Verifying Identity as a Social Intersection

Nicole Immorlica, Matthew O. Jackson, Glen Weyl*

April 2019

Abstract

Most existing digital identity solutions are either centralized (e.g., national identity cards) or individualistic (e.g., most "self-sovereign" identity proposals). Outside of digital life, however, identity is typically social (for instance, "individual" data such as birthdate is shared with parents) and intersectional (viz., different data and trust are shared with different others). We formalize these ideas to provide a more robust and realistic framework for decentralized identity. We build upon the concepts web-of-trust and social collateral, from cryptography and economics, to provide a system of defining, verifying, and making use of identity through networks. We exploit the redundancy created by intersectionality together with the fragmentation of identity suggested by self-sovereign schemes to minimize social collateral required for verification. We exploit the probabilistic structure of Bloom filters to provide uniqueness proofs to prevent Sybil attacks while conveying minimal compromising information to verifiers. We discuss applications to "proof-of-personhood" blockchains and Radical Markets.

1 Introduction

Since the emergence of the modern state, identities have typically been verified by credentials such as a passport or social network account issued by a central authority such as a state or corporation (Scott, 1998). Yet such systems have significant capacity limitations (e.g., a passport cannot be used to verify present occupation as often requested by border agents) and security vulnerabilities (e.g., hacking a single database or stealing a single token are often sufficient to compromise much of an individual's identity). Alternative, "decentralized" identity paradigms have emerged to address these concerns, but are generally much more demanding on users or even more limited in their capabilities. In this paper we sketch a different paradigm for identity verification that relies on formalizing pervasive features of preformal human identity – where a person's "identity" is embodied in what is known about

*Immorlica and Weyl are from Microsoft Research, and Jackson is from Stanford University and the Santa Fe Institute. Jackson gratefully acknowledges support under NSF grant SES-1629446 and from Microsoft Research New England. We thank Vitalik Buterin, Andrew Dickson, Lucas Geiger, and Kaliya Young, for helpful conversations and comments on earlier drafts.

them by others – highlighted in the classical sociology of Georg Simmel (1908). Present technology makes an identity system based on more sociologically accurate notions feasible with minimum burdens on users, allowing systems that are simultaneously much more secure, capable and, in the relevant senses, private.

The crucial aspects of identity we draw on, identified by Simmel, are its *redundancy*, *sociality*, and *intersectionality*. By redundancy, we mean that every individual's uniqueness is over-determined by countless features from the locations she has traversed, the relationships she has established, the things she has done, the knowledge she has accumulated, her fingerprints, etc. There is no-one else with the same set of features and past interactions known about her by others. Not only is her set of features and interactions unique, but it is also greatly over-determined. There is a long history of interactions and thus there are an exponentially large number of combinations of different interactions that the person has had that uniquely determine that person. Not only is Matt the only person who talked to Nicole on October 23 and also with Glen on October 15, but he is also the only person who was at Microsoft New England on November 6 and in Ecuador on November 8. Many such combinations of facts about Matt are unique to him.

By sociality, we mean that most if not all of these data are, naturally in the course of social life or by their very nature, shared with and known by others: at most times we are in the vicinity of others, a date of birth is obviously shared with many other family members, etc. This sociality is interpreted broadly, as a relationship may be that a person was employed by a company for some period of time, and that company (in addition to many of its employees) can verify that relationship.

By intersectionality, we mean that the set of others with whom these identity markers of any given individual are shared differs for each datum and thus the individual may be seen as (in large part, at least) the intersection of the social groups with whom the constituents of her identity are shared.[1]

The features of the *pre-formal* identity are stored, in the normal course of human lives, in minds and other emergent personal records. Our central claim is that if such features can be formalized and accessed efficiently, they are a powerful foundation for many identity applications. Redundancy is valuable because it implies an individual can show her uniqueness using a small subset of the data that constitute her identity and thus may often avoid revealing too much about herself. Sociality is valuable because it implies that, to build trust in a claim about some of a person's data, it is usually sufficient for that individual to use preexisting social sharing of data, thereby largely avoiding compromises of privacy or security. Intersectionality is valuable because it ensures that individuals can, for different applications, rely on a range of different social connections and thus avoid making any other

[1]The uniqueness of the individual is then that she is the only member of the intersection of the people who worked with some company during some period, is a member of a particular family, attended a particular high school, has brown eyes, etc. To be clear, we use "intersectionality" in this classical sociological sense which is relatively neutral to the power relations of the society in which these markers are embedded, rather than in the (related) contemporary sense (Crenshaw, 1989). For a thoughtful discussion of these relationship between these ideas see Stoetzler (2016) and for a recent, brief exposition of the Simmelian view of identity see Hitzig and Weyl (Forthcoming).

individual or group a central "chokepoint" for verifying her identity, avoiding the security failings of centralized identity systems and allowing for a degree of incompleteness of any social connection's view of an individual that can substitute for "pseudonymity".

In what follows, we begin by providing background on both sociological and computational thinking about identity, focusing on how our paradigm contrasts with pre-existing paradigms we are aware of. We then develop, in Section 3, the formal primitives of our approach. In Section 4, we develop the protocol over which identity verification messages would run. In Section 5, we develop a number of extensions and elaborations of this basic protocol that will be necessary for a range of applications. In Section 6 we develop applications to several canonical uses of identity verification in the digital arena (such as voting and blockchain participation) and provide an example of an easily-implemented subset of identity information (location) that could be the basis of a practical system. We conclude with a discussion of important directions for future development.

2 Background

As Simmel (1908) observed, "pre-formal" identity systems, the way we think of ourselves or describe ourselves to others in everyday conversation, differ greatly from formal identity systems.[2] We might think to ourselves "I am the partner of A, a member of X, the sister of B, spent five years living in Q, a researcher in field L, etc." or describe a friend in a similar manner. We would almost never use language that matches with formal identity verification such as "my Social Security Number is XXX-XX-XXXX".

In contrast to formal verification protocols, such informal descriptions have several distinguishing factors. First, they are highly redundant and thus in any given description, highly incomplete. In an introduction to any person, one highlights only certain aspects of oneself, which vary across context, sometimes reflecting "personal", sometimes "professional" and sometimes "political" aspects of who one is. Even when sufficiently detailed to uniquely identify one, such descriptions are very incomplete and may have quite limited overlap across contexts, so much so that when contexts meet acquaintances are usually surprised by what they learn about someone they thought they knew.

Second, pre-formal identity is mostly relational and social. Unlike a unique number, it almost always reflects connections to other humans and human communities. Even surnames, being carried through families, are fundamentally relational.

Third, pre-formal identity is, as a result and expression of these first two points, overwhelmingly intersectional, in the sense that while (nearly) all aspects of identity are social, usually no social groups or connection contains or even knows about all or even most of the aspects of any individual. Instead, the individual sits at the intersection of the social groups and relationships to which she belongs. In the language of Zelizer (2005), each of these circles with internally public but externally private information is a "circle of intimacy".

[2]The narrative in this section, regarding identity systems and their history, draws from Scott (1998) and the Economist (2018), which give more background and history. For other discussions, see **?**Chango (2012); Young (n.d.).

Prior to the emergence of modern centralized authorities, such as nation states and corporations, such pre-formal identity was the only form that existed and intersections were sufficiently large that coherent communities shared large fractions of each individual's identity. We can label this regime "communal identity". This tightness of circles of intimacy, however, made identity very local, implying that anyone from far away would have difficulty distinguishing or trusting community members.

To permit longer-distance relationships, mobility, taxation and other features of modern society, states and corporations had to find ways to mark unique identity without the ability to record the richness and complexity of highly redundant social identity within communities. As such they created abstract simplifications of identity, such as given name-surname pairs, fingerprints, places of birth, identification numbers, etc., that were as simple as possible while achieving unique identification. To achieve this simplicity, such systems are far less redundant and have a far simpler structure than pre-formal identity. We can call this regime "centralized identity".

Such identity systems gave citizens a basis for transitioning out of small, closed, heavily-intersecting communities into large, open, and more urban environments with much thinner intersections – where the fraction of an individual's identity known to any other given individual, conditional on that individual knowing anything, is smaller. This process made pre-formal identity gradually far more intersectional in the sense we describe above, eroding the communal regime. As such, in formal settings individuals came to rely increasingly on centralized identity to navigate these thinner social relations.

Yet such identity systems have several interrelated flaws. First, they are highly *insecure*, because their simplicity and lack of redundancy meant that crucial data such as an ID number constantly had to be given out and yet was also sufficient to impersonate an individual. Furthermore, because all data is stored in a single repository managed by the state or a corporation, such a repository becomes a natural locus for external hacking or internal corruption. Second, they are highly *thin* in that they reduce an individual to a small dimensional object (in system or out, criminal or not, a credit score, etc.) as the central database has little use for more information than this. This limits the functionality of the system to a relatively small range of cases or degrades performance, often in highly unequal ways (e.g. convicted individuals find it hard to re-enter society as this is the only information about themselves they can reliably convey). Third, they are *artificial*, in the sense that the central information stored for verification usually bears little relation to the social or personal conception of identity of the relevant individual and her communities. Thus all such information is added on top of the information the individual would naturally store about herself and thus incurs a cost in security, data generation and storage costs or, usually, both.

Recently, new identity paradigms have tried to get around some of these elements. One approach, adopted by "big data" platforms like Facebook and Google, is to overcome thinness by storing enormous amounts of detailed information about each individual. we might call this "panoptic identity". However, such solutions have greatly exacerbated the other two problems, as they require extremely artificial compromises to intimacy through the

global sharing of data with platforms that would not otherwise store it, creating exceptional potential security risks.

Another approach has been "self-sovereign identity", in which identity claims are owned by and stored local to a citizen. Properly engineered, this has the potential to partly enhance security against centralized attacks; however it still suffers from having much of a person's identity stored in a single place and so if an individual is hacked and loses access to their "wallet", they can be compromised. In addition, these frameworks typically suffer from far more extreme thinness and artificiality. In particular, almost all identity claims pertain to many individuals; for example, a mother's date of birth gives information about half a dozen or more people. Given this, it is not clear what self-sovereignty even means. Furthermore, identity claims can only be verified by other citizens, requiring the storage of attestations by other citizens to specific facts that create additional, artificial burdens.

