information theoretic perspective," *IEEE Trans. Inf. Theory*, vol. 61, no. 2, pp. 996–1009, Feb. 2015.
[159] K. Chatzikokolakis, "Probabilistic and information-theoretic approaches to anonymity," Ph.D. dissertation, Laboratoire d'Informatique, École Polytechn., Paris, France, Oct. 2007.
[160] L. O'Connor, "Entropy bounds for traffic confirmation," IACR Cryptol. ePrint Arch., Tech. Rep. 2008/365, Oct. 2008.
[161] P. Mittal and N. Borisov, "ShadowWalker: Peer-to-peer anonymous communication using redundant structured topologies," in *Proc. 16th ACM Conf. Comput. Commun. Secur.*, 2009, pp. 161–172.
[162] B. Niu, Q. Li, X. Zhu, G. Cao, and H. Li, "Achieving k-anonymity in privacy-aware location-based services," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr./May 2014, pp. 754–762.
[163] R. Annessi and M. Schmiedecker, "NavigaTor: Finding faster paths to anonymity," in *Proc. IEEE Eur. Symp. Secur. Privacy (EuroS&P)*, Mar. 2016, pp. 214–226.
[164] F. Fei, S. Li, H. Dai, C. Hu, W. Dou, and Q. Ni, "A K-anonymity based schema for location privacy preservation," *IEEE Trans. Sustain. Comput.*, to be published.
[165] L. Tennant. *Improving the Anonymity of the IOTA Cryptocurrency*. Accessed: Oct. 2017. [Online]. Available: http://iotafeed.com/wp-content/uploads/2017/08/anonymityiota.pdf
[166] G. Tóth, Z. Hornák, and F. Vajda, "Measuring anonymity revisited," in *Proc. 9th Nordic Workshop Secure IT Syst.*, Espoo, Finland, 2004, pp. 85–90.
[167] G. Tóth and Z. Hornák, "Measuring anonymity in a non-adaptive, real-time system," in *Proc. Int. Workshop Privacy Enhancing Technol.* (Lecture Notes in Computer Science), vol. 3424. Berlin, Germany: Springer, 2004, pp. 226–241.

[168] V. Shmatikov and M.-H. Wang, "Measuring relationship anonymity in mix networks," in *Proc. 5th ACM Workshop Privacy Electron. Soc.*, 2006, pp. 59–62.
[169] Y. Zhu and R. Bettati, "Anonymity vs. information leakage in anonymity systems," in *Proc. 25th IEEE Int. Conf. Distrib. Comput. Syst.*, Jun. 2005, pp. 514–524.
[170] K. Chatzikokolakis, T. Chothia, and A. Guha, "Statistical measurement of information leakage," in *Proc. Int. Conf. Tools Algorithms Construct. Anal. Syst.* (Lecture Notes in Computer Science), vol. 6015. Berlin, Germany: Springer, 2010, pp. 390–404.
[171] Y. Deng, J. Pang, and P. Wu, "Measuring anonymity with relative entropy," in *Proc. Int. Workshop Formal Aspects Secur. Trust*, in Lecture Notes in Computer Science, vol. 4691. Berlin, Germany: Springer, 2006, pp. 65–79.
[172] Y. Guan, X. Fu, R. Bettati, and W. Zhao, "An optimal strategy for anonymous communication protocols," in *Proc. 22nd Int. Conf. Distrib. Comput. Syst.*, Jul. 2002, pp. 257–266.
[173] M. E. G. Moe, "Quantification of anonymity for mobile ad hoc networks," in *Proc. 4th Int. Workshop Secur. Trust Manage. (STM)*, 2008, pp. 95–107.
[174] A. Rényi, "On measures of entropy and information," in *Proc. 4th Berkeley Symp. Math. Statist. Probab.*, 1961, pp. 547–561.
[175] S. Clauß and S. Schiffner, "Structuring anonymity metrics," in *Proc. 2nd ACM Workshop Digit. Identity Manage.*, 2006, pp. 55–62.
[176] B. Gedik and L. Liu, "Location privacy in mobile systems: A personalized anonymization model," in *Proc. 25th IEEE Int. Conf. IEEE Distrib. Comput. Syst. (ICDCS)*, Jun. 2005, pp. 620–629.
[177] L. Huang, H. Yamane, K. Matsuura, and K. Sezaki, "Towards modeling wireless location privacy," in *Privacy Enhancing Technologies*. Berlin, Germany: Springer, 2006, pp. 59–77.
[178] J. Krumm, "A survey of computational location privacy," *Pers. Ubiquitous Comput.*, vol. 13, no. 6, pp. 391–399, 2009.
[179] G. Danezis and C. Diaz, "A survey of anonymous communication channels," Microsoft Res., Bengaluru, India, Tech. Rep. MSR-TR-2008-35, 2008.
[180] R. Jansen and A. Johnson, "Safely measuring Tor," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, Vienna, Austria, Oct. 2016, pp. 1553–1567.

TIANBO LU received the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, in 2006. He is currently an Associate Professor with the School of Software Engineering, Beijing University of Posts and Telecommunications, China. His technical interests include information hiding and watermarking, network security, anonymous communication, P2P computing, and software assurance. He received funding from the National Natural Science Foundation of China and attracted the industry to work on the above areas and to build programmable network testbeds.

ZEYU DU received the B.S. degree in hydrology and water resources engineering from the China University of Geosciences, China, in 2017. He is currently pursuing the M.S. degree with the School of Software Engineering, Beijing University of Posts And Telecommunications. His research interests include privacy enhancing technologies, cyber-physical systems security, applied cryptography, and network security.

Z. JANE WANG received the B.Sc. degree in electrical engineering from Tsinghua University, China, in 1996, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Connecticut, in 2000 and 2002, respectively. She has been a Research Associate with the Electrical and Computer Engineering Department, University of Maryland, College Park. Since 2004, she has been with the Department Electrical and Computer Engineering, The University of British Columbia, Canada, where she is currently a Professor. Her research interest includes statistical signal processing theory and applications, with focus on multimedia security and biomedical signal processing and modeling. While at the University of Connecticut, she received the Outstanding Engineering Doctoral Student Award. She co-received the *EURASIP Journal on Applied Signal Processing* (JASP) Best Paper Award, in 2004, and the IEEE Signal Processing Society Best Paper Award, in 2005. She is the Chair and Founder of the IEEE Signal Processing Chapter at Vancouver. She is serving as Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, and the IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING.