Improving Unlinkability of Attribute Based Authentication through Game Theory

YEVHEN ZOLOTAVKIN AND JAY JEONG*, VERONIKA KUCHTA, MAKSYM SLAVNENKO, and ROBIN DOSS, Deakin University

This paper first formalizes the problem of unlinkable attribute-based authentication in the system where each user possess multiple assertions and uses them interchangeably. We state that existing definition of unlinkability of the assertions that users submit to Relying Party (RP) to get authenticated is limited. This also does not allow to provide recommendations as for optimal usage of available assertions. To mitigate this issue we use conditional entropy to measure uncertainty for RP who attempts to link observed assertions to user labels. Conditional entropy is the function of usage statistics for all assertions in the system. Personal *decisions* made by the users about usage of assertions contribute to this statistics. This collective effect from all the users impacts unlinkability of authentication and must be studied using game theory. We particularize several instances of the game where context information that is provided to the users differs. Through game theory and based on conditional entropy we demonstrate how each user optimizes usage for the personal set of assertions. In the experiment we substantiate advantage of the proposed rational decision making approaches: unlinkability that we obtain under Nash equilibrium is higher than in the system where users authenticate using their assertions at random. We finally propose an algorithm that calculates equilibrium and assists users with the selection of assertions. This manifests that described techniques can be executed in realistic settings. This does not require modification of existing authentication protocols and can be implemented in platform-independent identity agents. As a use case, we describe how our technique can be used in Digital Credential Wallets (DCW): we suggest that unlinkability of authentication can be improved for Verifiable Credential (VC).

CCS Concepts: • **Security and privacy** → Security services.

Additional Key Words and Phrases: attribute based authentication, unlinkability, game theory

ACM Reference Format:
Yevhen Zolotavkin and Jay Jeong, Veronika Kuchta, Maksym Slavnenko, and Robin Doss. 2020. Improving Unlinkability of Attribute Based Authentication through Game Theory. 1, 1 (March 2020), 35 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Verifiable digital credentials play an important role in the authentication and authorization of modern users [59, 66]. Multiple initiatives have originated in the professional community with the aim to address privacy issues associated with the usage of attribute-rich credentials. Data models such as Verifiable Credentials (VC) and IRMA allow to implement principles of data minimization

*Both authors contributed equally to this research.

Authors' address: Yevhen Zolotavkin and Jay Jeong, first.last@deakin.edu.au; Veronika Kuchta, v.kuchta@uq.edu.au; Maksym Slavnenko, Maksym.Slavnenko@deakin.edu.au; Robin Doss, robin.doss@deakin.edu.au, Deakin University, Geelong, Victoria.

XXXX-XXXX/2020/3-ART $15.00
https://doi.org/10.1145/1122445.1122456

[2, 68]. In addition, recently introduced standards such as the ISO 2755x family provide definitions, frameworks and best practices capable to improve users' privacy. For instance, standard ISO 27551 "Requirements for attribute based unlinkable entity authentication" analyses a broad range of conditions that negatively affects user privacy [43]. However both ISO 27551 and VC have technological and conceptual limitations: practicality of unlinkability tests and known techniques for its improvement are unsatisfactory. The standard defines *unlinkability* which is capable to counter privacy threats under these conditions.

Definition 1 (Unlinkability [21])**.** *Within a particular set of information, the inability of an observer or attacker to distinguish whether two items of interest are related or not (with a high enough degree of probability to be useful to the observer or attacker) is called unlinkability.*

From the definition, it is tempting to assume that *indistinguishability* is sufficient for unlinkability. Unfortunately, this has the risk of being incorrect. As we demonstrate in the following example, special attention must be paid to the clauses *'within a particular set of information'* and *'with a high enough degree of probability to be useful to the observer or attacker'*.

Example #0 Consider an attribute-based authentication system where two different entities, *Alice* and *Bob* authenticate with identical assertions (i.e. some abstract proof) α to a Relying Party (RP). This warrants that indistinguishability at RP side is achieved. The role of the attacker is played by the RP who is trying to *link* the users without their consent. The 'particular set of information' for RP consists of the knowledge that a random authentication event is initiated by *Alice* with probability $\Pr(A) = 0.9$, and it is initiated by *Bob* with $\Pr(B) = 0.1$. RP observes 2 random authentication events and needs to decide whether they were initiated by the same entity. By saying that 2 events were initiated by the same entity RP is wrong with probability $1 - 0.9^2 - 0.1^2 = 0.18$. This probability of error is *low* which means that the logical clause 'with a high enough degree of probability to be useful to the observer or attacker' is valid. As a result, unlinkability is not satisfied.

Example #0 highlights one of the reasons why approaches built exclusively upon cryptography are limited - their sole goal is to improve indistinguishability (i.e. to extend anonymity set) [13, 61]. Game theory is in clear contrast to that: it studies incentives to use different options for which one can find many interpretations [25, 48]. Some of these interpretations may have statistical meaning such as the frequency of authentication.

In the provided example, the probability of error also agrees with *Simpson diversity* (or Gini-Simpson index) measure which has a direct relation to information entropy [47]. In general, greater entropy provides better unlinkability. For instance, if the above example is considered as a *game* where players decide how frequently they authenticate to RP we conclude that the **best** decision is $\Pr(A) = \Pr(B) = 0.5$. In such case, probability of error for any answer produced by RP is 0.5 (Simpson and Shannon entropy also achieve their maxima). Such error rate is high which implies that unlinkability is satisfied. Despite the fact that the solution exists, the practicality of that example is limited. This is because the frequency of authentication events are dictated by the user's basic needs (not privacy), which implies that marginal probabilities $\Pr(A)$, $\Pr(B)$ can rarely be changed.

Also, importance of indistinguishability should not be undermined: this quality may be insufficient for unlinkability, but distinguishability is clearly detrimental. If, for instance, *Alice* and *Bob* possess different assertions α and β, respectively, RP is able to link them with an absolute certainty (probability of error and information entropy is 0).

However, aspects of linkability threat can change when *Alice* and *Bob* possess multiple assertions that they can use *interchangeably*. Further we will demonstrate that other notions of entropy are also

important for unlinkability which becomes even more apparent in the case when indistinguishability can not be ensured. As a result, different kind of decisions is needed to improve unlinkability.

For a more granular description of the threats and unlinkability in attribute based authentication system, we define (AP, RP, U)-model: it consists of Attribute Provider (AP), Relying Party (RP), and user (U) (see fig. 1). Further in the text, we will use terms *credential*, *assertion*, and *attribute*.

Definition 2 (Credential [33])**.** *An object or data structure that authoritatively binds an identity - via an identifier or identifiers - and (optionally) additional attributes, to at least one authenticator possessed and controlled by a subscriber.*

Definition 3 (Assertion [34])**.** *A statement to a relying party that contains identity attributes about a subject. Assertions may also contain authentication or other identity information about the subject.*

Definition 4 (Attribute [34])**.** *A reference of a named quality or characteristic inherent in or ascribed to someone or something.*

Terminology will be used in the following context throughout this study. User registers with AP. Upon a request from the user, AP provides *credential(s)*. These credential(s) usually contain *attribute(s)*. A user then derives an *assertion* from this credential(s), and submits this assertion to RP to get authenticated. Among all the unlinkability definitions in ISO/IEC 27551 standard 'RP+AP-U unlinkability' is characterized by Unlinkability Level (UL) 5 which corresponds to the *highest degree of anonymity* [43].

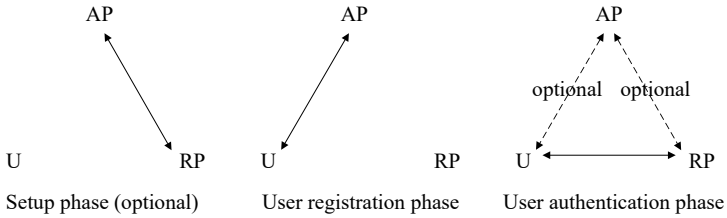


Fig. 1. Attribute based authentication phases in (AP, RP, U)-model [43]

Definition 5 (RP+AP-U unlinkability)**.** *Is the unlinkability in the system where adversary can observe, actively intercept and modify exchanged messages and additionally plays the role of both RP and AP. The target entity role is U.*

A number of assumptions about protocol *correctness* and its *unforgeability* are made within the ISO standard. It also exclusively targets application communication layer, and we make identical assumptions in this paper. For instance, parameters such as Round Trip Time (RTT) or IP address of a user may be extracted by malicious RP from the details of authentication/authorization protocol. These parameters can be further used to link users. However, these problems fall behind the scope of our paper: RP+AP-U-unlinkability is necessary but not sufficient for anonymity. We demonstrate that this necessary condition has not been addressed to date for the systems where users possess multiple attributes. In addition, the complexity that is associated with solving the problem is likely to increase in the future due to the diversity among the assertions used by the users increasing. This implies that importance of consistent assumptions about attribute based authentication systems, models for unlinkability, and efficient algorithms will increase over time.

Regrettably, the positive effect of ISO 27551 is impeded due to the lack of specification that describe *conformance* procedures for the definitions provided there. In other words, the standard defines *what* unlinkability is but it does not explain *how* this can be achieved. In this paper, we

address this limitation by developing so-called *security assurance argument* using indirect Security Assurance Conformity Assessment (SACA) method which is based on information theory and game theory [41, 42].

Th motivation behind a game-theoretical inquiry into unlinkability in multi-attribute authentication systems stems from several observations. First, number of various credentials admissible by Relying Parties (RPs) grows due to the expanding number of digital activities and sources that produce them. For example, increasing demand for new high-assurance eKYC procedures entails that assertions produced as a part of these procedures will be reused on different occasions [14, 44]. Second, assertions that are derived from these credentials differ: even advanced privacy preserving techniques such as selective disclosure and Zero Knowledge Proofs (ZKP) do not provide indistinguishability. Third, user preferences over this ever-growing set of assertions are non-trivial: user *Alice* is hardly able to predict which assertion (and how often) will be used by user *Bob* even if *Alice* knows the complete set of *Bob's* assertions.

We demonstrate that user decisions impact *unlinkability* in the authentication system and should not be random. As a result, users face a difficult dilemma under uncertainty which has not been discussed in the literature to date. To articulate importance of the problem we will further support our point with greater details.

1.1 The growing number of credentials

The number of credentials in use is defined by the variety of credentials issued and supplied to the users by APs on one hand, and the demand for these credentials from RPs on the other hand.

1.1.1 Increasing variety of issued credentials. A number of new governmental and private initiatives have emerged in the domain of attribute-based authentication over recent years. These initiatives facilitate the expansion on the set of attributes that are commonly used by users. Electronic identification, authentication and trust services (eIDAS) is among these initiatives [5, 27]. This regulation was designed to ensure that up-to-date procedures and techniques for digital identity and trust verification and validation can be used by EU citizens across the borders of their native countries. As a part of the regulation, citizens can produce assertions about their personal attributes such as their name, date of birth and address in the form which is verifiable by all accredited RPs within the EU. RPs who participate in eIDAS benefit from a high Level of Assurance (LoA) of the attributes that are derived from eID during eIDAS session. Commonly, these assertions are derived from a cryptographically protected hardware chip which is embedded into personal plastic eID cards. This procedure requires certified card-readers. Moreover, some of the EU member states have recently introduced NFC-enabled eID cards: high LoA attributes can now be extracted using compatible models of smartphones [16, 17]. Currently, the majority of EU citizens still consider their eID cards as a tool for physical (in-person) identification only. As such, it is expected that cross-border interoperability championed by eIDAS will also change users' attitude and will create incentives to use their personal digital eID attributes more often [20].

With the aim to facilitate that process EU constantly monitors adoption of eIDAS and supports it. A recent initiatives prepared by European Commission proposes to revise and extend eIDAS [26]. One of the options discussed in the initiative includes among others legislative intervention aiming to extend the scope of eID regulation under eIDAS to the private sector. Notably, this would introduce new trust services for identification, authentication and for the provision of attributes, credentials and attestations. Yet another option would introduce a European Digital Identity scheme (EUid) complementary with eIDAS for citizens to access online public and private services, when identification is necessary.

In addition to governmental initiatives opportunities of eIDAS are being leveraged by private sector which eagerly adopts principles and tools of Self Sovereign Identity (SSI). For example, a number of private companies are working on solutions able to 'fill the gap between SSI and eIDAS compliance' [24, 32, 40]. Their concepts often rely on 'Trust Over IP' metamodel which creates a context specific Governance Framework that enables reliable deployment of identity ecosystem [22]. This is accomplished by utilizing concepts such as Decentralized Identifiers (DIDs) and Verifiable Credentials (VCs) [65, 68].

1.1.2 Increasing demand for attributes with high LoA. Gaining access to the customers' credentials brings multiple benefits for RPs [3]. These benefits often include: *a*) reducing transactional cost for identification; *b*) development of new services that depend on automatic and inexpensive identification; and *c*) decrease in the need for private companies to collect and store customers' personal information themselves. For instance, electronic Know Your Customer (eKYC) procedures were confirmed to have substantial advantages in a study conducted by ING group. In the loan applications governed by fully digital eKYC with real time scoring and instant disbursement 'time to cash' was only 5 minutes in average. This contrasts with 13 days period in the case where paper signature and documentation were used which required big customer's effort and often caused poor customer experience.

A growing number of countries are allowing third parties to access their digital ID databases to carry out a variety of functions and services, including for elections, financial services, and healthcare [63]. For instance, the International Telecommunications Union (ITU) identified 22 governments that enable third parties to access their digital ID systems for the purpose of conducting KYC. ITU provides a number of recommendations to the countries with a national identity system, or another similar market-wide identity system. These countries should recognize these systems as a public resource. Access to this directory, and use of it, should be open to all regulated digital financial services providers (e.g. RPs) at a reasonable cost [44]. As a result of implementing these recommendations the number of credentials that a user can submit to RP within attribute-based authentication systems will inevitably increase.

1.2 Indistinguishability is infeasible

Due to cost and budgetary constraints, a substantial number of RPs tend to integrate functions of authentication and access control. In addition, RPs encourage the usage of richer assertions to improve quality in the access control decisions and reduce risks. It is therefore intuitively clear that indistinguishability and richness of assertions are contrasting qualities that can not be satisfied concurrently.

Attribute-based access control systems rely upon attributes to not only define access control policy rules but also enforce the access control. Among the major attribute considerations for Access Control Systems *veracity* plays momentous role in supporting RPs confidence [38]. Veracity establishes the policy and technical underpinnings for semantic and syntactic correctness of attributes, and ensures that the obtained attributes are trustworthy. Attribute trustworthiness considers how well the sources of attributes are authenticated, identified, and validated. In order to achieve substantial level of veracity it is recommended that the governance body develop a "trust model" [37]. It is used to illustrate the trust chain and help determine ownership and liability of information, services, policies, and requirements for technical solutions to validate or enforce trust relationships (see fig. 2).

As per fig. 2, the root of trust is derived from many sources, such as Subject Attribute Authorities and/or Policy Developers. It is therefore recommended that allowable values of attributes as well as the methods for their provisioning are established [37]. Authoritative subject attribute provisioning

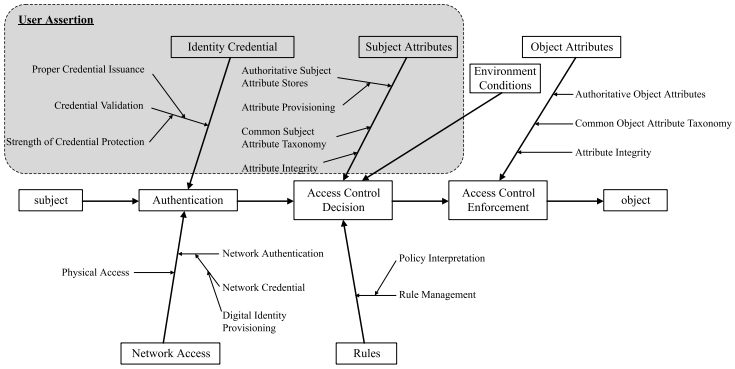


Fig. 2. ABAC trust chain [37]

capabilities should be appropriately dependable in regards to quality, assurance, privacy, and service expectations. Within specific user assertion, these characteristics can be understood through Attribute Value Metadata.

Definition 6 (Attribute Value Metadata (AVM) [34])**.** *Data describing an asserted value for an associated attribute.*

The description provided by AVM may differ depending on the requirements of user assertions as specified by RP. For example, `IssuerName` is a part of AVM in SAML2 assertions which are standard for eIDAS. This parameter describes the name of the issuer on the `MetadataList`. This list also contains attribute `Territory` which describes territory covered by metadata locations found in this element [5].

This implies that information that is communicated through AVM to RP can discriminate users based on the issuer and territory if these parameters differ among users. Such distinguishability can not be averted by the users even if selective disclosure and/or ZKP are used. This is because such privacy-preserving techniques can be applied by the user exclusively to subject attributes while metadata can not be modified: its integrity is usually protected by the means of cryptography (e.g. signed).

Business operation requirements developed by RP often address efficiency of operations. To meet efficiency requirement a special attention must be paid to attribute creation and sharing mechanism, as well as rules for maintaining attributes' privacy between APs and access control functions [38]. Unfortunately, RP does not always favor user privacy in meeting his efficiency goals [1]. For example, in order to increase his market competitiveness RP may extend the range of online services. This may result in a policy where subject is capable of performing various operations on system objects which requires assigning one or more attributes to a subject. In spite that extension of attribute set may bring convenience to the customers, it also increases customer distinguishability due to extended set of AVM which is observable by RP. The situation exacerbates even further if the same operation can be performed on object using various kinds of subject attributes.

The remainder of this paper is structured as follows. In section 2, we analyze aspects that influence decisions of the users in attribute-based authentication system. Based on this, we define **R**esearch **Q**uestion (**RQ**) which we attempt to address through Game-theoretical Approach to Privacy in Authentication Systems (GAPAS) in section 3 where we consider two variants of the non-cooperative coordination games with incomplete information: *naïve* game, and *tenable* game. This is followed

by experimental evaluation and comparison of unlinkability in these games as well as alternative scenarios of random interchangeable usage of assertions in section 4. With the aim to emphasize practical importance of GAPAS we then present and analyze in section 5 the use case involving Verifiable Credentials (VC) which is becoming a widely used format. This section also highlights on how decision making algorithms stemming from GAPAS can be implemented in Digital Credential Wallets (DCW). An extended discussion about theoretical and practical contributions of our paper is provided in section 6. We provide overview of relevant literature and conclude the paper in section 7 and section 8, respectively.

2 DEFINING THE RESEARCH QUESTION

In this section, we analyze factors that affect unlinkability in attribute-based authentication systems. This allows us to formalize Research Question (RQ) which we attempt to address in the latter parts of the paper.

Due to unavoidable differences between user assertions it is important to understand the rational principles behind assertion selection. Among privacy-aware users it is uncommon to share with each other their attributes and AVM. While it is natural to assume that coordinated usage of assertions with identical AVM is in the best interest of the users the lack of communication makes it hardly achievable. Further, we will answer the following questions: (*a*) how to measure unlinkability; (*b*) what is the guidance for the selection of attributes by the users to maximizes unlinkability; (*c*) why is unlinkability impeded by the lack of communication between the users; (*d*) why should this problem be studied using game theory.

2.1 Measuring unlinkability

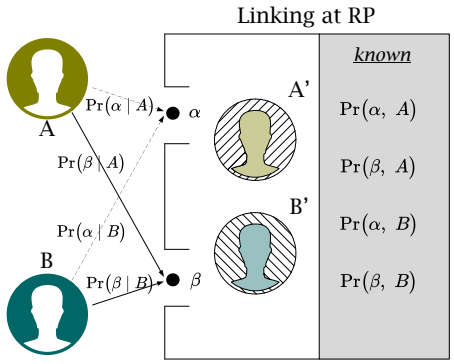
Example #1. Lets consider a simplified description of attribute-based authentication system consisting of 2 users *Alice* (further *A*), *Bob* (further *B*) and RP. We assume that RP establishes policy P which allows to perform certain actions over an object (e.g. *'to buy alcohol online'*) given that certain condition is satisfied (e.g. *'the subject is older than 18'*). Yet another restriction of P is that RP accepts assertions of 2 types only: AVM must be either **'digital driver license'** (further denoted as α) or **'digital graduate certificate'** (further denoted as β). It is presumed that *A* and *B* have both types of the credentials. Both driver license and graduate certificate may contain multiple *subject attributes* within *claim* section of the credential. These attributes may include `FirstName` and `LastName` of the subject, his/her `DOB`, `Address` etc.. In addition, degree certificate may contain information such as `Degree`, `AwardedDate`, `Institution` and so on. Despite of the differences in their *subject attributes* both *A* and *B* aim to reduce distinguishability. For this they apply privacy preserving techniques (such as selective disclosure and ZKP) to the *claim* section of the corresponding credential. This allows them to demonstrate that they satisfy P without revealing any additional information [11]. As a result, assertion with AVM α submitted by *A* can not be distinguished by RP from assertion with AVM α submitted by *B*. On the other hand, RP can distinguish assertions with α from assertions with β.

Both *A* and *B* then decide on how often they use α versus β in their authentication sessions to RP. The relative frequency of usage of α by *A* is denoted as $\Pr(\alpha \mid A)$, and, $\Pr(\beta \mid A) = 1 - \Pr(\alpha \mid A)$. Similarly, the relative frequency of usage of α by *B* is denoted as $\Pr(\alpha \mid B)$, and, $\Pr(\beta \mid B) = 1 - \Pr(\alpha \mid B)$. The goal of *A* and *B* is to produce decisions which maximize unlinkability at RP. In contrast, RP tries to link authentication sessions. He observes realizations of random variable $l \in \ell$, $\ell = \{\alpha, \beta\}$ but does not know whether *A* or *B* has initiated authentication. From the standpoint of RP, this observed realization of l is the result of the decision which is made by the user whose label is random variable $\mathsf{L} \in \mathcal{L}$, $\mathcal{L} = \{A, B\}$ (see fig. 3a). For instance, it is intuitively clear that

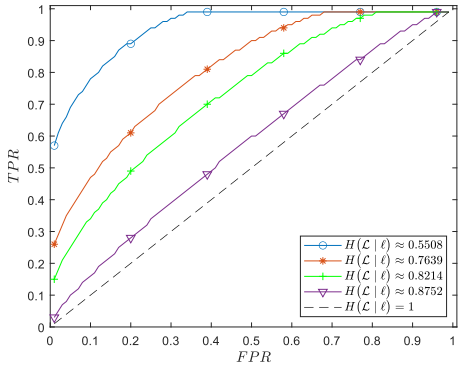
unlinkability is high if $\Pr(\alpha \mid A) = \Pr(\alpha \mid B) = 1$, e.g. both A and B always authenticate to RP with their digital driver license. This is because RP can not distinguish their AVM. The opposite happens if $\Pr(\alpha \mid A) = \Pr(\beta \mid B) = 1$, e.g. A always uses driver license, but B always uses degree certificate when they authenticate to RP. In that case RP can distinguish different entities behind authentication events as soon as he observes α in one event and β in the other.

To differentiate the result of 'linking' which is conducted by RP from the actual label of the user we will use L' instead of L. Questions *'How does RP build his linking detector?'* and *'What are the characteristics of it?'* are out of the scope of our paper. Instead, we demonstrate that the upper performance of such linking at RP is limited by conditional entropy $H(\mathsf{L} \mid l)$.

Lemma 1. *Best linking performance is limited by* $H(\mathsf{L} \mid l)$ *(for details see appendix A).*



(a) Scheme of linking based on attribute selection.



(b) ROC curves for optimal linking function built by RP.

Fig. 3. Linking in attribute based authentication system with 2 users.

In addition to lemma 1 the plots of Receiver Operating Characteristics (ROC) curves illustrate effect of conditional entropy on fig. 3b: for a given False Positive Rate (FPR) the highest possible True Positive Rate (TPR) decreases with $H(\mathsf{L} \mid l)$. Due to this property the measure of conditional entropy can be used by A and B to produce decisions which improve unlinkability. To further understand benefits of the proposed measure we compare it with the RP+AP-U-unlinkability test in ISO 27551 (see listing 1). This test specifies on whether definition 5 is satisfied. We further regard that entities U_0 and U_1 in the test are played by *Alice* and *Bob*, respectively, from the coordination game on fig. 3a. Among the major similarities between the coordination game and the test are: (1) realizations α, β in the game cohere with the set of attributes that satisfy access policy P in the test; (2) authentication events (protocol executions) repeat over time; (3) performance is measured based on the conditional probability of correct guess (made by RP) given the value of realization/attribute. In spite of these similarities the test in ISO/IEC DIS 27551 does not allow to evaluate the performance for the game. This is because the test: (*a*) demands that $\Pr(A) = \Pr(B) = 0.5$ which is not always satisfied on practice; (*b*) details of '*guessing*' procedure (line 7, listing 1) remain unclear. In addition, unlinkability conformance procedures remain unaddressed by the standard: it is unclear whether unlinkability can be improved if A and B can select among different assertions, and *how* this selection should be done. Next, we will analyze if $H(\mathsf{L} \mid l)$ can be used to provide optimal unlinkability recommendations (e.g. conformance) to A and B.

Listing 1. RP+AP-U unlinkability, ISO/IEC DIS 27551

Output: `true` or `false`
Test:
  Adversary $\mathfrak{A}$ chooses the set of attributes for U_0, U_1 and policy P ;
  AP and RP execute the setup phase (if any) ;
  AP and U_0 execute the user registration phase (if any) ;
  AP and U_1 execute the user registration phase (if any) ;
  RP, AP and U_0 execute the authentication phase ;
  RP, AP and U_b execute the authentication phase, $b \in \{0,1\}$, $\Pr(b = 0) = 0.5$;
  $\mathfrak{A}$ returns a *guess* $b' \in \{0,1\}$ on the value of b ;
  if $\Pr(b' = b) \to 0.5$ return `true` ;
  else return `false` .

2.2 Optimal selection of assertions

As per lemma 1, in order to improve unlinkability at RP users A and B may coordinate with one another to increase $H(\mathsf{L} \mid l)$. From **Example #1**, best response for A and B is $\Pr(\alpha \mid B) = \Pr(\alpha \mid A)$. This also implies that $\Pr(\beta \mid B) = \Pr(\beta \mid A)$. In contrast, if $\Pr(\alpha \mid A) = 1$ and $\Pr(\beta \mid B) = 1$ RP can link users with 100% accuracy. Coordination in the context of the described best response example is simple. This is because (a) the number of users is just 2; (b) they both use the same set of attributes; and (c) there is an implicit assumption that information for coordination is known to the users. Further we will consider more realistic scenarios where unlinkability is measured using criterion $C = H(\mathsf{L} \mid l)$ for the system with $n \gg 2$ users whose assertions may differ, and who use a specific coordinating agreement to improve unlinkability.

2.3 Unlinkability and the lack of communication

It is intuitive to suggest that in the system with $n \gg 2$ users that they may coordinate by committing to some sort of de-facto agreement. This does not require direct communication between the users. On the other hand, such agreement may be sufficient to govern relative frequencies of utilization of the assertions possessed by the users. In the following example we shall see how adherence to such an agreement (and possible deviations from it) affect unlinkability criterion $C = H(\mathsf{L} \mid l)$.

Example #2. We amend setting of **Example #1** to make them more realistic. One observation is that for randomly selected 2 out of n users probability that both of their assertions match must be less than 1. This, for instance, is likely to happen due to substantial granularity of AVM which reflects variety of authoritative sources referenced in user assertions (see fig. 2). We encompass this by considering 2 kinds of driver licenses and 2 kinds of graduate certificates that are distributed among n users (see table 1).

Table 1. Credential ownership by category and AVM realization

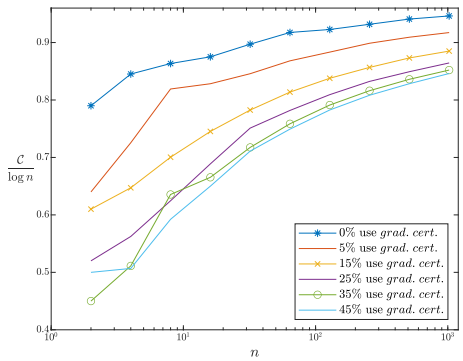
Category	**AVM realization**		**Total**
Drivers License	`Restricted` 771, 855(12.2%)	`Unrestricted` 5, 580, 224(87.8%)	6, 319, 611
Tertiary Qualification	`Undergraduate` 4, 030, 835(86.5%)	`Postgraduate` 631, 121(13.5%)	4, 661, 956

Drivers License that are issued in the state of Victoria (Australia) follows a Graduate Licensing Scheme (GLS) where a subject must go through a '`Restricted`' phase. This phase lasts until a specific level of driving experience is accumulated by the subject at which stage an appropriate exams needs to be passed. Successful completion of these steps would allow an individual to

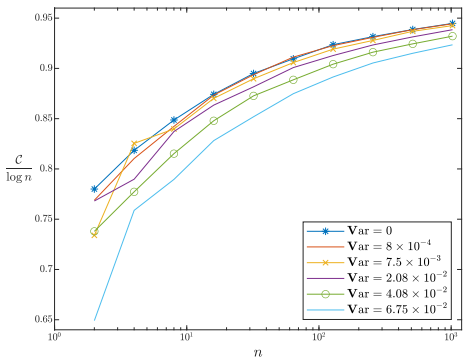
transition into a 'Unrestricted' phase. The described flow implies that an individual can have a driver license assertion with AVM which refers to either an restricted or non-restricted type. However, an individual can not have digital driver licenses of both types. The data in table 1 stipulates that as of 2019, 12.2% of users had restricted category license, whereas 87.8% of the users had unrestricted (i.e. full) license.

Similarly, an AVM for digital graduate certificate may designate that it is either 'Undergraduate' or 'Postgraduate'. Due to the same reasons pertaining to the digital drivers license for Victoria, a person can not possess a postgraduate certificate if they have only finished their undergraduate program. On the other hand, your final qualification level cannot be displayed at undergraduate level if you have finished a postgraduate program. From table 1 it can be observed that 86.5% and 13.5% of users have undergraduate and postgraduate certificates, respectively.

We now formalize allocation of assertions among n users. Category $\mathcal{A} = \{\alpha_1, \alpha_2\}$ represents **driver licenses** where α_1 stands for 'Restricted' AVM, and α_2 stands for 'Unrestricted' AVM. In addition, category $\mathcal{B} = \{\beta_1, \beta_2\}$ represents **graduate certificates** where β_1 denotes 'Undergraduate' AVM, and β_2 denotes 'Postgarduate' AVM. We use indices $i \in \{1, ..., n\}$ to differentiate users. A pair of random variables $(\alpha^{(i)}, \beta^{(i)})$, $\alpha^{(i)} \in \mathcal{A}$, $\beta^{(i)} \in \mathcal{B}$ encodes *type* of a user with index i. According to table 1 $\Pr(\alpha^{(i)} = \alpha_1) = 0.122$, $\Pr(\alpha^{(i)} = \alpha_2) = 0.878$, and $\Pr(\beta^{(i)} = \beta_1) = 0.865$, $\Pr(\beta^{(i)} = \beta_2) = 0.135$. Assuming that $\alpha^{(i)}$ and $\beta^{(i)}$ are independent we obtain the following joint probabilities: $\Pr(\alpha^{(i)} = \alpha_1, \beta^{(i)} = \beta_1) \approx 0.106$, $\Pr(\alpha^{(i)} = \alpha_1, \beta^{(i)} = \beta_2) \approx 0.017$, $\Pr(\alpha^{(i)} = \alpha_2, \beta^{(i)} = \beta_1) \approx 0.759$, $\Pr(\alpha^{(i)} = \alpha_2, \beta^{(i)} = \beta_2) \approx 0.118$. Further we will use matrix $\aleph$ to represent this distribution of types across the users. For table 1 we obtain $\aleph = [0.122,\ 0.878]^{\mathrm{T}} \times [0.865,\ 0.135]$. Since every user i has only two alternatives, their decision can be represented using scalar value $s_i = \Pr(\alpha^{(i)} \mid i)$ which is a continuous strategy, in general.



(a) Users select and use 1 attribute only. Majority selects *driver license*.



(b) Users use both of their attribute realizations interchangeably with $\mathbb{E}[s_i] = 0.5$.

Fig. 4. Unlinkability under various conditions.

Next, we need to analyze how adherence to the de facto agreement affects unlinkability, and how this unlinkability depends on the number n of users in the system. For this we consider two kinds of agreements where the results of corresponding experiments are presented on fig. 4. For simplicity, we assume that every user authenticates to RP with the probability equal to $\frac{1}{n}$. As we shall see further in the text, this implies that $C \leq \log n$. From this stems our motivation to measure performance against normalized unlinkability rate $\frac{C}{\log n}$.

Results for the agreement to use only 1 out of the 2 available assertions are presented on fig. 4a. The agreed strategy for user i is denoted as $\dot{s}_i$ while actual strategy played by the user is s_i. According to the agreement, users must use driver license only, meaning that $\forall i(\dot{s}_i = 1)$. In this instance, we consider that strategy of each user is a discrete random variable, $s_i \in \{0, 1\}$. As per figure, unlinkability increases with the percentage of users who adhere to the agreement. Results for a different agreement are depicted on fig. 4b. In this instance, users agree to use both of their assertions interchangeably and with equal probability, meaning that $\forall i(\dot{s}_i = 0.5)$. The strategy played by each user is a continuous random variable $s_i \in [0, 1]$. Contrary to the previous agreement, the degree of deviations is reflected with the variance. This is because the mere fact that i deviates from the agreement does not communicate '*how strong*' the deviation is. As can be seen from the figure, performance of the system is better for the cases with lower variance (e.g. better adherence to the agreement). For the both experiments on fig. 4 ratio $\frac{C}{\log n}$ increases with n meaning that importance of the de facto agreement reduces. Despite this fact importance of coordinating agreement for small-and-medium-sized RPs can not be underestimated.

While the aforesaid agreements are beneficial (and departures are harmful), they are exemplar only. It is therefore premature to consider these special cases optimal since many other agreements may be arranged for a system of n users whose types are distributed in accordance with $\aleph$. Hence, the following research question remains:
RQ: How should users use their assertions to maximize RP+AP-U-unlinkability?

2.4 Game theory for unlinkable authentication

Based on the evidence presented, we argue that game theory is a suitable tool for exploration of the RQ. This is because firstly, the conditional entropy which is used for criterion C is an integral measure where each summand may depend on the decisions of many users. As such, for every user i information about decisions of other users (we denote them $-i$) must be communicated or assumed in the form of belief. As we have previously demonstrated in fig. 4, an arbitrary de-facto agreement between the users may enable them to communicate such information. Nonetheless, a central question for i is whether his/her best interest is to produce decision that supports such agreement (or deviates from it). These questions are studied by decision making and game theory. In particular, principle of non-deviation (equilibrium) is famous due to Nash [57]. Also, questions of information, its meaning, origins and tools for its dissemination are often asked within game theoretical contexts. Moreover, a deeper line of thoughts guides us through the situations where tools for information synchronization (such as agreement) are not existent or can not be trusted by i. Some of the concepts are proven to remain consistent irrespective of the decisions made by others and are often applied in robust decision making [67]. Finally, results of non-cooperative game theory can enhance unlinkability on individual level and do not require modifications of existing authentication protocols. This game-theoretical results can therefore become applicable in Digital Credential Wallets (DCW) which store and manage user credentials [64].

3 GAME-THEORETICAL APPROACH TO PRIVACY IN AUTHENTICATION SYSTEMS (GAPAS)

In this section, we formalize game theoretical model for the attribute-based authentication system with n users who use their assertions interchangeably. Decisions of the players are guided by the principle of *best response*. To calculate it we find expressions for the expected utilities of the players. For this we utilize results of lemma 1 and make assumption about information that is known to the players. This information is in the form of priors (or beliefs) over the set $\mathbf{M}$ of marginal probabilities at RP, and can, for example, be communicated to the players through mediator M. To ensure that

this information is consistent with the principle of best response we express conditions for Nash equilibrium and use these conditions to define **M**. However, for the case when information about **M** can not be communicated (trusted M does not exist) to the players we utilize Wald maximin approach.

In this section, we consider that all users are players in the game. We also presume that terms 'attributes', 'credentials' and 'assertions' have the same meaning here.

Table 2. Important Notations

Notation	Description
GAPAS	Game-theoretical Approach to Privacy in Authentication Systems
VC, VP	Verifiable Credentials, Verifiable Presentation
RP, AP	Relying Party, Attribute Provider
ROC	Receiver Operating Characteristics
DCW	Digital Credentials Wallet
$n \geq 2$	Number of players in the game
$\mathcal{A} = \{\alpha_1, \ldots, \alpha_l\}$, $\mathcal{B} = \{\beta_1, \ldots, \beta_m\}$	Categories of attribute realizations.
$\ell = \mathcal{A} \cup \mathcal{B}$	Full set of user attribute realizations.
$l \in \ell$	Discrete random variable in set ℓ
$\mathcal{L} = \{A, B\}$	Set of user labels A, B in 2-player game.
$\mathsf{L} \in \mathcal{L}$	Discrete random variable in set $\mathcal{L}$.
$I = \{1, ..., n\}$	Set of indices of the players in n-player game, $n \gg 2$.
$i \in I$	Discrete random variable (player's index) in set I.
$\alpha^{(i)} \in \mathcal{A}$, $\beta^{(i)} \in \mathcal{B}$	Attributes of a random player i
$H(\cdot\|\cdot)$	Conditional entropy.
$\mathbf{t}_i = (\alpha^{(i)}, \beta^{(i)})$	Type of player i.
$\mathcal{T} = \{\mathbf{t}_1, ...\mathbf{t}_n\}$	Set of types for all n players.
$\Pi := \boldsymbol{\pi}_1 \times \cdots \times \boldsymbol{\pi}_n$	Set of pure strategies $\boldsymbol{\pi}_i$ of all players $i \in I$.
$S_i, i \in I$	Subset consisting of continuous strategies $\boldsymbol{s}_i$
$\mathcal{S} := S_1 \times \cdots \times S_n$	Set of all continuous strategies for the players.
u_i	Payoff function of player i, $i \in I$
$\Im = \langle I, \mathcal{T}, \Pi, \mathcal{S}, u \rangle$	Game over the sets $I, \mathcal{T}, \Pi, \mathcal{S}, u$.
$\Omega_{(\cdot)}$	Set of players with realization $(\cdot)$
$\aleph$	Discrete joint probability mass function over $(\alpha^{(i)}, \beta^{(i)})$.
$\beth$	Distribution over $\mathcal{T} \times \mathcal{S}$.
$\Pr_{\mathcal{S}}(\alpha^{(i)})$, $\Pr_{\mathcal{S}}(\beta^{(i)})$	Marginal probabilities at RP (for $\mathbf{t}_i$)
M	Complete set of marginal probabilities at RP
$\mathbb{E}[\cdot]$	Expected value
$\mathrm{Var}(\cdot)$	Variance

3.1 Game theoretical model

We define a game $\Im = \langle I, \mathcal{T}, \Pi, \mathcal{S}, u \rangle$, where $I = \{1, 2, ..., i, ..., n\}$ denotes the set of indices i for the players (we will use $-i$ to denote all other players except i); $\mathcal{T} = \mathbf{t}_1 \times ... \times \mathbf{t}_i \times ... \times \mathbf{t}_n$ is the set of all players' types where $\mathbf{t}_i = (\alpha^{(i)}, \beta^{(i)})$. Random variables $\alpha^{(i)}$ and $\beta^{(i)}$ are drawn from distinctly different categories $\mathcal{A} = \{\alpha_1, ..., \alpha_\iota, ..., \alpha_l\}$ and $\mathcal{B} = \{\beta_1, ..., \beta_\rho, ..., \beta_m\}$, respectively, $\mathcal{A} \cap \mathcal{B} = \varnothing$ and $\ell = \mathcal{A} \cup \mathcal{B}$. Discrete joint **p**robability **m**ass **f**unction (**pmf**) $\aleph$ describes distribution of $(\alpha^{(i)}, \beta^{(i)})$ over $\mathcal{A} \times \mathcal{B}$ such that $\Pr(\alpha^{(i)} = \alpha_\iota, \beta^{(i)} = \beta_\rho) = \aleph_{\iota,\rho}$. $\Pi = \boldsymbol{\pi}_1 \times \cdots \times \boldsymbol{\pi}_n$ is the set of discrete pure strategies for all players, e.g. $\pi_{i,t_i} \in \boldsymbol{\pi}_i$ denotes t_i-th strategy available to player i; $\mathcal{S} = S_1 \times \cdots \times S_n$ is the set of all continuous strategies for the players, containing subsets S_i including (perhaps infinitely many) probability vectors $\mathbf{s}_i : \boldsymbol{\pi}_j \to [0, 1]^{|\boldsymbol{\pi}_i|}$ defining continuous strategies, such that

$\mathbf{s}_i(\pi_{i,t_i}) \geq 0$ and $\sum_{\boldsymbol{\pi}_i} \mathbf{s}_i(\pi_{i,t_i}) = 1$; $u_i : \mathcal{T} \times \mathcal{S} \to \mathbb{R}$ is a payoff function of player i over a profile of types and continuous strategies.

Throughout the paper, the attribute selection is restricted to just 2 alternatives. As a result, pure strategies of player i will be represented by $\boldsymbol{\pi}_i = \{\pi_{i,1}, \pi_{i,2}\}$ where $\pi_{i,1}$ should be interpreted as 'player i authenticates to RP with the realization of $\alpha^{(i)}$', and $\pi_{i,2}$ should be interpreted as 'player i authenticates to RP with the realization of $\beta^{(i)}$'. Then, continuous strategy can be expressed with a scalar $s_i \in [0, 1]$ where s_i is the probability $\Pr(\alpha^{(i)} \mid i)$, while $1 - s_i$ is the probability $\Pr(\beta^{(i)} \mid i)$. This should be interpreted as: 'player i authenticates to RP with the realization of $\alpha^{(i)}$ in $100s_i\%$ of authentication sessions (randomly selected by him) while in the rest $100(1 - s_i)\%$ of all sessions he uses realization of $\beta^{(i)}$'.

Definition 7 (GAPAS). *Game-theoretical Approach to Privacy in Authentication Systems enables user i to increase unlinkability. This is done by maximising $\mathbb{E}[u_i]$ in game $\Im$ which requires adjusting s_i.*

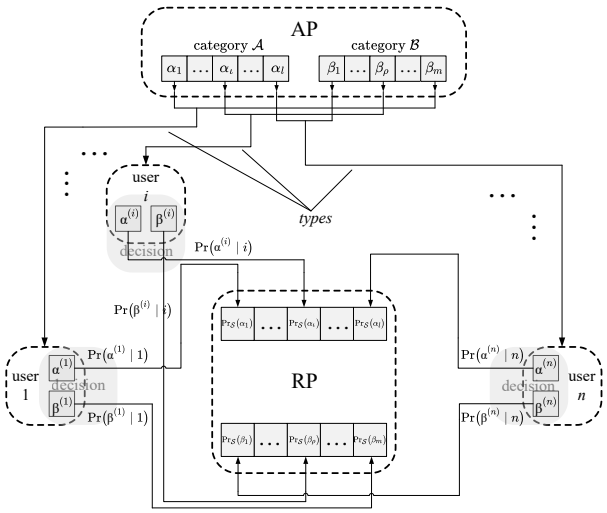


Fig. 5. Schematic explanation of interactions between n users, AP, and RP

The overview of (AP, RP, U)-model where n users produce their decisions is provided on fig. 5. It describes the following interactions:

(a) Attribute Provider (AP) supplies ready-to-use attributes/assertions to n players;

(b) for every player i, the pair of received attributes $(\alpha^{(i)}, \beta^{(i)})$ specifies his/her type $\mathbf{t}_i$. Every i decides upon continuous strategy $s_i = \Pr(\alpha^{(i)} \mid i)$. All n users then authenticate to RP on multiple occasions during some prolonged time period t and use their attributes interchangeably according to s_i;

(c) RP authenticates users and registers all authentication events. As such, he knows how frequently every of attribute realization from $\mathcal{A} = \{\alpha_1, ..., \alpha_\iota, ..., \alpha_I\}$ and $\mathcal{B} = \{\beta_1, ..., \beta_\rho, ..., \beta_m\}$, is used, but does not know the label i (or '*id*') of the user who authenticates to RP with specific attribute realization at specific authentication session. We call these relative frequencies $\Pr_{\mathcal{S}}(\alpha^{(i)})$, $\Pr_{\mathcal{S}}(\beta^{(i)})$ '*marginal probabilities at RP*': every user i needs to estimate them to produce his/her *best response.*

Definition 8 (Marginal probabilities at RP). *The set* **M** *of marginal probabilities at RP is given as follows:*

$$\mathbf{M} = \bigcup_{i=1}^{n} \left\{ \Pr_{\mathcal{S}}\left(\alpha^{(i)}\right), \Pr_{\mathcal{S}}\left(\beta^{(i)}\right) \right\} \text{ where}$$

$$\begin{aligned} \Pr_{\mathcal{S}}\left(\alpha^{(i)}\right) &= \sum_{j=1}^{n} \Pr\left(\alpha^{(i)} = \alpha^{(j)}\right) \Pr(j) s_j \,, \\ \Pr_{\mathcal{S}}\left(\beta^{(i)}\right) &= \sum_{j=1}^{n} \Pr\left(\beta^{(i)} = \beta^{(j)}\right) \Pr(j) (1 - s_j) \,. \end{aligned}$$

We stress the difference between: (*i*) marginal probabilities $\Pr\left(\alpha^{(i)}\right)$, $\Pr\left(\beta^{(i)}\right)$ at AP which is the probability that a randomly supplied attribute (to a randomly selected player i) takes certain realization from $\mathcal{A}$, $\mathcal{B}$, respectively - this can be obtained from $\aleph$; and (*ii*) marginal probabilities $\Pr_{\mathcal{S}}\left(\alpha^{(i)}\right)$, $\Pr_{\mathcal{S}}\left(\beta^{(i)}\right)$ at RP. In deterministic settings, this can be either directly **provided** by RP (or by independent mediator M) **to** the players, or can be **calculated by** the players if the set of all decisions $\{s_1, ..., s_i, ..., s_n\}$ is known to the players. Also, information about elements in **M** may be non-deterministic for players and, hence, we will talk about **priors** over **M**.

3.2 Model analysis

Here we establish relation between $\Pr_{\mathcal{S}}\left(\alpha^{(i)}\right)$, $\Pr_{\mathcal{S}}\left(\beta^{(i)}\right)$ and player decisions on one hand, and C on the other hand. We then use C to derive expected utilities $\mathbb{E}\left[u_i\right]$, and to calculate best responses s_i^{b}.

Based on lemma 1, collective unlinkability of the entire authentication system with n users is $C = H\left(i \mid l\right)$:

$$\begin{aligned} C = -\sum_{i=1}^{n} \Pr(i) \Bigg(& \Pr\left(\alpha^{(i)} \mid i\right) \log \frac{\Pr\left(\alpha^{(i)} \mid i\right)}{\Pr_{\mathcal{S}}\left(\alpha^{(i)}\right)} \Pr(i) \\ & + \Pr\left(\beta^{(i)} \mid i\right) \log \frac{\Pr\left(\beta^{(i)} \mid i\right)}{\Pr_{\mathcal{S}}\left(\beta^{(i)}\right)} \Pr(i) \Bigg) = \sum_{i=1}^{n} \Pr(i) \log \frac{1}{\Pr(i)} \\ & - \sum_{i=1}^{n} \Pr(i) \Bigg(\Pr\left(\alpha^{(i)} \mid i\right) \log \frac{\Pr\left(\alpha^{(i)} \mid i\right)}{\Pr_{\mathcal{S}}\left(\alpha^{(i)}\right)} \\ & \qquad + \Pr\left(\beta^{(i)} \mid i\right) \log \frac{\Pr\left(\beta^{(i)} \mid i\right)}{\Pr_{\mathcal{S}}\left(\beta^{(i)}\right)} \Bigg) \,, \end{aligned} \tag{1}$$

where $\Pr\left(\alpha^{(i)} \mid i\right) = s_i$ and $\Pr\left(\beta^{(i)} \mid i\right) = 1 - s_i$. Assuming that $\forall i \Pr(i) = \frac{1}{n}$ we can rewrite eq. (1) as

$$C = \log n - \frac{1}{n} \sum_{i=1}^{n} \left(s_i \log \frac{s_i}{\Pr_{\mathcal{S}}(\alpha^{(i)})} + (1 - s_i) \log \frac{1 - s_i}{\Pr_{\mathcal{S}}(\beta^{(i)})} \right) . \tag{2}$$

Our task is then to propose utilities for the players such that every player maximizing his/her utility will maximize $\mathbb{E}\left[C\right]$.

Assumption 1. *Priors over* **M** *satisfy the following scaling constraint for all* $i \in I$*:*

$$\frac{\mathrm{Var}\left[\Pr_{\mathcal{S}}\left(\alpha^{(i)}\right)\right]}{\mathbb{E}\left[\Pr_{\mathcal{S}}\left(\alpha^{(i)}\right)\right]^2} = \frac{\mathrm{Var}\left[\Pr_{\mathcal{S}}\left(\beta^{(i)}\right)\right]}{\mathbb{E}\left[\Pr_{\mathcal{S}}\left(\beta^{(i)}\right)\right]^2} = \text{const.}$$

Lemma 2. *Expected utility for player i is (for details see appendix A)*

$$\mathbb{E}\left[u_i\right] \approx -s_i \log \frac{s_i}{\mathbb{E}\left[\Pr_{\mathcal{S}}\left(\alpha^{(i)}\right)\right]} - \left(1-s_i\right) \log \frac{1-s_i}{\mathbb{E}\left[\Pr_{\mathcal{S}}\left(\beta^{(i)}\right)\right]} . \tag{3}$$

Based on lemma 2 we derive best response for player $i \in I$.

Corollary 1. *For **continuous** strategies, **best response of player** i is defined as*

$$s_i^{\mathrm{b}} = \frac{\mathbb{E}\left[\Pr_{\mathcal{S}}\left(\alpha^{(i)}\right)\right]}{\mathbb{E}\left[\Pr_{\mathcal{S}}\left(\alpha^{(i)}\right)\right] + \mathbb{E}\left[\Pr_{\mathcal{S}}\left(\beta^{(i)}\right)\right]} , \tag{4}$$

*while for i playing **discrete** strategies*

$$s_i^{\mathrm{b}} = \begin{cases} 1 & \text{if } \mathbb{E}\left[\Pr_{\mathcal{S}}\left(\alpha^{(i)}\right)\right] > \mathbb{E}\left[\Pr_{\mathcal{S}}\left(\beta^{(i)}\right)\right] ; \\ 0 & \text{if } \mathbb{E}\left[\Pr_{\mathcal{S}}\left(\alpha^{(i)}\right)\right] < \mathbb{E}\left[\Pr_{\mathcal{S}}\left(\beta^{(i)}\right)\right] , \end{cases} \tag{5}$$

*and, i is **indifferent*** if $\mathbb{E}\left[\Pr_{\mathcal{S}}\left(\alpha^{(i)}\right)\right] = \mathbb{E}\left[\Pr_{\mathcal{S}}\left(\beta^{(i)}\right)\right]$.