The Basic Logic of Our Approach Our approach builds on pre-formal identity and mirrors how people check on other people's claims outside of a formal system. As an example, if a manager wants to hire someone and verify that an applicant worked at some company during some time period in some role – something that would qualify the applicant for the position for which he is applying – then that manager might make some phone calls. If the manager knows and trusts someone at the company where the applicant claims to have worked previously, then the manager could call that person directly and ask the person to verify that the applicant's claim is true. If the job was particularly important, and the manager wanted a high level of confidence, then she might verify the applicant's claim with more than one person. If the manager did not know someone directly at the company, she might call a friend whom she knows and trusts and who she thinks would know someone at the company. She might ask how well the friend trusts her acquaintance at that company, and if there was enough trust, then ask the friend to check on the applicant's claim for her. The more important it was to the manager to verify the applicant's claim, the more trust she might require on a such a chain, and she might even use several independent chains.

In an application like voting, the person running an election might not only want to check up on some fact, such as a voter being a resident of a particular district, but also ensure that the voter does not vote twice. This would not only involve checking not only that the voter resides in the district, but also that the voter is the unique resident at a particular address within the district with a particular age and gender or other combination of attributes. In a small community this could be done informally via chains of checking. By tracking even just a few characteristics that uniquely identify a given voter, then someone whose characteristics overlapped with that voter would not be able to vote again.

In these examples, there are aspects of both authentification and authorization – as the individuals are checked to have characteristics that make them eligible for employment or voting, and in the second case that they are the only person with certain characteristics – so that they are both eligible and not the same as someone else who has already voted.

Our idea is very much to build directly on these ideas and structure, but to augment it so that it works in a decentralized network. The main difference between these examples and

what we are proposing is that this could be done via a system that would preserve privacy, allow for much longer and richer chains, to check on multiple facts, and to track and update trust in the system – but the basic structure and logic is identical to the pre-formal checking .

The identity paradigm most closely related to the one we propose is called "Web of Trust". This approach was developed in the early 1990s to counteract these fundamental problems. In these solutions, each individual has a cryptographic key which serves as a proxy for identity. These keys are signed by members of the community, often at physical parties, creating a web of trust. Keys with many trusted signatures are themselves trusted and can be used as proof of identity. While these systems are effective in shifting the gate-keeping functionality of identity systems from the governing body to the society itself, they are still deeply artificial, requiring the creation and signing of dedicated keys, which may then be stolen. For these reasons, these systems never saw widespread use.

Our basic approach is to draw on some of the formal structure of web of trust while relying much more heavily on the structure pre-formal identity discussed in the employment and voting examples mentioned above. Each individual is associated with a large set of features corresponding to information that would, pre-formally, already be stored in their minds and used to imagine their own identity (e.g. the identity of a first kiss, place of marriage, past places of residence, employment, coworkers, siblings, etc.) *as well as in the minds of others with whom this information is naturally shared*. Individuals also have pairwise friendships with varying degrees of trust. The friends of an individual know a subset of that individual's features. When individual A wishes to verify the identity of individual B, A chooses a subset of features and asks B to reveal their values. A then verifies B's answers by searching the network of friendships for individuals that can verify these answers and are connected to A with a path of highly trusted friendships. If this process reveals any lies, the trust on links transmitting the lies is decreased, thus penalizing the individuals stating or transmitting lies. The number of features used in verifying identity should be large enough that the answers of individuals will be unique and small enough that A can't use B's answers to steal B's identity.

This proposal maintains security by distributing information and by avoiding a single, alienable key as a method of identification, instead relying on information the individual would know in any case. It is thick, because most of the crucial information about an individual is available and can be verified somewhere in the network. And, most centrally, it is natural, because information used is already known to the individuals involved, ensuring a minimum of artificial information sharing. Security is further bolstered because redundancy ensures that no individual has to share, for verification, more than a small part of her identity with a verifier. This fragmentation also ensures an information structure that may be superior to pseudonimity or anonymity, in which a thin slice of true identity is used without revealing unnecessary parts. It also is very resilient to having one's identity stolen, since the volume of facts and features about an individual is so enormous that it is impossible for anyone else to mimic – in strong contrast to relying on just a few or single identifiers such as a national identity number or social security number identify an individual. In what

follows we formalize this paradigm and discuss a range of applications.

3 Primitives

Identity is a set of features, such as birth date, gender, ethnicity, city, education, employment, education, skills, experiences, family, etc., henceforth interpreted as a string of bits. Let there be K features (K will typically be very large, hence redundancy), and let k denote a generic feature. A question of the form "In what year was individual j born?" thus becomes a question about the value of some feature, which would take on a value such as 1971. It could also be the answer of whether j was in a given place or at some meeting at a specific date and time. This means that there are an exponentially large number of potential bits about any given individual that might be of use and interest.

Each individual has a set of friends and organizations (local governments, employers, hospitals, banks...) who know things about the individual (viz. sociality). Each friend or organization will typically know only some subset of an individual's features (viz. intersectionality).[3] Individuals and their friendships and relationships create a weighted, directed, and multi-layered social network (e.g., see (Kivelä et al., 2014)).

Formally, this social network is a set of individuals $\{i\}$, a *trust network* consisting of a weighted set of edges $\{w_{ij}^k\}$ representing relationships, and a set of *knowledge states* $\{(s_{ij}^k, c_{ij}^k)\}$ for each i representing what individual i knows about others in the network. These are all indexed by the bits as they may differ depending on the type of knowledge in question. We defer the question of the initiation and labeling of individual nodes and the network for an actual protocol to Section 5.1, and for now take it as given.

- The weight $w_{ij}^k \in \mathbb{R}^+$ indicates the *degree of trust* that i places in j regarding evaluation of direct or indirect information about bit k.

- The set s_{ij}^k represents i's *belief about the value(s)* of j's k'th bit. For example, s_{ij}^k might contain the set of programming languages that i believes j to be proficient in, or the years of experience that i believes j has programming. If i has no knowledge of j's k'th bit, we say $s_{ij}^k = \perp$.

- The value $c_{ij}^k \in \mathbb{R}^+$ represents i's *confidence in belief* about s_{ij}^k. Here we allow for errors. For instance, an employment record may have an out-dated address, while a roommate or spouse may be fully confident in someone's address.[4] Alternatively, c_{ij}^k might be the amount of confidence i has regarding j's ability to deliver some skill, e.g.,

[3] This knowledge can be obtained exogenously or endogenously. Exogenous information is obtained when a feature naturally concerns multiple individuals. For example, if A is the oldest child of B, then A's birth year is also the year that B's first child was born. These are, respectively, features of A and B and the knowledge of these features are shared by both A and B. Our system also allows for individuals to share information endogenously. If individual A has a high degree of trust in individual B, then A can reveal some of A's features to B, making it easier for others to verify A's identity.

[4] This could be operationalized by using something like the number of days since the value of the bit was last verified if it is simple factual information that might change.

complete a routine programming assignment within a reasonable time. If $s_{ij}^k = \perp$, then we assume $c_{ij}^k = 0$.

In general, these values may be time-varying and modified by the identity protocol, and may be further refined based on individuals' relationships and beliefs, as discussed in Section 5; for now we take them as fixed.

Trust can vary by the type of information. For instance, it may be that j is an experienced programmer, and so i trusts j at evaluating information about programming skills, but not about accounting or medicine. It is also a direct extension to have different trusts for whether j is being used as an intermediate node in a chain, so whether i trusts j's evaluation of some other node h's reliability, as opposed to j's knowledge directly about bit k.[5] We do not formalize that distinction in order to save on notation and transparency.

Trust can also be in different units depending on the application. If it is evaluating the financial solvency, it might be some amount of value - as sort of credit limit, up to which one node trusts another to evaluate. Instead, it could simply be a probability number indicating how likely i thinks it is that information that comes through j is correct.

Beliefs about specific bits can be private or shared by various combinations of individual . For instance, it could be that only individual i knows the value of s_{ij}^k, for instance if it is a parent i's knowledge about a child j's allergy; or it could be common knowledge between i and j if it concerns whether they met at a certain time and place on a certain date; or it could be commonly known by a set of coworkers if it concerns whether j was a co-worker of i at a certain company on a certain date.

4 Protocol Calls

Our system verifies individuals' features by searching the trust network according to the following protocol. Suppose individual j, the *claimant*, wishes to convince some i, the *verifier*, that the value of her k'th bit is σ. For example, if j is applying for a programming job at institution i and claims to have 10+ years of programming experience, the k might be "years of programming experience" and σ might be 10+.

The protocol proceeds iteratively until the verifier is either satisfied or decides to quit searching. In the τth iteration, the verifier identifies a target "stake" parameter, say κ_τ. This is the total trust and confidence value that the verifier needs in order to be satisfied of the claim that the value of j's k-th bit is σ, or in this case 10+.

The protocol then looks to find a sequence of individuals, $V = \{v_1, \ldots, v_p\}$, such that

- v_p has $s_{v^p j}^k = \sigma$, and
- the chain of trusts $w_{iv_1}^k, w_{v_1 v_2}^k, , w_{v_{p-1} v_p}^k$, together with $c_{v_p j}^k$, have a total "evidence value" exceeding the threshold κ_τ.

[5] We can all think of friends who are very knowledgeable about some topic, but a poor judge of people, or vice versa.

More generally, we allow the protocol to find collections of sequences $V_1, \ldots V_\ell$, possibly overlapping, where each sequence ends up with some node that has nontrivial knowledge about j's kth bit. For instance, it could include several sequences: $V_1 = \{1, 2, 3, 4\}$, $V_2 = \{5, 2, 3, 6\}$, $V_3 = \{2, 8, 4\}$. Here, nodes 3 and 4 are called on twice, and 2 is used in all of the sequences, and the particular link 23 is used twice, and 4's information about the value of j's kth bit is used twice.