The proof for corollary 1 is straightforward and we omit it here.

Remark 1. ***Players of the same type*** *produce identical best responses and have identical best expected utilities:*

$$\begin{gathered} \forall i, j (\mathbf{t}_j = \mathbf{t}_i) \implies \mathbb{E}\left[u_j^{\mathrm{b}}\right] = \mathbb{E}\left[u_i^{\mathrm{b}}\right] \\ = \log\left(\mathbb{E}\left[\Pr_{\mathcal{S}}\left(\alpha^{(i)}\right)\right] + \mathbb{E}\left[\Pr_{\mathcal{S}}\left(\beta^{(i)}\right)\right]\right) . \end{gathered} \tag{6}$$

The result of remark 1 follows directly from the expressions within definition 8 and corollary 1. Next, we need to address question of consistency of expectations [36].

Definition 9 (Consistent expectations)**.** *Best responses $\mathcal{S}^{\mathrm{b}}$ of all n players must satisfy for all $i \in I$:*

$$\Pr_{\mathcal{S}^{\mathrm{b}}}\left(\alpha^{(i)}\right) = \mathbb{E}\left[\Pr_{\mathcal{S}}\left(\alpha^{(i)}\right)\right] \wedge \Pr_{\mathcal{S}^{\mathrm{b}}}\left(\beta^{(i)}\right) = \mathbb{E}\left[\Pr_{\mathcal{S}}\left(\beta^{(i)}\right)\right]. \tag{7}$$

One way to enable eq. (7) is to find the conditions supporting Nash equilibrium in the form

$$s_i^{\mathrm{b}} = \frac{\Pr_{\mathcal{S}^{\mathrm{b}}}\left(\alpha^{(i)}\right)}{\Pr_{\mathcal{S}^{\mathrm{b}}}\left(\alpha^{(i)}\right) + \Pr_{\mathcal{S}^{\mathrm{b}}}\left(\beta^{(i)}\right)} \quad \text{for all } i \in I, \tag{8}$$

and then to make sure that the priors on $\mathbf{M}$ are formed accordingly. This can be communicated to the players through a mediator M. We will next formulate eq. (8) using information of players' types. This is possible because the whole set of players' indices I can be represented using subsets that have direct reference to all possible realizations of $\alpha^{(i)}$, $\beta^{(i)}$. We use Ω_{α_ι} and Ω_{β_ρ} such that $\forall \phi, \xi \in I$

$$\left(\alpha^{(\xi)} = \alpha_\iota\right) \iff \left(\xi \in \Omega_{\alpha_\iota}\right) , \ \left(\beta^{(\phi)} = \beta_\rho\right) \iff \left(\phi \in \Omega_{\beta_\rho}\right) .$$

We denote $\Omega_{\alpha_\iota,\beta_\rho} = \Omega_{\alpha_\iota} \cap \Omega_{\beta_\rho}$ for which $\Omega_{\alpha_\iota} = \bigcup_{\rho=1}^{m} \Omega_{\alpha_\iota,\beta_\rho}$ and $\Omega_{\beta_\rho} = \bigcup_{\iota=1}^{l} \Omega_{\alpha_\iota,\beta_\rho}$ holds. For $i \in \Omega_{\alpha_\iota,\beta_\rho}$ we set $\Pr_{\mathcal{S}^{\mathrm{b}}}\left(\alpha^{(i)}\right) = \frac{1}{n} \sum_{\xi \in \Omega_{\alpha_\iota}} s_\xi^{\mathrm{b}}$ and $\Pr_{\mathcal{S}^{\mathrm{b}}}\left(\beta^{(i)}\right) = \frac{1}{n} \sum_{\phi \in \Omega_{\beta_\rho}} \left(1 - s_\phi^{\mathrm{b}}\right)$.

In line with remark 1 all players whose indices are in $\Omega_{\alpha_\iota,\beta_\rho}$ produce the same best response denoted as $\theta_{\iota,\rho}$, and, as a result $\Pr_{\mathcal{S}^{\mathrm{b}}}\left(\alpha^{(i)}\right) = \frac{1}{n} \sum_{\nu=1}^{m} |\Omega_{\alpha_\iota,\beta_\nu}| \theta_{\iota,\nu}$, while $\Pr_{\mathcal{S}^{\mathrm{b}}}\left(\beta^{(i)}\right) = \frac{1}{n} \sum_{\tau=1}^{l} |\Omega_{\alpha_\tau,\beta_\rho}| \left(1 - \theta_{\tau,\rho}\right)$. Without loss of generality, for large $n >> l \times m$, we have $\aleph_{\iota,\rho} = \frac{1}{n} |\Omega_{\alpha_\iota,\beta_\rho}|$. Validity of eq. (8) is guaranteed if for all ι, ρ:

$$\theta_{\iota,\rho} = \frac{\sum_{\nu=1}^{m} \aleph_{\iota,\nu}\theta_{\iota,\nu}}{\sum_{\nu=1}^{m} \aleph_{\iota,\nu}\theta_{\iota,\nu} + \sum_{\tau=1}^{l} \aleph_{\tau,\rho}(1-\theta_{\tau,\rho})} \ . \tag{9}$$

Next, we discuss solutions for eq. (9) in pure continuous (e.g. authentication with 2 attributes) and pure discrete (e.g. authentication with 1 attribute) strategies. We will also consider the maximin scenario for the case when neither **M** nor priors over **M** are known. We will further separate these results by referring to 'Naïve game' when **M** is provided (by mediator *M*, for instance) and 'Tenable game' when nothing is known about **M**, respectively.

3.3 Naïve game and its equilibria

This game is based on assumption that **M** is known and some of its properties were discussed in the previous subsection.

3.3.1 Players use 2 attributes interchangeably. We transform eq. (9) and take into account that sums $\sum_{\nu=1}^{m} \aleph_{\iota,\nu}\theta_{\iota,\nu}$ and $\sum_{\tau=1}^{l} \aleph_{\tau,\rho}(1-\theta_{\tau,\rho})$ have terms with common $\aleph_{\iota,\rho}$. All the possible Nash equilibria in (pure) continuous strategies are represented by the following system of nonlinear equations:

$$\begin{aligned}
&\theta_{1,1}\Big(\sum_{\nu=2}^{m} \aleph_{1,\nu}\theta_{1,\nu} + \sum_{\tau=2}^{l} \aleph_{\tau,1}(1-\theta_{\tau,1})\Big) = \sum_{\nu=2}^{m} \aleph_{1,\nu}\theta_{1,\nu} \ , \\
&\quad\vdots \\
&\theta_{\iota,\rho}\Big(\sum_{\substack{\nu=1\\ \nu\neq\rho}}^{m} \aleph_{\iota,\nu}\theta_{\iota,\nu} + \sum_{\substack{\tau=1\\ \tau\neq\iota}}^{l} \aleph_{\tau,\rho}(1-\theta_{\tau,\rho})\Big) = \sum_{\substack{\nu=1\\ \nu\neq\rho}}^{m} \aleph_{\iota,\nu}\theta_{\iota,\nu} \ , \\
&\quad\vdots \\
&\theta_{l,m}\Big(\sum_{\nu=1}^{m-1} \aleph_{l,\nu}\theta_{l,\nu} + \sum_{\tau=1}^{l-1} \aleph_{\tau,m}(1-\theta_{\tau,m})\Big) = \sum_{\nu=1}^{m-1} \aleph_{l,\nu}\theta_{l,\nu} \ .
\end{aligned} \tag{10}$$

3.3.2 Players use 1 attribute. As a result, each player only plays a discrete pure strategy. If all players with the same (ι,ρ) produce the same best response, their averaged response is also discrete, e.g $\theta_{\iota,\rho} \in \{0,1\}$. However, averaged response $\bar{\theta}_{\iota,\rho}$ of the players of the same type may be a continuous value if different players of type (ι,ρ) play different pure discrete strategies. This is only possible if players of type (ι,ρ) are indifferent as to which among 2 strategies to play (see corollary 1). Taking into account all possible attribute realizations, this latter condition is represented by the following linear system:

$$\begin{cases}
\sum_{\nu=1}^{m} \aleph_{1,\nu}\bar{\theta}_{1,\nu} &= \sum_{\tau=1}^{l} \aleph_{\tau,1}(1-\bar{\theta}_{\tau,1}) \ , \\
&\vdots \\
\sum_{\nu=1}^{m} \aleph_{\iota,\nu}\bar{\theta}_{\iota,\nu} &= \sum_{\tau=1}^{l} \aleph_{\tau,\rho}(1-\bar{\theta}_{\tau,\rho}) \ , \\
&\vdots \\
\sum_{\nu=1}^{m} \aleph_{l,\nu}\bar{\theta}_{l,\nu} &= \sum_{\tau=1}^{l} \aleph_{\tau,m}(1-\bar{\theta}_{\tau,m}) \ .
\end{cases} \tag{11}$$

3.4 Tenable game and its equilibria

Here, we presume that player i makes decisions under uncertainty about the priors on $\mathbf{M}$. One way to address this uncertainty is to apply Wald maxi-min principle requiring i to consider the worst-case scenario played by $-i$ [67]. The expected utility of i is then

$$\mathbb{E}_w[u_i] = \max_{s_i} \min_{\mathcal{S}_{-i}} \mathop{\mathbb{E}}_{\beth_i} [u_i] \ , \tag{12}$$

where $\beth_i$ is the distribution over $\mathbf{t}_i \times S_i$. This can be ruminated as a special case of the 'naïve' game where i calculates her best response using corollary 1 in which instead of $\mathbb{E}\big[\Pr_{\mathcal{S}}(\alpha^{(i)})\big]$ and $\mathbb{E}\big[\Pr_{\mathcal{S}}(\beta^{(i)})\big]$ she substitutes worst possible estimates $\mathbb{E}\big[\Pr_{\mathcal{S}^w}(\alpha^{(i)})\big]$ and $\mathbb{E}\big[\Pr_{\mathcal{S}^w}(\beta^{(i)})\big]$, respectively. The following assumption explains $\beth_i$.

Assumption 2. *Neither the exact type of player i nor the number of attributes that she uses is known to $-i$.*

We justify assumption 2 by referring to common settings in attribute-based authentication systems where users do not normally share information about the number of their personal attributes as well as their properties. According to eq. (12) and assumption 2 $-i$ minimizes the best utility $\mathbb{E}\big[u_i^{\mathrm{b}}\big]$ which is given by eq. (6) while $\beth_i$ reduces to the distribution over $\mathbf{t}_i$ only (which is $\aleph$). In order to calculate value of $\mathbb{E}_w\big[u_i^{\mathrm{b}}\big]$ we require $\mathcal{S}_{-i}^w$:

$$\mathcal{S}_{-i}^w = \arg\min_{\mathcal{S}_{-i}} \mathop{\mathbb{E}}_{\beth_i}\big[u_i^{\mathrm{b}}\big] \sim \arg\min_{\mathcal{S}_{-i}} \mathop{\mathbb{E}}_{\beth_i}\big[\Pr_{\mathcal{S}}(\alpha^{(i)}) + \Pr_{\mathcal{S}}(\beta^{(i)})\big]. \tag{13}$$

In order to complete our calculations for the expectation over $\aleph$ we, as previously, use notation $\theta_{\iota,\rho}$:

$$\begin{aligned} &\mathop{\mathbb{E}}_{\beth_i}\big[\Pr_{\mathcal{S}}(\alpha^{(i)}) + \Pr_{\mathcal{S}}(\beta^{(i)})\big] = \\ &\mathop{\mathbb{E}}_{\beth_i}\big[\aleph_{\iota,\rho} + \sum_{\substack{\nu=1\\ \nu\neq\rho}}^{m} \aleph_{\iota,\nu}\theta_{\iota,\nu} + \sum_{\substack{\tau=1\\ \tau\neq\iota}}^{l} \aleph_{\tau,\rho}(1-\theta_{\tau,\rho})\big] = \\ &\sum_{\iota=1}^{l}\sum_{\rho=1}^{m}\Big(\aleph_{\iota,\rho}^2 + \aleph_{\iota,\rho}\sum_{\substack{\nu=1\\ \nu\neq\rho}}^{m} \aleph_{\iota,\nu}\theta_{\iota,\nu} + \aleph_{\iota,\rho}\sum_{\substack{\tau=1\\ \tau\neq\iota}}^{l} \aleph_{\tau,\rho}\big(1-\theta_{\tau,\rho}\big)\Big) \ . \end{aligned} \tag{14}$$

In the last line of eq. (14) we ignore $\aleph_{\iota,\rho}^2$ for the calculation of

$$\arg\min_{\boldsymbol{\theta}} \sum_{\iota=1}^{l}\sum_{\rho=1}^{m} \aleph_{\iota,\rho}\Big(\sum_{\substack{\nu=1\\ \nu\neq\rho}}^{m} \aleph_{\iota,\nu}\theta_{\iota,\nu} + \sum_{\substack{\tau=1\\ \tau\neq\iota}}^{l} \aleph_{\tau,\rho}(1-\theta_{\tau,\rho})\Big) \ , \tag{15}$$

from which we conclude that for all ι, ρ:

$$\theta_{\iota,\rho} = \begin{cases} 0, & \text{if } \sum_{\substack{\nu=1\\ \nu\neq\rho}}^{m} \aleph_{\iota,\nu} \geq \sum_{\substack{\tau=1\\ \tau\neq\iota}}^{l} \aleph_{\tau,\rho} \ ; \\ 1, & \text{otherwise} \ . \end{cases} \tag{16}$$

4 EXPERIMENT

To evaluate the impact of GAPAS on privacy in attribute-based authentication we assess our game-theoretical results by conducting numerical evaluations for the system with $n \gg 2$ users.

The goal of experiment. We compare: (*i*) unlinkability in naïve game (e.g. game with mediator) with the unlinkability in tenable game (e.g. maximin); (*ii*) unlinkability in naïve and tenable games with the unlinkability in the system where users make '*alternative*' decisions. To find solutions for nonlinear systems, we run our experiment in Matlab using the trust region algorithm [19]. For the experiment, we require $\Pr(i)$, $\aleph$. Based on eq. (9) we derive best response expressions that are identical among players i whose types $\mathbf{t}_i = \{\alpha^{(i)}, \beta^{(i)}\}$ match. As such, we further use $\theta_{\iota,\rho} = s_i$ for all i whose $\mathbf{t}_i$ realization is $(\alpha_\iota, \beta_\rho)$. We then define the systems of equations for equilibria in *naïve* as well as *tenable* game settings.

4.1 Experiment organization

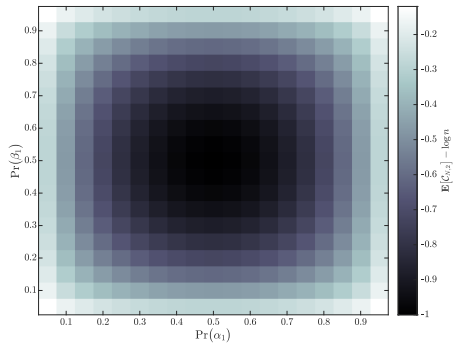
For our baseline scenarios, considerations are made for 'unrestricted rationality' where 2 attribute realizations $\{\alpha^{(i)}, \beta^{(i)}\}$ available to player i can be used interchangeably in naïve and tenable games (see section 3). We also analyze some of alternative scenarios with different kinds of '*irrationality*'. While the discussion of many possible alternative decisions goes beyond the scope of our paper we identify: (*a*) 'restricted rationality' where users play naïve or tenable game but (in contrast to interchangeable usage) select and always use the same realization out of 2 realizations available to them (see section 3.3.2); (*b*) 'random move' scenario where users use both of their realizations interchangeably but in random manner, $\forall i,\ \varrho(s_i) = 1,\ s_i \in [0, 1]$. We use compact notation for the unlinkability which is obtained in different scenarios. Expected unlinkability $\mathbb{E}[C] = \log n + \sum_i \mathbb{E}[u_i]$ in rational scenarios is denoted by $\mathbb{E}[C_{\kappa,\mu}]$ where $\kappa \in \{N, T\}$ denotes either naïve (letter 'N') or tenable (letter 'T') game, respectively. $\mu \in \{1, 2\}$ indicates the number of attribute realizations used by each player: $\mu = 1$ specifies games with restricted rationality; $\mu = 2$ specifies games with unrestricted rationality. Notation $\mathbb{E}\left[C_{\{\kappa,\mu\}^r}\right]$ is for expected unlinkability measured under random moves scenario (index 'r').