The way in which such collections of sequences of trust and confidence are evaluated can be done in many ways, and can be adjusted depending on the specific context and type of information k. We mention two.

1. The *maximum flow* (equivalently, minimum cut) that the verifier i can route along the collection of sequences used. More precisely, let the verifier be the source node. Connect a sink node t to each terminal individual in $v_p \in V_\tau$, and have the weight from that v_p to the terminal node be $c^k_{v_p j}$. For each τ in $1, \ldots, \ell$ and for any consecutive pair vv' in sequence V_τ set the capacity between them equal to $w^k_{vv'}$. The maximum flow f in this graph is the amount of evidence provided.[6]

2. The *probability* of the claim. More precisely, suppose each w^k_{ij} and c^k_{vj} are in the unit interval (or scale them accordingly). Instead of calculating the max flow, now let $w^k_{vv'}$ be the probability that v and v' successfully transmit, and $c^k_{v_p j}$ the probability that v_p communicates its information. Then look at the overall probability that information is successfully transmitted.

If the amount of evidence acquired exceeds κ_τ, then the protocol terminates. Otherwise, the verifier starts a new iteration, should she desire to do so. She may also increase or decrease the stake.

There are many variations that one can imagine. For instance, it might be that one ends up with ten different sequences and 9 of them give information that j has 10+ years of experience, but one claims that there are only 5 years of experience. The verifier could decide that is sufficient, or could end up thinking that that is not enough and want to do more checks.

We mention two desired properties of potential implementations.

Limiting Congestion: First, to avoid flooding the network with verification requests, it is recommended that the protocol set limits on the number of iterations, and perhaps on the size of the collection of chains used. The protocol would choose a minimal collection to provide sufficient evidence, if there are multiple possibilities - and more on this below.

Maintaining Privacy: Second, for privacy purposes, it is recommended that intermediate nodes not learn σ. For extremely sensitive claims, cryptographic techniques such as zero knowledge proofs can even let the verifier check a claim without learning its statement.[7]

[6] See Mobius and Szeidl (2007); Karlan, Mobius, Rosenblat and Szeidl (2009) for more discussion of using paths for trust purposes. In our setting, a min cut becomes the key limiter.

[7] A simple way to do this is to ask the claimant and all nodes in V_τ to report a hash of the value of the k'th feature rather than the feature itself.

This might be particularly useful in preventing identity theft. The protocol could also be worked in such a way that the actual sequences are not known to any of the participants, only the outcome.

Authentification and Authorization Above, we have described how i checks on a bit of j having a certain value. There is also a question of making sure that the j that is being reported is actually the same j that i is talking to.

The protocol would simultaneously check things as follows: First, i has to choose some bit(s) k' that uniquely identify j as the person that they are thinking of hiring – the person they are actually talking to, for instance that this is the same person that was interviewed on a certain day, or exchanged some specific messages with i. The protocol, could verify that j not only had bit k having 10+ as a value, but also that this intersects with the same j having bit k' indicating that j had an interview with i at some particular time.

Thus, the protocol would be checking the value of multiple bits about some j simultaneously, and some of those would be identifying j as the person in question, while others would be verifying that j has some desired characteristics – thus both authenticating and authorizing at the same time.

5 Extensions and Elaborations

There are a number of important issues that arise regarding incentives in being sure that such a system can work, and here we discuss a few most critical issues.

5.1 Initializing the Network and Nodes

In order to run this on a decentralized computer system, the i's and j's we described have to somehow be initiated to store information about a person's knowledge and history, including information about others, and then set trust levels, confidences, and so forth. If Cleo wants to find which node corresponds to her friend Mark, she could initiate the system by asking nodes to answer questions that only Mark would know, and then offer the trust to the node that correctly answers the questions. Nodes would not need to have permanent addresses, as their identity could always be established by answering questions to other nodes in this way.

There is also some danger of having all the identifying information of some person stored in some single place - like a wallet or a node, as whomever "controlled" that node could then assume that person's identity. To avoid this, some information would not be placed on the node, but kept by the person. For instance, Mark's node might not be told what his favorite food or song was, or what he said to Cleo at the last meeting, but might have other information that Cleo could check such as a list of their various rendezvous. Then, Cleo could ask questions both of the node and the controller, and check that they both correspond to Mark. The node itself could ask many questions of Mark in order to be operated, but this

extra level would also allow Mark to be sure that nobody else could operate the node even if they hacked into it.

This level of extra security could be costly in terms of communication, and so might be adjusted depending on the importance of the verification.