In order to produce *outputs* in the form of expected unlinkability, our experiment requires the following *inputs*: 1) $\{\kappa, \mu\}$ or $\{\kappa, \mu\}^r$; and 2) $\Pr(i)$, for all players i and the **pmf** $\aleph$. For all the instances of experiment, we consider n users and $\Pr(i) = \frac{1}{n}$ for all i. We aim at conducting numerical evaluations for a wide range of various joint **pmf**s $\aleph$. For the purpose of convenient presentation and comparison of the outputs from the experiment we depict corresponding unlinkability using two-dimensional heat maps (see Figures 6-8). Coordinates $\big(\Pr(\alpha_1), \Pr(\beta_1)\big)$ of each point on the map define a corresponding 2×2 matrix $\aleph$: $\aleph = \big[\Pr(\alpha_1),\ 1 - \Pr(\alpha_1)\big]^{\mathrm{T}} \times \big[\Pr(\beta_1),\ 1 - \Pr(\beta_1)\big]$ where both $\Pr(\alpha_1)$, $\Pr(\beta_1)$ were quantized with 0.05 step on interval $[0,\ 1]$. Color intensity corresponds to unlinkability.

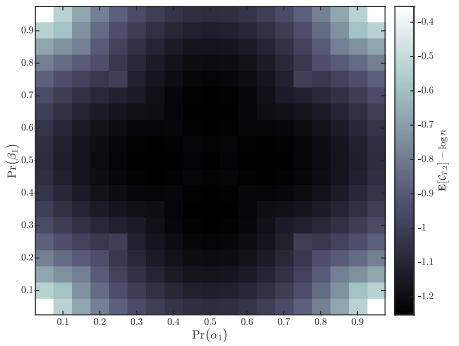
4.2 Results

We first calculated the equilibria for our baseline scenarios on the naïve and tenable games where players can use both of their attribute realizations interchangeably (see fig. 6). For each possible $\aleph$ in naïve game we solved complete information Nash equilibria (see eq. (10)) to find $\mathbf{M}$ that need to be communicated to the players by mediator. Among all the possible solutions we selected those maximizing $\mathbb{E}[C_{N,2}]$. For each possible $\aleph$ in tenable game we calculated worst case condition that may be created for player i by others $n - 1$ players (see eq. (16)). Then, best response of i, and $\mathbb{E}[C_{T,2}]$ are calculated (see eq. (4)). As can be observed from comparison of fig. 6a and fig. 6b naïve game provides substantially better unlinkability.

To compute equilibria for naïve games with single attribute usage (e.g. restricted rationality) we solved a linear system representing mixed and pure discrete equilibria (see eq. (11)). The benefits of using 2 attributes (unconstrained rationality) versus 1 attribute (constrained rationality) can



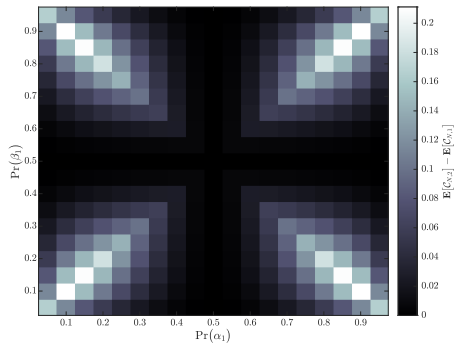
(a) Naïve game, players use 2 attributes interchangeably, e.g. $\theta_{\iota,\rho} \in [0,1]$.



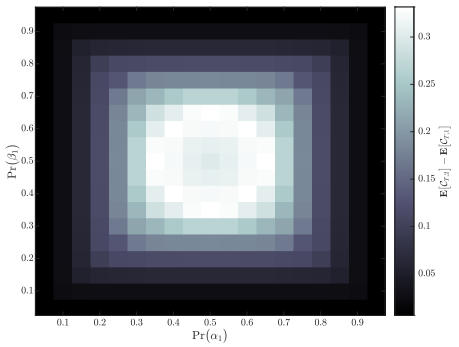
(b) Tenable game, players use 2 attributes interchangeably, e.g. $\theta_{\iota,\rho} \in [0,1]$

Fig. 6. Comparison of expected unlinkability in naïve and tenable baseline scenarios.

be observed by comparing residual unlinkabilities on fig. 7 which are greater than 0 for the both heatmaps.



(a) Difference between expected unlinkability $\mathbb{E}\left[C_{N,2}\right]$ and $\mathbb{E}\left[C_{N,1}\right]$.

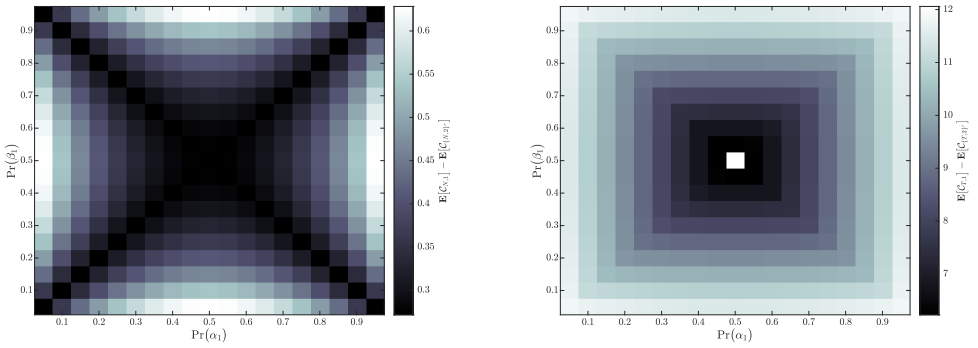


(b) Difference between expected unlinkability $\mathbb{E}\left[C_{T,2}\right]$ and $\mathbb{E}\left[C_{T,1}\right]$.

Fig. 7. Residual expected unlinkability in 'Naïve game' and 'Tenable' games.

We conducted a range of experiments with randomized moves which results are presented on fig. 8. For the 2-attribute randomized game, each player i decides $0 \leq s_i \leq 1$ at random in accordance to uniform distribution on $[0,1]$. As can be seen from the residuals of expected unlinkabilities, even constrained rationality (1 attribute usage) scenario outperforms scenario where 2 realizations are used randomly (chaotically).

As a special case, we analyzed example described in section 2.3 where $\aleph = \left(\begin{smallmatrix} 0.106 & 0.017 \\ 0.759 & 0.118 \end{smallmatrix}\right)$. We compared unlinkability in the system where users use 1 or 2 attributes in two variants of the game as well as randomized scenario (see fig. 9). Few important observations can be made in that regard. First, unlinkability achievable in *tenable* game is better than in randomized scenario, but it is worse than *naïve* game performance. Second, differences between relative performance $\frac{\mathbb{E}[C]}{\log n}$ for all these variants of the game change with n: for smaller n this becomes even more apparent.



(a) Difference between expected unlinkability $\mathbb{E}\left[C_{N,1}\right]$ and $\mathbb{E}\left[C_{\{N,2\}^r}\right]$.

(b) Difference between expected unlinkability $\mathbb{E}\left[C_{T,1}\right]$ and $\mathbb{E}\left[C_{\{T,2\}^r}\right]$.

Fig. 8. Residual expected unlinkability in 'Naïve' and 'Tenable' games.

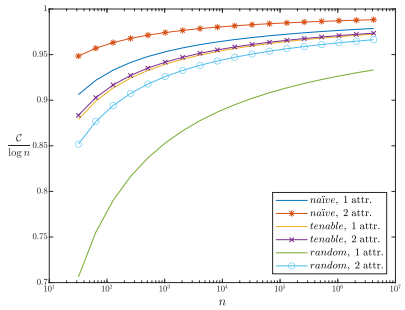


Fig. 9. Unlinkability for the example in section 2.3 under various games

5 INTERCHANGEABLE USAGE OF VERIFIABLE CREDENTIALS

In this section, we further analyze details of practical implementation for **Example #2** where Verifiable Credentials (VCs) are used as a format for assertions. Due to multiple benefits (including advanced privacy preserving tools) VC format becomes widely appreciated [15, 32, 68]. We demonstrate *why* interchangeable usage of assertions is plausible for the systems where VC format is accepted. In addition, we suggest an algorithm that governs usage of these VC-compatible assertions within Digital Credential Wallets (DCW).

5.1 VCs and Verifiable Presentations

Verifiable Credentials (VC) are a standard created by W3C. It allows a trusted issuer to issue tamper resistant statements (credentials and assertions) about attributes of the entities known to the issuer. Validity of assertions that comply with VC requirements can be unambiguously verified (e.g. RP). The primary utility of VC is that it expedites privacy which contrasts with current practices of Identity Providers (IdP) in federated Identity Management systems. This privacy preserving features of VC are supposed to be enabled by the holders of the credentials which aligns with user-centric paradigm: they decide upon Personally Identifiable Information (PII) and control its disclosure in independent manner. This control is, however, not arbitrary since resulting assertions must satisfy

RP's trust requirements as discussed in section 1.2. Improved security of identity management is among other advantages of user-centric paradigm: it mitigates the risk pertaining to a single point of failure (which is common for IdPs).

In VC-compatible assertions, control over information that is disclosed to RP can be exercised through Zero Knowledge Proofs (ZKP), which enables a legitimate holder to produce *Verifiable Presentation* (VP). The benefits of such approach can facilitate anonymous authentication for the addressee (e.g. future VC holder) the issuer can sign VC using, for example, Camenisch-Lysyanskaya (CL) or Boneh-Boyen signatures (BBS) [9, 12]. This allows holder (e.g. user) to produce multiple VPs where he/she selectively discloses the information about the attributes in the VC. The signatures produced by the holder for different VPs do not correlate which has been widely advocated by the community as the main privacy preserving feature. This often creates a ground for the narrative that RP can not link assertions from the same user if ZKP is used.

Unfortunately, benefits of anti-correlation properties of ZKP system are undermined by Attribute Value Metadata (AVM) and Attribute Schema Metadata (ASM) which are present in VP [34]. This causes distinguishability: interchangeable usage of attributes should therefore be practiced to avoid adverse effects on users' unlinkability in VC-compatible attribute-based authentication systems. To see why, let us analyze the case depicted on fig. 10 (see appendix B for VC/VP details).

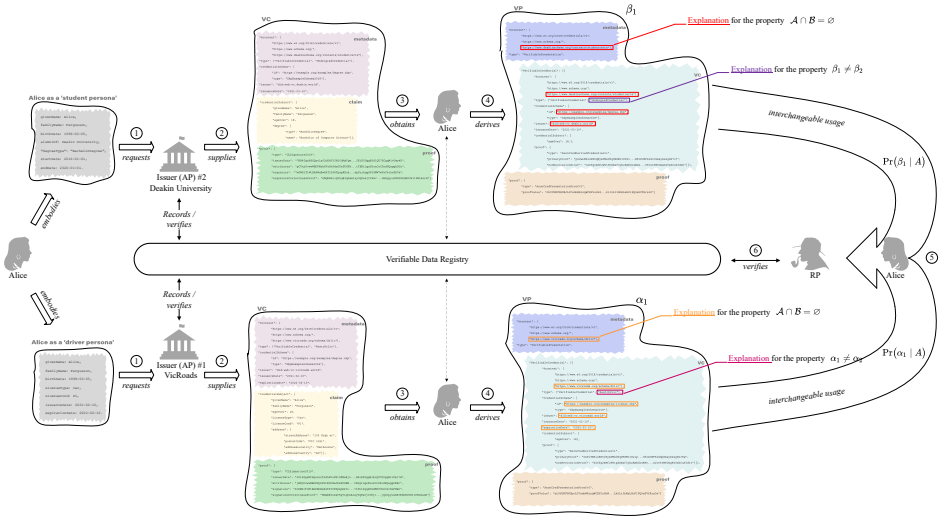


Fig. 10. Application of the proposed model for VC usage scenario

Taking into account the flow in (AP, RP, U)-system where user *Alice* obtains her credentials from APs, and RP grants her access to the products (e.g. allows to buy alcohol online) based on the policy P requiring that any user must be *'older than 18'*. In this specific scenario, RP accepts (e.g. trusts the issuers of) the digital driver licenses (issued by VicRoads, Australia) and digital tertiary qualification certificates (issued by Australian universities). *Alice* embodies different personae which allows her to claim credentials from two different issuers (see fig. 10). At ①, she requests VCs from AP #1 and AP #2. All the aspects of user authentication demanded by APs, and consider that these steps are successfully executed are omited to ensure simplicty. This eventuates in ② during which VCs are supplied to *Alice*. For better privacy, VCs can remain completely encrypted during that transmission step. This can be accomplished by using, for instance, JSON Web Encryption (JWE) file format [46]. In order to support verifiability of VCs, information about public keys of both APs

Table 3. Main objects and properties of VC and VP

Data concept	Object/property	Purpose	Presence	Category	Format example
VC, *meta*	`@context`	maps aliases to Uniform Resource Identifiers (URI)	compulsory	static	ordered set of URIs [8]
VC, *meta*	`id`	unambiguously refers to the credential	optional	static/dynamic	single URI (including DID) [8, 33]
VC, *meta*	`type`	defines type for objects within VC	compulsory	static	terms defined according to JSON-LD grammar [50]
VC, *claim*	`credentialSubject`	specifies the subject of the claim	compulsory	static/dynamic	a set of objects with properties related to the subject
VC, *meta*	`issuer`	specifies the issuer of VC	compulsory	static	URI or other id-object such as JWK or DID [4, 45]
VC, *meta*	`issuanceDate`	expresses the date and time when a VC becomes valid	compulsory	dynamic	combination of date and time strings, RFC3339 [51]
VC, *proof*	`proof`	issuer's assertion about information in VC	optional/compulsory	dynamic	RSA digital signatures, such as RsaSignature2018 [55]
VP, *meta*	`id`	unambiguously refers to the presentation	optional	static/dynamic	single URI (including DID) [8, 33]
VP, *meta*	`type`	defines type for objects within VP	compulsory	static	terms defined according to JSON-LD grammar [50]
VP, *meta*	`verifiableCredential`	construction of one or more VCs or derived VCs	optional	dynamic	VC data model v1.0 [68]
VP, *meta*	`holder`	specifies the entity that is generating VP	optional	static/dynamic	URI or other id-object such as JWK or DID [4, 45]
VP, *proof*	`proof`	authenticates VP holder to the verifier	optional/compulsory	dynamic	RSA digital signatures, such as RsaSignature2018 [55]

must be recorded in a publicly accessible medium. The function of such medium is performed by *Verifiable Data Registry* which can be implemented using Decentralized PKI (DPKI), for example [53].

At ③, *Alice* obtains VCs and decrypts them if necessary. Major components within VC include *metadata*, *claims*, and *proof* (see table 3). Properties of these components explain why assertions from different APs may differ. A claim is a statement about a subject, and a synonymous term *'subject attribute'* can be used as well. These statements are composed using attribute-value pairs: they are placed within `credentialSubject` and define its properties. In spite of being the most informative part of VC, claims (e.g. `credentialSubject`) may be insufficient for unambiguous interpretation which is required by RP. Therefore, metadata (*meta* for short) carries functions of interpretation and validation of attribute-value pairs within `credentialSubject` [68]. As such, functions of meta in VC can be explained through the combination of AVM and ASM [34]. The purpose of the proof is to assert claims and meta in a way which is unique for the specific issuer (e.g. to uniquely authenticate AP). This is required to establish trust to the information in VC for both *Alice* and RP (since they already trust AP).

It can be observed (③, fig. 10) that multiple objects within *meta* and `credentialSubject` differ when VCs from VicRoads and Deakin University are compared. For example, for the 2 VCs obtained by *Alice* at least one object within `@context` property differ: credential for driver license contains "`https://.../drlic`" while credential for tertiary qualification contains "`https://.../studentcerts`". Each of the other properties within metadata including `type`, `credentialSchema`, `issuer`, and `issuanceDate` also differ for those 2 VCs. It is remarkable that driver license VC contains `expirationDate` property while tertiary qualification certificate does not. Differences between those 2 VCs become even more apparent if contents of `credentialSubject` are compared.

To improve her privacy in attribute-based authentication system *Alice* eliminates from her assertions as much of PII as possible. For this she first modifies existing VC, $\mathrm{VC} \Rightarrow \dot{\mathrm{V}}\mathrm{C}$. This modification is due to the changes in original claims cl of that VC: *Alice* removes redundant information from cl, $cl \Rightarrow \dot{cl}$. Since structural parts of original VC change, a new proof $\dot{pr}\left(\dot{cl}, \dot{mt}\right)$ must be produced for $\dot{\mathrm{V}}\mathrm{C}$ where meta remains unchanged, e.g. $\dot{mt} = mt$. This modified $\dot{\mathrm{V}}\mathrm{C}$ then becomes a part of $\ddot{\mathrm{V}}\mathrm{P}$. In order to compose a valid $\ddot{\mathrm{V}}\mathrm{P}$ it is also required to add $\ddot{mt}$ and to produce $\ddot{pr}\left(\ddot{mt}, \dot{\mathrm{V}}\mathrm{C}\right)$ (see ④, fig. 10):

$$\ddot{\mathrm{V}}\mathrm{P} = \left[\ddot{pr}\left(\ddot{mt}, \underbrace{\left[\dot{pr}\left(\dot{cl}, \dot{mt}\right), \dot{cl}, \dot{mt}\right]}_{\dot{\mathrm{V}}\mathrm{C}}\right),\ \ddot{mt}, \dot{\mathrm{V}}\mathrm{C}\right].$$

As per the derivation procedure, $\dot{cl}$ becomes the only part of the resulting assertion where attributes of the subject are stated explicitly. The rule $\dot{cl} \subseteq cl$ must be obeyed during the derivation procedure.

For instance, it can be seen that for both VPs derived by *Alice* that she has $\dot{cl} \leftarrow \{$"ageOver" : 18$\}$. This conforms with the derivation rule and greatly reduces PII which makes claims alone indistinguishable when driver license VP is compared with tertiary qualification VP. Nevertheless, the rest of the both VPs contain enough information to clearly differentiate assertions possessed by *Alice*.

This is because ZKP and selective disclosure can be performed by the holder over claims only (e.g. credentialSubject). As a result, *Alice* assertions that originate from VicRoads and Deakin University will always differ: this supports property $\mathcal{A} \cap \mathcal{B} = \varnothing$ in GAPAS. In addition, assertions that originate from AP #1 may differ for *Alice* and other users. This is due to the meta property type containing object "RestrDrLic" which is used only with restricted driver licenses. In contrast, for unrestricted driver licenses type will contain "UnrestrDrLic" (not displayed on fig. 10) instead of "RestrDrLic" (for the details see **Example #2**).

This observation supports $\alpha_1 \neq \alpha_2$ which is important for our decision making model. Also, assertions produced from restricted and unrestricted driver licenses belong to the same category $\mathcal{A}$ because they are issued by AP #1, but *Alice* can not hold both valid α_1 and α_2. This also supports our assumption for joint distribution of assertions among players which is defined by matrix $\aleph$. Similarly, we observe that VP produced by *Alice* from tertiary qualification certificate contains UndergradCredential as part of the type. There are other users who possess credentials from AP #2 and whose degree is postgraduate. As a result, VPs of those users will have type which is different (not shown on fig. 10) to *Alice's* type. This is because those users' types contain PostgradCredential in contrast to UndergradCredential. This conforms with $\beta_1 \neq \beta_2$ while both β_1, β_2 belong to the same category $\mathcal{B}$. *Alice* can have only one assertion from $\mathcal{B}$.

Based on the discussed properties of assertions we can uniquely specify *Alice* type in the game (not to be confused with type in VC/VP) as $\mathbf{t} = (\alpha_1, \beta_1)$ (see fig. 10). Depending on the information that is available to *Alice* about other players she can play either *naïve* or *tenable* variant of the game. This implies that during the sessions when she authenticates to RP she will use her assertions α_1 and β_1 *interchangeably* according to probabilities $\Pr(\alpha_1 \mid A)$ and $\Pr(\beta_1 \mid A)$, respectively (see ⑤, fig. 10). Finally, in each of authentication sessions RP can verify whether assertion submitted by *Alice* is valid (see ⑥, fig. 10).

5.2 Interchangeable usage of assertions in DCW

VCs and VPs can be held within a piece of software or a hardware device - also known as a Digital Credential Wallet (DCW) [64, 70]. Although different types of DCW's exist, the majority of them satisfy basic requirements related to the task of entity authentication. This includes: (*i*) receiving and securely storing credentials (also includes requesting a credential from AP in some cases); and (*ii*) selective disclose of credential information, e.g. VC $\rightarrow$ VP. This implies that stages ③ and ④ on the diagram (see fig. 10) comprise common functionalities of DCWs which provides a user-friendly yet secure way of authentication through Verifiable Credentials [64].

With the aim to improve user unlinkability while maintaining ease of use for VCs we propose to also incorporate *interchangeable usage* of assertions which corresponds to stage ⑤ (see fig. 10) into the design of DCW. For example, best responses for the player in *tenable* variant of the game can be governed by algorithm 1 which is derived from corresponding equilibrium conditions (see corollary 1 and eq. (16)). It is based on maximin principle and does not require any additional information except $\aleph$. From fig. 10, *Alice* with type $\mathbf{t} = (\alpha_1, \beta_1)$ uses matrix $\aleph = \left(\begin{smallmatrix} 0.106 & 0.017 \\ 0.759 & 0.118 \end{smallmatrix}\right)$ (see **Example #2** in section 2.3). According to algorithm 1 she calculates $\theta(1,1) = 1$, $\theta(1,2) = 1$, and $\theta(2,1) = 0$, from which she obtains $s = \dfrac{\big(\theta(1,1)\ \ \theta(1,2)\big) \times \big(0.106\ \ 0.017\big)^{\mathrm{T}}}{\big(\theta(1,1)\ \ \theta(1,2)\big) \times \big(0.106\ \ 0.017\big)^{\mathrm{T}} + \big(1-\theta(1,1)\ \ 1-\theta(2,1)\big) \times \big(0.106\ \ 0.759\big)^{\mathrm{T}}} = \frac{0.123}{0.123+0.759} \approx 0.139$. To ensure that for each of authentication sessions assertion (VP) α_1 is selected randomly with

probability $\Pr(\alpha_1 \mid A) = 0.139$, and β_1 is selected randomly with probability $\Pr(\beta_1 \mid A) = 0.861$ user *Alice* runs uniform random generator with support on $[0, 1]$. If random number s^* from the generator is less or equal to s *Alice* authenticates to RP with the assertion derived from her driver license (e.g. α_1). Otherwise, she authenticates with the assertion derived from her tertiary degree certificate (e.g. β_1). If proposed algorithm is implemented in DCW all the mentioned decisions will be made by the software which does not require *Alice* participation.

Algorithm 1: Maximin decision algorithm

input : $\aleph$, $\{\alpha_\iota, \beta_\rho\}$
output: $\mathrm{VP}(\alpha_\iota)$ or $\mathrm{VP}(\beta_\rho)$

begin

$\mathbf{h}(\iota, \cdot) \leftarrow \left\{\bigcup_{\nu=1}^{m} \aleph_{\iota,\nu}\right\}$, $\mathbf{h}(\cdot, \rho) \leftarrow \left\{\bigcup_{\tau=1}^{l} \aleph_{\tau,\rho}\right\}$;
$\boldsymbol{\theta}(\iota, \cdot) \leftarrow \{0\}_m$, $\boldsymbol{\theta}(\cdot, \rho) \leftarrow \{0\}_l$, $s \leftarrow 0$, $s^* \leftarrow 0$;
for $\nu \leftarrow 1$ **to** m **do**
  if $\sum_{\substack{\psi=1 \\ \psi \neq \nu}}^{m} \aleph_{\iota,\psi} \geq \sum_{\substack{\gamma=1 \\ \gamma \neq \iota}}^{l} \aleph_{\gamma,\nu}$ **then** $\theta(\iota, \nu) \leftarrow 0$;
  else $\theta(\iota, \nu) \leftarrow 1$;
for $\tau \leftarrow 1$ **to** l **do**
  if $\sum_{\substack{\psi=1 \\ \psi \neq \rho}}^{m} \aleph_{\tau,\psi} \geq \sum_{\substack{\gamma=1 \\ \gamma \neq \tau}}^{l} \aleph_{\gamma,\rho}$ **then** $\theta(\tau, \rho) \leftarrow 0$;
  else $\theta(\tau, \rho) \leftarrow 1$;
$s \leftarrow \frac{\boldsymbol{\theta}(\iota,\cdot) \times \mathbf{h}(\iota,\cdot)^{\mathrm{T}}}{\boldsymbol{\theta}(\iota,\cdot) \times \mathbf{h}(\iota,\cdot)^{\mathrm{T}} + \left(\{1\}_l - \boldsymbol{\theta}(\cdot,\rho)\right) \times \mathbf{h}(\cdot,\rho)^{\mathrm{T}}}$;
$s^* \leftarrow UniRand([0, 1])$;
if $s^* \leq s$ **then** *output* $\mathrm{VP}(\alpha_\iota)$;
else *output* $\mathrm{VP}(\beta_\rho)$;

6 DISCUSSION

Metadata is often a part of a credential from which assertion is obtained: it is important for RP and can not be altered by a user. Our work contrasts with the cryptographical approaches that aim to achieve unlinkability through indistinguishability. We **advocate** that unlinkability can also be substantially improved if assertions are revealed to RP by the users in coordinated manner. The novelty of GAPAS is due to *interchangeable and coordinated* usage of assertions. This interchangeable usage is often possible because many RPs have flexible access policies.

6.1 Importance for theory

From a theoretical perspective, we defined criterion C from which users derive their utilities in non-cooperative coordination games. This development was motivated by our analysis of existing criteria and measures for user unlinkability during which we observed a substantial gap. For example, test for RP+AP-U-unlinkability described in ISO/IEC 27551 (see listing 1) does not specify

how users select attributes for authentication if multiple attribute realizations (e.g. α_ι and β_ρ) are available [43].

In addition, it neither specifies *how* RP links users nor *what* is the estimate of the rate of successful linking. The proposed criterion C is based on the information-theoretical measure of conditional entropy. The main advantage of C is that it can be used to estimate best linking performance of malicious RP (see lemma 1). To explain linking procedure, on fig. 3a we assumed that certain statistics about user decisions must be known to calculate $C = H(\mathrm{L} \mid l)$. However, the results of lemma 1 do not depend on *whether* this statistics is known to RP and *how* he can use it: detailed discussions about properties of RP's linking detector are out of the scope of this paper and we only utilize the upper estimate C of its linking ability.

lemma 2 further demonstrates how personal expected utilities of the players can be derived from C as well as the assumptions needed for that. To calculate their optimal strategies (e.g. best responses) a player i who controls their personal assertions $\alpha^{(i)}$ and $\beta^{(i)}$ (e.g. whose type is $\mathbf{t}_i = (\alpha^{(i)}, \beta^{(i)})$) must know marginal probabilities at RP, $\Pr_{\mathcal{S}}(\alpha^{(i)})$, $\Pr_{\mathcal{S}}(\beta^{(i)})$ (see fig. 5). These marginal probabilities depend on statistical distribution $\aleph$ of types of other players in the system as well as their best responses. Inter-dependencies of personal best responses for multiple players explain why unlinkability in attribute based authentication systems needs to be addressed using GAPAS. The way of *how* information about $\Pr_{\mathcal{S}}(\alpha^{(i)})$, $\Pr_{\mathcal{S}}(\beta^{(i)})$ is communicated to i is of paramount importance to the game and its equilibria. As such, this '*how to communicate information*' question is a stepping- stone to answer the **R**esearch **Q**uestion

"How should users use their assertions to maximize RP+AP-U-unlinkability?"

GAPAS provides the answer to RQ. To accommodate likely scenarios, two types of non-cooperative coordination games with incomplete information were analysed: (*i*) *naïve* game where mediator M provides necessary information to the players; and (*ii*) *tenable* game where no information about $\Pr_{\mathcal{S}}(\alpha^{(i)})$, $\Pr_{\mathcal{S}}(\beta^{(i)})$ is available. To the best of our knowledge, we are the first to apply game-theoretical approaches to the task of interchangeable attribute usage. For the naïve game, we applied principles of correlated equilibria [6, 49]. This allows us to achieve consistency of priors (e.g. information about $\Pr_{\mathcal{S}}(\alpha^{(i)})$, $\Pr_{\mathcal{S}}(\beta^{(i)})$).

In addition, multiple equilibria are possible in the system (see eq. (10): by selecting this information we maximize the overall unlinkability C for the system of n players. One of the limitations of this approach is that information provided by mediator M needs to be trusted by the players. This is avoided in tenable game where we apply Wald maximin principle to maximize utility of each player i. That player calculates 'worst case scenario' values $\Pr_{\mathcal{S}}(\alpha^{(i)})$, $\Pr_{\mathcal{S}}(\beta^{(i)})$ based on $\aleph$ [67]. This worst case scenario guarantees that utility of i can not be lower even if other $n-1$ players attempt to minimize expected utility (over $\aleph$) of i whose type is unknown to them. Such approach is widely used in robust decision making and is suitable for the situations that are characterized by severe uncertainty such as the situation when a player i does not know $\Pr_{\mathcal{S}}(\alpha^{(i)})$, $\Pr_{\mathcal{S}}(\beta^{(i)})$. On the downside of this approach is that utility of i is lower than in the case with correlated equilibrium.

From section 4 we observe that playing tenable game comes at the cost (see fig. 6) [67]. Also, there is a clear contrast between unlinkability when users are guided by different principles of assertion usage. Both naïve and tenable games belong to *rational* principle of assertion usage. In section 4 we compare them with *alternative* principles of usage (see figs. 7 and 8). From the results it is clear that rational principles of interchangeable usage where users *coordinate* have substantial benefit over other alternative scenarios. This is because GAPAS optimizes impact on unlinkability (through best responses) produced by every individual user i.

We also gain insights into how the distribution $\aleph$ of attribute realizations (and, hence, user types $\mathbf{t}_i$) used by the users impacts their unlinkability. From **Example #0** we learned that indistinguishability

may be insufficient for unlinkability. On the other hand, better indistinguishability (e.g. larger anonymity sets) can provide better unlinkability. This is because the task of coordination in GAPAS becomes easier if user assertions belong to larger anonymity sets. The size of anonymity sets can be calculated from $\aleph$, number of users n in the system, and their decisions $\mathcal{S}$. For instance, it can be seen that for naïve and tenable games expected unlinkabilities are lower towards the center of corresponding heatmaps on figs. 6a and 6b. This is because that area represents more diverse distributions $\aleph$ (more equally-sized anonymity sets) which further constrains coordination effect. In contrast, outer areas of these maps represent the cases when majority of the players have the same type.

6.2 Importance for practice

The results obtained in this paper also have important implications for practice.

First, our study contributes by using indirect Security Assurance Conformity Assessment (SACA) method which is based on information theory and game theory [41, 42] to develop a security assurance argument. More specifically, unlinkability definitions provided in ISO 27551 lack clarity as it neither stipulates on *how* users need to use their assertions (if a user has more than 2 of them) during the test nor *how* RP makes his *'guess'*. Criterion C coupled with naïve or tenable games can now be used to address this limitation. This is because rational decision making unambiguously specifies actions of the users the resulting unlinkability can now be rated. A certain threshold T_C can be used to decide whether attribute-based authentication system conforms to a standard or not.

Second, recommendations are provided for the issuance of credentials as well as for setting of access policies by an RP. As we have observed in section 5 metadata can become the reason for distinguishability of assertions. Therefore, any unnecessary detalization or redundancy should be avoided by AP who issues credentials. For example, content of the properties `@context`, `type` within VC should be kept as minimal as possible. Across all APs it is better to use the same `credentialSchema`. If it is not critical for LoA of credentials or internal protocols and procedures of AP, time must not appear in `issuanceDate` property – e.g. it should always be in the format `YYYY-MM-DD`. This would ensure that larger number of users have identical metadata. Role of of RP is also important: he defines access policies P which also affects user types $\mathbf{t}_i$ in attribute based authentication system. For example, if for the use case in section 5 RP only allows assertions from AP VicRoads (and does not allow assertions from AP Deakin University) the distribution of user types in the system will become $\aleph = \begin{pmatrix}0.122 & 0.878\end{pmatrix}$. This will affect decisions that can be made by the users, equilibria, and resulting unlinkability C. Hence, GAPAS can be used to recommend (or 'benchmark') policies. Privacy-respecting RPs can use this benchmarking to the greater benefit of the users. On the other hand malicious RPs may deliberately constrain access policies with the aim to make linking easier (e.g. to reduce unlinkability). Therefore, additional mechanisms that may incentivize or penalize RP for such malicious behavior are yet to be further studied [52].

Finally, GAPAS can assist firms that are developing various identity agents and DCW solutions. For example, software implementation of algorithm 1 would not require changes in authentication protocols and procedures such as OAuth 2.0 and OIDC. In addition, communication and computation overheads associated with GAPAS are minimal. To make a decision in a naïve game, user's DCW that incorporates GAPAS would require information set from mediator M. Such set is provided once only, its size equals the number of different realizations in the system, e.g. $|\mathcal{A}| + |\mathcal{B}|$. A tenable game does not require communication. Also, only naïve game requires that equilibrium is computed by M. This needs to be done once only: obtained information set (e.g. $\Pr(\alpha^{(i)})$ and $\Pr(\beta^{(i)})$) is communicated to every user i. To calculate that set M solves eq. (10). This can be done by the trust-region algorithm [19]. In both naïve and tenable variants of the game, each player produces the best response, which is a single operation (see corollary 1). However, tenable game calculates

'worst-case scenario' (see eq. (16) and algorithm 1). For every user, the total complexity of that procedure is $x(l+m)$, where x is the number of different realizations controlled by a user, $l \times m$ is the dimensionality of matrix $\aleph$. In both naïve and tenable games each player produces the best response only once: the decision is then applied across all authentication sessions.

7 RELATED WORK

Here we provide a brief analysis of sources contributing to the questions of privacy, unlinkability and anonymity. First, we outline works that rather formalize the above mentioned definitions. Second, we deliberate upon game-theoretical approaches that aim at improving some of these characteristics through non-cooperative interactions.

7.1 Definitions of Unlinkability

In common privacy context unlinkability is strongly related to anonymity. For example, [62] states " ...Unlinkability is a sufficient condition of anonymity, but it is not a necessary condition." while [43] equals anonymity to RP+AP-U-unlinkability.