5.2 Liens and Limiting Fraud

One way in which a system of identity could be manipulated would be for some single individual to gain someone's trust and then abuse it. For instance, say that j gains i's trust and gets a high w_{ij}^k, and then j claims to have knowledge about many other individuals and that knowledge is actually false. For instance, j claims that a number of j's friends have some feature k that is valuable for them to have (e.g., having a high credit score, etc.) even though none of them actually have that feature. If many verifiers use a path that relies on i's trust of j to verify claims about j's friends, then the single weight w_{ij}^k is being heavily leveraged. One erroneous w_{ij}^k could be exploited arbitrarily by a fraudulent j. A way to avoid this is to track traffic through given links, and to put temporary liens on them. For instance, the system could note that link ij has been used for a verification where the verifier required some amount of trust (e.g., flow in the above algorithm) to be used on that link $x_{ij}^k \leq w_{ij}^k$. While that lien was in place, that link would have a reduced value of $w_{ij}^k - x_{ij}$.

How much of a lien would be put in place, and for how long, would depend on the application. It would depend on things like the risks being taken by the verifier (for instance, how big a loan they are making to the person whose credit score they are attempting to verify), how many past times the link ij has been successfully used, and whether other paths are also being used. But by tabulating the traffic through a given link, one can then control its use and prevent it from being abused so that one node's misplaced trust in another cannot be exploited to more than some capped amount that relates to the original weight in place.

5.3 Re-routing

In cases in which there are multiple routes, which links end up being used can make a difference, especially in cases in which liens may be placed on a link. For instance, there may be multiple paths that a verifier could use to check an identity, and which one they choose could end up making a difference to a later verifier. For instance, it might be that a first verifier could use either link ij or link $i'j'$ as part of a verification path, while a second verifier could only use ij. Thus, if the first verifier ends up using ij and placing a lien on it, that could prevent the second verifier from being able to do their check. Given that it is not easy to predict which conflicts might arise (for instance, some third verifier might end up only being able to use $i'j'$), it makes sense to be able to reoptimize. For instance, if the first verifier has used link ij and placed a lien on it, and later the second verifier needs it, then then it would be possible to change the path of the first verifier to use $i'j'$ and to shift the lien, and then free up the trust and link ij to be used by the second verifier. This constant reoptimization could become much more involved as the combinations of paths that

are best suited for any given set of verifications can become an (NP) hard problem, but by reoptimizing, even approximately, a much more efficient set of checks may be possible.

5.4 Evolution of Trust

This brings us to a related question, which is how trust evolves over time. The more some link ij is used successfully (so that it leads to a verification that turns out to be deemed useful and truthful), it appreciates and leads to additional trust.[8] The more it leads to incorrect information, the more it would depreciate.

The details of how much it would appreciate or depreciate, to properly align incentives, would have to relate to the social marginal value of the information that it provides. That will depend on the particular circumstances and the ways in which trust is used, and so we have little in particular to say at this point.

5.5 Incentives to Participate

Providing accurate trust information, and maintaining information about bits and confidence in those bits about others is a costly activity. In order to provide incentives for people to do this, they need to be rewarded for doing it and doing it accurately.

The basic incentives for participation in the system would be that one accrues "capital" that one can then use to pay for future queries. In order to make queries, one needs to provide services as well. The use of proper scoring rules, and other methods of ascertaining whether people's trust and confidence numbers are accurate over time, would help provide incentives for calibration. There would also need to be some rewards for providing breadth of information: that is, providing information about as many others as one has identity information about. Properly calibrating this will involve some care, as one does not want to reward people from simply creating false profiles just to claim to know about them; and so the rewards would come from actual use of posted information in proportion to its usefulness.

5.6 Adversarial Attacks

Identity systems are prone to several common attacks. We explain how our system defends against these attacks.

5.6.1 Multiple Identities

In an identity system, it is important that each individual corresponds to a single identity. If an individual gains access to multiple identities, then she can subvert many identity applications, e.g., by voting multiple times or receiving multiple loans. In what is commonly called a Sybil attack, an individual forges multiple fake identities. Our system of stakes and liens is naturally resistant to such an attack. Identity is stored and verified in a highly

[8] This applies to whether it is useful in verifying that someone has some value of some bits, or else showing correctly that they fail to have them. The key to appreciation is that the link resulted in useful information.

distributed manner, and so Sybil attacks either require collusion from a large collection of individuals, or require a network of forged identities which will quickly be alienated by the system of liens.

In particular, in order to establish a new identity, one has to either bribe/collude with many individuals who are already trusted in the network, or work with a group of new and marginal individuals who are not already connected in the network and then rely on a small number of bridges to the rest of the network. It is easy to flag that some group is essentially isolated except for a few bridges - and liens on those bridges limit the damages that can be done over them.

5.6.2 Identity Theft

Another common concern in identity systems is identity theft, in which one individual assumes the identity of another by "stealing" identifying information. Identity systems are particularly prone to this attack, since verifiers justifiably gain access to individuals' identifying information in the identity verification process. This problem is exacerbated in applications such as voting where verifiers must collect enough identifying information to prove a database consists of unique individuals.

This is actually a much worse problem in more centralized system in which individuals are identified by a small number of tags, such as a social security number, drivers license number, or voter id number, and the authorities who issue those numbers do not have other ways of verifying who is whom, other than matching those names with a name, birthdate, and address. In such systems, gathering a few numbers is enough to assume someone else's identity.