Some definitions of unlinkability are called 'games' [43, 58, 60, 69]. Such selection of terminology seems unjustified: these constructs are barely related to game theory since observability of the strategies played by the players as well as their payoffs remain unclear. In order to avoid confusion with game theory we will further call them 'tests'. One common idea behind these tests is to define unlinkability as the result of an interaction between an attacker (whose goal is to distinguish between the actions of different agents) and a challenger. Depending on the variant of the test, either 2 or 3 agents are selected by the challenger. He then presents to an attacker results of the sessions with explicit assurance that either 1 or 2 among these agents were selected, respectively. The attacker then needs to guess the label of the entity, or the relation between the labels (e.g. 'same' or 'different'), respectively. Unlinkability is achieved if the attacker's performance is not statistically better than a random guess.

Models based on logical description for unlinkability have also been used by the community [10, 35, 56]. For example, in [35] authors propose a framework for reasoning about anonymity in particular. Their framework employs the modal logic of knowledge within the context of the runs and the systems framework but does not consider quantitative measurements of anonymity. Authors of [10] abstract model based on epistemic logic, with natural and intuitive definitions in terms of the attacker's knowledge. In addition to unlinkability they also identified a dual notion of inseparability which may be of importance for some special cases on practice. However, majority of the authors in this category do not examine probabilistic descriptions of unlinkability. This sets forth impossibility of a single (or integral) measure for unlinkability. As a result, many realistic scenarios with complex usage patterns can not be encompassed by these descriptions for the sake of further optimization, for instance.

Works [7, 28, 69] quantify the linkability of items in the system using information-theoretic descriptions. For example, a basic information-theoretic notion for unlinkability is given in [69] where authors utilize Shannon entropy to measure unlinkability of elements within one set as well as between the sets. One limitation of this approach is that no context is considered by the authors. This was rectified in [28] where 7 special cases of context information where considered. Nevertheless, this information is provided in the form of 'hints' (known to the attacker) about target partitions, and can not be easily generalized for all kinds of context. Authors in [7] measure unlinkability using mutual information between the actual communication that took place, and the information the adversary knows about it. In that way their definition obviously considers context information in the most general way. On the other hand, mutual information may be heavily

affected by priors which is impractical for decisions making in non-cooperative environment where players may not communicate.

Finally, a number of papers survey criteria and measures that may be useful in a wide range of privacy applications [18, 23, 71]. For example survey [71] spans across multiple privacy domains and can serve as a general framework for privacy measurements. In particular, the authors propose an extensive taxonomy of privacy metrics which is classified by output and describes 17 entropy-based measures, to name a few.

Summary: one of the main limitations of the analyzed sources is the lack of attention to the problem of interchangeable usage of assertions. Some of the information-theoretic measures such as in [71] are universal. However, possible application of these measures to the problem of interchangeable usage is not suggested by the authors. Existing definitions are therefore insufficient to optimize unlinkability in the environment where multiple assertions of a user can satisfy access policy P that is established by RP.

7.2 Game Theoretical Approaches

Game theory studies incentives and interactions between rational agents [30]. In some cases these interactions can be characterized by a stable state of Nash equilibrium where agents do not deviate from the strategies they play. These outcomes are also considered as the most likely and hence users should evaluate their anticipated privacy level with an equilibrium in mind.

A number of papers apply game theory to address privacy issues that occur on practice [29, 39, 54]. It should, however, be taken into account that their solutions are usually problem-specific and can not be easily applied across multiple domains. For example, problems of pseudonym change in mobile networks were investigated by the authors of [29, 39]. In [29], the authors elaborate on user-centric location privacy model which takes into account the beliefs of users about the tracking power of the adversary, the degree of anonymity that users obtain in the mix zones as well as the cost and time of pseudonym change. Results from their study define an equilibrium where the strategies played by the users can be decided when their utilities are compared with a threshold value. In [39] authors analyze a game where local adversary is equipped with multiple eavesdropping stations to track mobile users who deploy mix zones in order to protect their location privacy. The authors predict the strategies of both players and derive the strategies at equilibrium in complete and incomplete information scenarios which is quantified based on real road-traffic information.

On the other hand, there is a number of game-theoretic papers with less distinct contribution to practical aspects of the privacy. They nevertheless may be useful for coordination scenarios in attribute based authentication [31, 49]. For instance, authors of [49] discuss a game with mediating mechanism that can improve the outcome of the game when compared to Bayes Nash Equilibrium (BNE). Authors demonstrate that any algorithm that computes a correlated equilibrium of a complete information game while satisfying a variant of differential privacy can be used as a recommended mechanism satisfying desired incentive properties.

Summary: an obvious limitation of the existing game-theoretical solutions is the absence of models that adequately cover interchangeable usage of assertions. Properties of information-theoretical measures command that games with continuous strategies (and not mixed strategies!) must be analyzed in the presence of multiple alternatives for the players. This component is missing from game-theoretical applications for privacy. Also, majority of the sources gravitate toward the games where information sets can be provided to the users. As such they ignore cases of severe uncertainty. There are several limitations for this sole line of thoughts. First, a mechanism that provides information to the players (similar to *mediator* in 'naïve game') must be designed. Second, players must place trust on that mechanism.

8 CONCLUSION

Attribute-based authentication offers a number of benefits for a user such as user-centric paradigm, effortless onboarding into online services, account-less money transfers, to name a few [14, 44, 59]. Due to growing number of high-LoA assertions RPs also start to appreciate advantages of attribute-based authentication and eagerly migrate to corresponding access platforms [66, 72].

However, the benefits of attribute based authentication come come with limitations. The privacy of interchangeable usage of assertions is a new problem which appears exclusively in attribute-based authentication and access control. This is because attribute-based assertions contain permanent (or static) elements that can not be altered for an entity who authenticates to RP throughout multiple sessions (see table 3). Based on this, we anticipate that importance of the **RQ** asked in this paper is set to increase over time.

Authentication and access control systems create a multi-user environment where assumption about absolute indistinguishability of assertions (e.g. single anonymity set) is impractical meaning that complementary approaches need to be taken [25, 48]. Game theory is one of the natural ways to address *unlinkability* problem. This is not a straightforward task.

We highlight the gap that exists between unlinkability definitions and standardization procedures on one hand and non-cooperative normative decision making on the other hand. This study addresses this gap by deriving user utilities from information-theoretical measures. This provides fruitful results by enforcing corresponding game equilibria which enables up to $\frac{68.3}{\log n}\%$ relative improvement to be obtained when compared with random (chaotic) interchangeable usage of a pair of assertions.

Naïve game that is proposed in this paper requires that priors about statistics are collected at RP side and communicated to the users in a trustworthy way. Whether this can be accomplished through a mediator remains an open question. There is a risk that users who place their trust in the system may be misled by a malicious mediator. To counter this problem, we propose a solution (e.g. *tenable* game) that does not require trust, but also provides reduced utility (compared to *naïve*). Hence, we foresee that design of a mediation mechanism where users are assured about higher (than tenable game) unlinkability comprises a promising research direction.

REFERENCES

[1] Carl Adams and Vasilios Katos. 2007. Exoinformation Space Audits: An Information Richness View of Privacy and Security Obligations. *Journal of Information Privacy and Security* 3, 3 (2007), 29–44. https://doi.org/10.1080/15536548.2007.10855820

[2] Gergely Alpár, Fabian van den Broek, Brinda Hampiholi, Bart Jacobs, Wouter Lueks, and Sietse Ringers. 2017. IRMA: practical, decentralized and privacy-friendly identity management using smartphones. *HotPETs 2017* (2017).

[3] World Bank. 2018. *Private Sector Economic Impacts from Identification Systems.* World Bank.

[4] A Beduschi, J Cinnamon, J Langford, C Luo, and D Owen. 2017. Building Digital Identities: The Challenges, Risks and Opportunities of Collecting Behavioural Attributes for new Digital Identity Systems. (2017).

[5] Diana Berbecaru and Cesare Cameroni. 2020. ATEMA: An attribute enablement module for attribute retrieval and transfer through the eIDAS Network. In *2020 24th International Conference on System Theory, Control and Computing (ICSTCC).* IEEE, 532–539.

[6] Dirk Bergemann and Stephen Morris. 2016. Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics* 11, 2 (2016), 487–522. https://doi.org/10.3982/TE1808

[7] Ron Berman, Amos Fiat, and Amnon Ta-Shma. 2004. Provable unlinkability against traffic analysis. In *International Conference on Financial Cryptography.* Springer, 266–280.

[8] T. Berners-Lee, R. Fielding, and L. Masinter. 2005. Uniform Resource Identifier (URI): Generic Syntax. (2005). https://tools.ietf.org/html/rfc3986

[9] Dan Boneh, Xavier Boyen, and Hovav Shacham. 2004. Short Group Signatures. In *Advances in Cryptology - CRYPTO 2004, 24th Annual International CryptologyConference, Santa Barbara, California, USA, August 15-19, 2004, Proceedings.* 41–55.

[10] Mayla Brusó, Konstantinos Chatzikokolakis, Sandro Etalle, and Jerry Den Hartog. 2012. Linking unlinkability. In *International Symposium on Trustworthy Global Computing*. Springer, 129–144.

[11] J Camenisch, S Krenn, A Lehmann GL Mikkelsen, G Neven, and MØ Pedersen. 2014. D3. 1: Scientific Comparison of ABC Protocols. *Part I-Formal Treatment of Privacy-Enhancing Credential Systems. Project deliverable in ABC4Trust* (2014).

[12] Jan Camenisch and Anna Lysyanskaya. 2002. A Signature Scheme with Efficient Protocols. In *Security in Communication Networks, Third International Conference, SCN 2002, Revised Papers (Lecture Notes in Computer Science)*, Vol. 2576. Springer, 268–289.

[13] Jan Camenisch and Els Van Herreweghen. 2002. Design and implementation of the idemix anonymous credential system. In *Proceedings of the 9th ACM conference on Computer and communications security*. 21–30.

[14] Sam Castle, Fahad Pervaiz, Galen Weld, Franziska Roesner, and Richard Anderson. 2016. Let's Talk Money: Evaluating the Security Challenges of Mobile Money in the Developing World. In *Proceedings of the 7th Annual Symposium on Computing for Development*. Association for Computing Machinery, New York, NY, USA, Article 4, 10 pages. https://doi.org/10.1145/3001913.3001919

[15] David W Chadwick, Romain Laborde, Arnaud Oglaza, Remi Venant, Samer Wazan, and Manreet Nijjar. 2019. Improved Identity Management with Verifiable Credentials and FIDO. *IEEE Communications Standards Magazine* 3, 4 (2019), 14–20.

[16] Mike Clark. 2019. German government adds iPhone NFC identity card reading to digital ID app. *NFC World* (2019). https://www.nfcw.com/2019/10/01/364573/german-government-adds-iphone-nfc-identity-card-reading-to-digital-id-app/

[17] Sarah Clark. 2020. Germany to begin rollout of open national digital identity service 'later this year'. *NFC World* (2020). https://www.nfcw.com/2020/07/29/367360/germany-to-begin-rollout-of-open-national-digital-identity-service-later-this-year/

[18] Sebastian Clauß and Stefan Schiffner. 2006. Structuring anonymity metrics. In *Proceedings of the second ACM workshop on Digital identity management*. 55–62.

[19] Thomas F Coleman and Yuying Li. 1996. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on optimization* 6, 2 (1996), 418–445.

[20] European Commission. 2018. Looking ahead: the user experience of eIDAS-based eID. *Value Proposition of eIDAS eID* (2018).

[21] Alissa Cooper, Hannes Tschofenig, Bernard D. Aboba, Jon Peterson, John Morris, Marit Hansen, and Rhys Smith. 2013. *Privacy Considerations for Internet Protocols*. Request for Comments IETF RFC 6973. The Internet Engineering Task Force, Wilmington, DE. https://doi.org/10.17487/RFC6973

[22] Matthew Davie, Dan Gisolfi, Daniel Hardman, John Jordan, Darrell O'Donnell, and Drummond Reed. 2019. The trust over ip stack. *IEEE Communications Standards Magazine* 3, 4 (2019), 46–51.

[23] Claudia Diaz, Stefaan Seys, Joris Claessens, and Bart Preneel. 2002. Towards measuring anonymity. In *International Workshop on Privacy Enhancing Technologies*. Springer, 54–68.

[24] Dizme. 2021. The key to Digital Identity. *https://www.dizme.io/* (2021).

[25] Yevgeniy Dodis, Tal Rabin, et al. 2007. Cryptography and game theory. *Algorithmic game theory* (2007), 181–207.

[26] European Commission. 2020-07-23. Proposal for a European Digital Identity (EUid) and Revision of the eIDAS Regulation. *Directorate-General for Communications Networks, Content and Technology* (2020-07-23).

[27] European Parliament. 2014-07-23. Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC. *Council of the European Union* (2014-07-23).

[28] Matthias Franz, Bernd Meyer, and Andreas Pashalidis. 2007. Attacking unlinkability: The importance of context. In *International Workshop on Privacy Enhancing Technologies*. Springer, 1–16.

[29] Julien Freudiger, Mohammad Hossein Manshaei, Jean-Pierre Hubaux, and David C Parkes. 2009. On non-cooperative location privacy: a game-theoretic analysis. In *Proceedings of the 16th ACM conference on Computer and communications security*. 324–337.

[30] Drew Fudenberg and Jean Tirole. 1991. *Game Theory* (11 ed.). The MIT Press.

[31] Arpita Ghosh and Katrina Ligett. 2013. Privacy as a coordination game. In *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 1608–1615.

[32] Sérgio Manuel Nóbrega Gonçalves, Alessandro Tomasi, Andrea Bisegna, Giulio Pellizzari, and Silvio Ranise. 2020. Verifiable Contracting. In *European Symposium on Research in Computer Security*. Springer, 133–144.

[33] Paul A. Grassi, Michael E. Garcia, and James L. Fenton. 2020. *Digital Identity Guidelines*. Standard NIST SP 800-63-3. National Institute of Standards and Technology, Gaithersburg, MD. https://doi.org/10.6028/NIST.SP.800-63-3

[34] Paul A. Grassi, Naomi B. Lefkovitz, Ellen M. Nadeau, Ryan J. Galluzzo, and Abhiraj T. Dinh. 2018. *Attribute Metadata: A Proposed Schema for Evaluating Federated Attributes*. Technical Report NISTIR 8112. National Institute of Standards

and Technology, Gaithersburg, MD. https://doi.org/10.6028/NIST.IR.8112

[35] Joseph Y Halpern and Kevin R O'Neill. 2005. Anonymity and information hiding in multiagent systems. *Journal of Computer Security* 13, 3 (2005), 483–514.

[36] John C. Harsanyi. 1967. Games with Incomplete Information Played by "Bayesian" Players, I–III Part I. The Basic Model. *Management Science* 14, 3 (1967), 159–182. https://doi.org/10.1287/mnsc.14.3.159

[37] Vincent C. Hu, David Ferraiolo, Rick Kuhn, Adam Schnitzer, Kenneth Sandlin, Robert Miller, and Karen Scarfone. 2019. *Guide to Attribute Based Access Control (ABAC) Definition and Considerations.* Technical Report NIST SP 800-162. National Institute of Standards and Technology, Gaithersburg, MD. https://doi.org/10.6028/NIST.SP.800-162

[38] Vincent C. Hu, David F. Ferraiolo, and D. Richard Kuhn. 2019. *Attribute Considerations for Access Control Systems.* Recommendation NIST SP 800-205. National Institute of Standards and Technology, Gaithersburg, MD. https://doi.org/10.6028/NIST.SP.800-205

[39] Mathias Humbert, Mohammad Hossein Manshaei, Julien Freudiger, and Jean-Pierre Hubaux. 2010. Tracking games in mobile networks. In *International Conference on Decision and Game Theory for Security*. Springer, 38–57.

[40] IDunion. 2021. An open ecosystem for trusted identities. *https://idunion.org/?lang=en* (2021).

[41] ISO Central Secretary. 2012. *Information technology – Security techniques – Security assurance framework – Part 1: Introduction and concepts.* Technical Report ISO/IEC TR 15443-1:2012(E). International Organization for Standardization, Geneva, CH.

[42] ISO Central Secretary. 2019. *Systems and software engineering — Systems and software assurance — Part 1: Concepts and vocabulary.* Standard ISO/IEC/IEEE 15026-1:2019(E). International Organization for Standardization, Geneva, CH.

[43] ISO Central Secretary. 2020. *Information technology – Requirements for attribute-based unlinkable entity authentication.* Standard ISO/IEC DIS 27551. International Organization for Standardization, Geneva, CH.

[44] TELECOMMUNICATION STANDARDIZATION SECTOR OF ITU. 2017. *ITU-T Focus Group Digital Financial Services: Main recommendations.* Standard. International Telecommunication Union, Geneva, CH.

[45] M. Jones. 2015. JSON Web Key (JWK). (2015). https://tools.ietf.org/html/rfc7517

[46] Michael Jones and Joe Hildebrand. 2015. *JSON Web Encryption (JWE).* Request for Comments IETF RFC 7516. The Internet Engineering Task Force, Wilmington, DE. https://doi.org/10.17487/RFC7516

[47] Lou Jost. 2006. Entropy and diversity. *Oikos* 113, 2 (2006), 363–375. https://doi.org/10.1111/j.2006.0030-1299.14714.x

[48] Jonathan Katz. 2008. Bridging game theory and cryptography: Recent results and future directions. In *Theory of Cryptography Conference*. Springer, 251–272.

[49] Michael Kearns, Mallesh Pai, Aaron Roth, and Jonathan Ullman. 2014. Mechanism design in large games: Incentives and privacy. In *Proceedings of the 5th conference on Innovations in theoretical computer science*. 403–410.

[50] Gregg Kellogg, Pierre-Antoine Champin, and Dave Longley. 2020. *JSON-LD 1.1: A JSON-based Serialization for Linked Data.* Recommendation. World Wide Web Consortium.

[51] G. Klyne and C. Newman. 2002. Date and Time on the Internet: Timestamps. (2002). https://tools.ietf.org/html/rfc3339

[52] Michael Kubach, Heiko Roßnagel, and Rachelle Sellung. 2013. Service providers' requirements for eID solutions: Empirical evidence from the leisure sector. *Open Identity Summit 2013* (2013).

[53] Loic Lesavre, Priam Varin, Peter Mell, Michael Davidson, and James Shook. 2019. A taxonomic approach to understanding emerging blockchain identity management systems. *arXiv preprint arXiv:1908.00929* (2019).

[54] Xinxin Liu, Kaikai Liu, Linke Guo, Xiaolin Li, and Yuguang Fang. 2013. A game-theoretic approach for achieving k-anonymity in location based services. In *2013 Proceedings IEEE INFOCOM*. IEEE, 2985–2993.

[55] Dave Longley and Manu Sporny. 2020. *RSA Signature Suite 2018.* Specification. World Wide Web Consortium.

[56] Kiraku Minami. 2020. Trace equivalence and epistemic logic to express security properties. In *International Conference on Formal Techniques for Distributed Objects, Components, and Systems*. Springer, 115–132.

[57] John F. Nash. 1950. Equilibrium Points in N-Person Games. *Proceedings of the National Academy of Sciences of the United States of America* 36,1.

[58] K. Nohl and D. Evans. 2009. Privacy through Noise: A Design Space for Private Identification. In *2009 Annual Computer Security Applications Conference*. 518–527.

[59] Nate Otto, Sunny Lee, Brian Sletten, Daniel Burnett, Manu Sporny, and Ken Ebert. 2019. *Verifiable Credentials Use Cases.* Guide. World Wide Web Consortium.

[60] Khaled Ouafi and Raphael C-W Phan. 2008. Privacy of recent RFID authentication protocols. In *International Conference on Information Security Practice and Experience*. Springer, 263–277.

[61] Christian Paquin. 2011. U-prove technology overview v1. 1. *Microsoft Corporation Draft Revision* 1 (2011).

[62] Andreas Pfitzmann and Marit Hansen. 2010. A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management. (2010).

[63] Michael Pisa and Jim Woodsome. 2019. Overcoming the "Know Your Customer" Hurdle with E-KYC. *Ideas to Action: Independent research for global prosperity* (2019). https://www.cgdev.org/blog/overcoming-know-your-customer-hurdle-e-kyc

[64] Mikerah Quintyne-Collins, Heather Vescent, Darrell O'Donnell, Greg Slepak, Michael Brown, Christoper Allen, and Michael Ruther. [n. d.]. Digital Credential Wallets. ([n. d.]).

[65] Drummond Reed, Manu Sporny, Dave Longley, Christopher Allen, Ryan Grant, and Markus Sabadello. 2021. *Decentralized Identifiers (DIDs) v1.0: Core architecture, data model, and representations.* Recommendation. World Wide Web Consortium.

[66] Daniel Servos and Sylvia L Osborn. 2017. Current research and open problems in attribute-based access control. *ACM Computing Surveys (CSUR)* 49, 4 (2017), 1–45.

[67] Moshe Sniedovich. 2016. Wald's mighty maximin: a tutorial. *ITOR* 23, 4 (2016), 625–653. https://doi.org/10.1111/itor.12248

[68] Manu Sporny, Noble Grant, Dave Longley, Daniel Burnett, and Brent Zundel. 2019. *Verifiable Credentials Data Model v1.0: Expressing verifiable information on the Web.* Recommendation. World Wide Web Consortium.

[69] Sandra Steinbrecher and Stefan Köpsell. 2003. Modelling unlinkability. In *International Workshop on Privacy Enhancing Technologies*. Springer, 32–47.

[70] Kalman C. Toth, Ann Cavoukian, and Alan Anderson-Priddy. 2020. Privacy by Design Architecture Composed of Identity Agents Decentralizing Control over Digital Identity. In *Open Identity Summit 2020*, Heiko Roßnagel, Christian H. Schunck, Sebastian Mödersheim, and Detlef Hühnlein (Eds.). Gesellschaft für Informatik e.V., Bonn, 163–170. https://doi.org/10.18420/ois2020_14

[71] Isabel Wagner and David Eckhoff. 2018. Technical privacy metrics: a systematic survey. *ACM Computing Surveys (CSUR)* 51, 3 (2018), 1–38.

[72] Zhiyi Zhang, Michał Król, Alberto Sonnino, Lixia Zhang, and Etienne Rivière. 2020. EL PASSO: Privacy-preserving, Asynchronous Single Sign-On. *arXiv preprint arXiv:2002.10289* (2020).

A PROOFS

Lemma 1. *Best linking performance is limited by $H(\mathsf{L} \mid l)$ (for details see appendix A).*

Proof. In order to link authentication sessions RP labels them with $\mathsf{L}' \in \mathcal{L}'$, where $\mathcal{L}' = \{A', B'\}$. We divide the proof in 2 parts: (*i*) we demonstrate that for the best linking performance RP aims to minimize $H(\mathsf{L} \mid \mathsf{L}')$; (*ii*) and, $H(\mathsf{L} \mid \mathsf{L}') \geq H(\mathsf{L} \mid l)$.

(*i*) We express linking performance $\mathfrak{P}$ of RP as the difference between True Positive Rate (TPR) and False Positive Rate (FPR): $\mathfrak{P} = \Pr(A' \mid A) - \Pr(A' \mid B) = \frac{\Pr(A',A)}{\Pr(A)} - \frac{\Pr(A',B)}{\Pr(B)}$ which is to be maximized and for which we demand that $\mathfrak{P} \geq 0$. In authentication systems, probability of A, i.e. $\Pr(A)$ and probability of B, i.e. $\Pr(B)$ are decided by the users and hence can not be affected by RP. We further demonstrate that either increase of the probability that both events A' and A occur, i.e. $\Pr(A', A)$ or decrease of the probability that both events A' and B occur, i.e. $\Pr(A', B)$ reduces $H(\mathsf{L} \mid \mathsf{L}')$. We note that conditional entropy

$$H(\mathsf{L} \mid \mathsf{L}') = \sum_{\mathsf{L} \in \mathcal{L}} \sum_{\mathsf{L}' \in \mathcal{L}'} \Pr(\mathsf{L}, \mathsf{L}') \log \frac{\Pr(\mathsf{L}')}{\Pr(\mathsf{L}, \mathsf{L}')} \ .$$

is unimodal on $\Pr(A', A)$ (similar must be stated about $\Pr(A', B)$) by analyzing its first derivative $\partial \frac{H(\mathsf{L}|\mathsf{L}')}{\partial \Pr(A,A')} = \log\big(\Pr(A, A') + \Pr(B, A')\big) + \log \Pr(A, B') - \log \Pr(A, A') - \log\big(\Pr(A, B') + \Pr(B, B')\big)$ and finding its unique extremum at $\frac{\Pr(A,A')}{\Pr(A,A')+\Pr(A,B')} = \frac{\Pr(B,A')}{\Pr(B,A')+\Pr(B,B')}$. The denominators in the latter equation are equal to $\Pr(A)$ and $\Pr(B)$, respectively. As a result, $\mathfrak{P} = 0$ at this extremum, and, due to unimodality of $H(\mathsf{L} \mid \mathsf{L}')$ on $\Pr(A', A)$ (and on $\Pr(A', B)$) maximization of $\mathfrak{P}$ requires minimization of $H(\mathsf{L} \mid \mathsf{L}')$.

(*ii*) For any deterministic linking algorithm $c : \ell \to \mathcal{L}'$ it is true that $H(\mathsf{L}' \mid l) = 0$, and hence, $H(\mathsf{L}', l) = H(l)$. Next, according to the properties of joint entropy, $H(\mathsf{L}, \mathsf{L}', l) \geq H(\mathsf{L}, l)$ from which follows that $H(\mathsf{L} \mid \mathsf{L}', l) \geq H(\mathsf{L} \mid l)$. According to the conditional entropy properties we also have $H(\mathsf{L} \mid \mathsf{L}') \geq H(\mathsf{L} \mid \mathsf{L}', l)$ which finally implies $H(\mathsf{L} \mid \mathsf{L}') \geq H(\mathsf{L} \mid l)$. □

Lemma 2. *Expected utility for player i is (for details see appendix A)*

$$\mathbb{E}[u_i] \approx -s_i \log \frac{s_i}{\mathbb{E}\big[\Pr_{\mathcal{S}}(\alpha^{(i)})\big]} - (1 - s_i) \log \frac{1 - s_i}{\mathbb{E}\big[\Pr_{\mathcal{S}}(\beta^{(i)})\big]} \ . \tag{3}$$

Proof. According to eq. (2) expected value of unlinkability for the entire system is $\mathbb{E}[C] = \log n + \frac{1}{n} \sum_i^n \mathbb{E}[u_i]$ where

$$\mathbb{E}[u_i] = -\mathbb{E}\Big[s_i \log \frac{s_i}{\Pr_{\mathcal{S}}(\alpha^{(i)})}\Big] - \mathbb{E}\Big[(1 - s_i) \log \frac{1 - s_i}{\Pr_{\mathcal{S}}(\beta^{(i)})}\Big] \ .$$

We further process the first term of the right hand side of the latter equation:

$$\begin{aligned} \mathbb{E}\Big[s_i \log \tfrac{s_i}{\Pr_{\mathcal{S}}(\alpha^{(i)})}\Big] &= \mathbb{E}\big[s_i \log s_i - s_i \log \Pr_{\mathcal{S}}(\alpha^{(i)})\big] \\ &= s_i \log s_i - s_i \mathbb{E}\big[\log \Pr_{\mathcal{S}}(\alpha^{(i)})\big] \ . \end{aligned}$$

Taylor expansion of $\log \Pr_{\mathcal{S}}(\alpha^{(i)})$ around $x_0 = \mathbb{E}\big[\Pr_{\mathcal{S}}(\alpha^{(i)})\big]$ produces

$$\begin{aligned} \log \Pr_{\mathcal{S}}(\alpha^{(i)}) \ \approx \ & \log \mathbb{E}\big[\Pr_{\mathcal{S}}(\alpha^{(i)})\big] \\ + \frac{\Pr_{\mathcal{S}}(\alpha^{(i)}) - \mathbb{E}\big[\Pr_{\mathcal{S}}(\alpha^{(i)})\big]}{\mathbb{E}\big[\Pr_{\mathcal{S}}(\alpha^{(i)})\big]} - \ & \frac{\big(\Pr_{\mathcal{S}}(\alpha^{(i)}) - \mathbb{E}\big[\Pr_{\mathcal{S}}(\alpha^{(i)})\big]\big)^2}{2\mathbb{E}\big[\Pr_{\mathcal{S}}(\alpha^{(i)})\big]^2} \ . \end{aligned}$$

By taking expectation over the right hand side of the equation we arrive at

$$\mathbb{E}\Big[\log \Pr_{\mathcal{S}}\big(\alpha^{(i)}\big)\Big] \approx \log \mathbb{E}\Big[\Pr_{\mathcal{S}}\big(\alpha^{(i)}\big)\Big] - \frac{\mathrm{Var}\Big[\Pr_{\mathcal{S}}\big(\alpha^{(i)}\big)\Big]}{2\mathbb{E}\Big[\Pr_{\mathcal{S}}\big(\alpha^{(i)}\big)\Big]^2} ,$$

and, we obtain that

$$\mathbb{E}\Big[s_i \log \frac{s_i}{\Pr_{\mathcal{S}}\big(\alpha^{(i)}\big)}\Big] \approx s_i \log \frac{s_i}{\mathbb{E}\Big[\Pr_{\mathcal{S}}\big(\alpha^{(i)}\big)\Big]} + s_i \frac{\mathrm{Var}\Big[\Pr_{\mathcal{S}}\big(\alpha^{(i)}\big)\Big]}{2\mathbb{E}\Big[\Pr_{\mathcal{S}}\big(\alpha^{(i)}\big)\Big]^2} .$$

Similarly, the following equation also holds:

$$\mathbb{E}\Big[(1-s_i) \log \frac{1-s_i}{\Pr_{\mathcal{S}}\big(\beta^{(i)}\big)}\Big] \approx (1-s_i) \log \frac{1-s_i}{\mathbb{E}\Big[\Pr_{\mathcal{S}}\big(\beta^{(i)}\big)\Big]} + (1-s_i) \frac{\mathrm{Var}\Big[\Pr_{\mathcal{S}}\big(\beta^{(i)}\big)\Big]}{2\mathbb{E}\Big[\Pr_{\mathcal{S}}\big(\beta^{(i)}\big)\Big]^2} .$$

Lastly, by considering assumption 1 we obtain that $\mathbb{E}\left[u_i\right]$ equals to

$$-\mathbb{E}\Big[s_i \log \frac{s_i}{\Pr_{\mathcal{S}}\big(\alpha^{(i)}\big)}\Big] - \mathbb{E}\Big[(1-s_i) \log \frac{1-s_i}{\Pr_{\mathcal{S}}\big(\beta^{(i)}\big)}\Big] - \frac{\mathrm{const}}{2} ,$$

and, due to indifference of the constant term to the actions of i, we exclude it from further considerations and, hence, demonstrate validity of eq. (3). □

B LISTINGS FOR VC/VP

Listing 2. *Alice*'s VP from Deakin University

```
{
"@context": [
    "https://www.w3.org/2018/credentials/v1",
    "https://www.schema.org/",
    "https://www.deakinschema.org/contexts/studentcerts"],
"type": "VerifiablePresentation",
    "VerifiableCredential": [{
        "@context": [
            "https://www.w3.org/2018/credentials/v1",
            "https://www.schema.org/",
            "https://www.deakinschema.org/contexts/studentcerts"],
        "type": ["VerifiableCredential", "UndergradCredential"],
        "credentialSchema": {
            "id": "https://example.org/examples/degree.zkp",
            "type": "ZkpExampleSchema2018"},
        "issuer": "did:web:vc.deakin.world",
        "issuanceDate": "2021-03-10",
        "credentialSubject": {
            "ageOver": 18,},
        "proof": {
            "type": "AnonCredDerivedCredentialv1",
            "primaryProof": "ps9wLNSi48K5qNyAVMwdYqVHSMv1Ur8i...Hf2ZvWF6zGvcSAsym2sgSk737",
            "nonRevocationProof": "ce6fg24MfJPU1HvSXsf3ybzKARib4WxG...VTce53M6UwQCxYshCuS3d2h
                "}}],
"proof": {
    "type": "AnonCredPresentationProofv1",
    "proofValue": "JkYdYMUYHURJLD7xdnWRinqWCEY5u5hG...k115Lt3hMzLHoPiPQ9sSVfRrs1P"}
}
```

Listing 3. *Alice*'s VP from VicRoads

```
{
"@context": [
    "https://www.w3.org/2018/credentials/v1",
    "https://www.schema.org/",
    "https://www.vicroads.org/schema/drlic"],
"type": "VerifiablePresentation",
    "VerifiableCredential": [{
        "@context": [
            "https://www.w3.org/2018/credentials/v1",
            "https://www.schema.org/",
            "https://www.vicroads.org/schema/drlic"],
        "type": ["VerifiableCredential", "RestrDrLic"],
        "credentialSchema": {
            "id": "https://example.org/examples/license.zkp",
            "type": "ZkpExampleSchema2018"},
        "issuer": "did:web:vc.vicroads.world",
        "issuanceDate": "2021-02-10",
        "expirationDate": "2022-02-10",
        "credentialSubject": {
            "ageOver": 18},
        "proof": {
            "type": "AnonCredDerivedCredentialv1",
            "primaryProof": "Ox8iTNSi48K5iHyAVMwdYqVHSMv1Uu1p...Uf2ZvWF6zGdpSAsym2sgSk35Q",
            "nonRevocationProof": "bI6fg24MfJPU1pZSXsf3ybzKARib4WPc...
                OJce53M6UwgFxYshCuS3dt3"}}],
"proof": {
    "type": "AnonCredPresentationProofv1",
    "proofValue": "aLYdYMUYHUpvLD7xdnWRinqWCEY5u9hW...Lk5Lt3hMzLHoFIPQ9sSVfRrsOz"}
}
```