In the more decentralized system that we describe, an individual's uniqueness and identity come from a large number of bits held by many others. Even if someone steals some of that information, the thief also needs to get others who are trusted in the network to also assert confidence that those stolen bits are associated with the thief. Stealing an identity becomes a much more complex and difficult task.

In addition, to prevent identity theft one can also limit the number of answers an individual provides to any particular verifier. While this effectively combats identity theft, it may create difficulties for verifiers that wish to prove uniqueness (although enormous redundancy, may render this a non-issue – as there could be so many things that over-determine someones uniqueness that having someone have a small fraction of them is not a threat.[9] Nevertheless, verifiers wishing to prove uniqueness could also employ hash functions to maintain individual privacy. For each of a (potentially large) set of questions, the verifier selects a hash function[10] and asks an individual for the hash of her answer. The verifier then asks other individuals on the network to also hash their knowledge of this answer. If the hashes match, then the answer is verified but not leaked (so long as the hash function is not inverted). Assuming sufficient heterogenity in the bit strings of individuals, and a sufficient number of

[9]This is a much richer version of the idea that someone who knows a person's password but not what the person's first pet's name was may still be unable to access the person's account.

[10]This hash function must be unique from any used previously to prevent copycat attacks.

questions per individual, then uniqueness of a vector of hashed answers implies uniqueness in the database.[11]

Finally, the system can also be built with different protocols for who has the right to release which information, and who is made aware of it. Some knowledge, for instance that a co-worker worked at a given company at the same time, is knowledge that is something that a person naturally acquired and is not prevented from spreading (unless it is an espionage organization). However, if a person ends up sharing some bits with someone else, one could place a lock on the use of those bits so that they are still the property of the owner. For instance, it may be important for a company to know something about a worker's age to know something about retirement benefit qualifications. However, that does not give the right for that company to share that bit with anyone else unless the employee authorizes it. The employer could become someone who could verify something about the employee's age, but only with the consent of the employee. This would add an extra dimension to our verification protocol - but an important one. Each bit of information would have different rights and permissions associated with it, indicating who has to consent to release that information, and who has the right to be informed of that bit's release. Some such laws exist in the current world, but quite haphazardly and incompletely, and this could be much more naturally and extensively built into the system that we are proposing.

6 Applications

In this section we briefly discuss a few potential applications of our identity system. Given the pervasiveness of identity and referral systems, the list is far from exhaustive.

6.1 Location Data: An Illustrative Example

Much of our protocol may seem abstract. How are all these data to be actually entered and where are they to be stored? Won't all of this be too burdensome on individuals? How will individuals prove they are the ones making the request at the time it is made? To illustrate how this might all be made to work in a simple and self-contained way, our first example discusses how a subset of information that could be recorded to achieve most of our goals. The application is from physical location.

Suppose that each smartphone recorded its possessor's location internally at regular intervals, say 10 minutes, using calls to Global Positioning Satellites. Furthermore, suppose that every time two individuals phones were in a line of sight of each other during one of these regular intervals, the phones established a link, stored on both phones, indicating they had not just been at that place at that time, but they had also "seen" the other.

[11] We note there are natural tradeoffs between the number of questions, the load of the hash function (its domain divided by its range), the uniqueness guarantees, and the risk of identity theft. The higher the load, the harder it is to invert by ranging over the domain or asking for multiple hashes of the same answer, but also the more questions a verifier must ask to guarantee uniqueness.

A secure log-in to the phone could then be accomplished, rather than by passwords, by questioning the user about where she was on such an such a date and time. Obviously some versions of this would be difficult for humans to recall, but could likely to be designed to provide sufficient contextual cues to aid recall without revealing critical information. As a user attempted to access more secure information, more intricate recall challenges could be posed.

Verifying one's location at points in time could then be accomplished in a semi-automated version of alibi verification in present forensics: a network search from the requester could identify others who (se digital representation) could verify the verifiee's location during the relevant time. To verify uniqueness using a Bloom filter, a requester would ask a series of scattered questions about location over time, sufficient to ensure that with high probability no two individuals would have had identical answers to these questions. Obviously the closer the social relationship between relevant individuals, the more questions would have to be ask to distinguish, but excluding conjoined twins it would be possible to relatively rapidly verify uniqueness, especially with some aid from the verifiee in choosing relevant points in time. To ensure that a verification is legitimate, a requestor could rely on paths of trust that (collectively) place high confidence that the verifier is being truthful, and could check with more verifiers.

If verifiers do not wish to reveal their location, for many applications it would be sufficient to rule out some specific location rather than pinpointing a location – for instance, it could be sufficient to check that the verifier was in close proximity to the verifiee at some time stamp and was not at the location where some crime took place. This can also be kept private if the algorithm simply reveals that path(s) of sufficient trust revealed that the verifiee was at a given location at a given time, but without revealing those paths or too much about the verifiers. This is feasible if users trust the trust system.

Obviously this example leaves many details to work out, and uses a metaphor of phone location that raises other privacy issues,[12] but it provides a concrete illustration of our broader approach.

6.2 Voting

A central role of identity in modern democratic societies is voting. Without some sort of identity system, one-person-one-vote voting systems would be hopelessly vulnerable to attack as an attacker could claim many identities and swamp the vote. Identity systems typically play two roles in voting that are related but separable. The first and most important is *uniqueness*; that is, ensuring someone votes no more than once. The second is *eligibility*, that someone voting has the right to do so. The first role is usually the most important one and more effort is usually invested in it, requiring that people vote in person, and in many developing countries some form of temporary indelible ink is used to mark that a citizen has

[12] Also, phones could be stolen or moved around by some surrogate. That takes us beyond the point of our discussion here. For a tighter metaphor, here one might think of a phone as a chip implanted in an individual that is harder to steal.

voted. The second requires maintenance of some centralized database with identifiers that can cross-check uniqueness of registrations; or more primitively relies on physical presence in some location and/or personal acquaintance with some voting authority.

Our procedure offers potential solutions to both of these challenges.

As we discussed in Subsection 5.6 above, our approach offers a natural method of verifying uniqueness of participants in a one-off manner, and could also do so in a persistent manner if the relevant query responses were maintained as the citizen's identity for the purpose of voting over time. Furthermore, as we discussed in Section 5.2, our approach is naturally designed to limit the amount of fraud possible: if there is an individual or small group attacking the system, the amount of fraud they can commit is limited to the trust placed in the weakest link connecting the to the verifier. By modulating κ_τ, we can modulate a trade-off between limiting fraud (by raising it) and increasing privacy and inclusion. Furthermore, for citizens with many pathways to the verifier for many of their characteristics, as should be the case in most open societies for most citizens given intersectionality, we should expect the trade-off to be relatively favorable.

Our procedure also seems natural for eligibility, at least for many of the eligibility criteria used by vote administrators at present. For example, to the extent that some length of residence in a region is part of an eligibility requirement, given that such long residence will inevitably lead to physical interpersonal proximity with a variety of people and organizations who can attest to the residence during that time period. For settings that rely on religious identity or conditions of birth, similar attestations by those present at or around birth or religious observances will be natural supporters.

One interesting extension is to systems of voting such as Storable Votes (Casella, 2006) or Quadratic Voting (Lalley and Weyl, 2018) in which electors have a persistent currency that may be used over several elections in exchange for votes. In this case, κ_τ could be adapted to the amount of votes a citizen desires to use in any given election, could be a more one-time initialization allowing an account to be created with a given vote endowment, or a mixture in which a flow of vote credits into the account requires the maintenance of a certain level of κ_τ. In any case, some form of persistent account based on a subset of identity traits would be necessary to ensure that the relevant elector could, but others could not, view her account balance to avoid coercion or undue influence.

6.3 Proof of Personhood Blockchains

Public and permissionless blockchains have become an increasingly popular data structure in recent years. Without digressing into technical details (Nakomoto, 2018), the basic idea of blockchains is for a shared public database (usually called a "ledger") to be redundantly stored on any computer willing to host it and the validation of transactions to take place (effectively) through the majority rule of the nodes hosting the database. To compensate the nodes for this service, nodes are allowed to (with some probability) add some of the value of the currency whose transactions are recorded on the ledger to their account. Any computer capable of running requisite computations ("proof of work" or PoW) is allowed to play this role; this proves that this node is not simply a duplicate of another.

This PoW consensus suffers from at least two limitations. First, allocating one vote on validation and one unit of earning to each unit of processing power tends to create a plutocracy governed by those with most access to the capital to purchase such processing power in markets external to the system. This can undermine the goal of decentralization that motivates blockchains by allowing dominance by a wealthy clique. Second, proving independence in this way requires the waste of computational resources (and ultimately energy) on puzzles that serve no productive role in the system.

Both of these limitations could be overcome by an independent identity system and in fact have been thus overcome in so-called "permissioned chains", such as Ripple (Schwartz et al., 2014) in which those allowed to participate in the consensus are granted access by a central validating authority. Such central validating authorities, however, seem to largely defeat the aim of decentralization that motivate blockchains.

A natural middle ground that might preserve the goal of decentralization could grow out of our protocol. A chain could build out from an initial creator based on trust and, up to lien capacity, those who join could either become validators themselves or introduce others whose uniqueness is proved through the protocol. With enough time to "repay" required liens, given the many ways in which most individuals are connected, this process would likely quite rapidly be effectively permissionless in that most world citizens would find a route to verify their uniqueness and be included. Validation could then occur by voting, including potentially by some of the voting techniques described above.

More ambitiously, one could imagine alternative data structures that more directly exploit the structure of trust and knowledge networks. For instance, in actual social networks, friends-in-common play strong roles, and a relationship between i and j can function better and be more trustworthy if they have a mutual friend (e.g., see the discussion in Jackson (2019); Jackson et al. (2012)). Furthermore, rather than global chains and private wallets being the two primary components, there could be a variety of different community data stores holding data shared within a relevant community and exposed quasi-publicly to that community but not outside. This could be a digital formalization of the Zelizer (2005)'s spheres of intimacy or something approaching it. While it might be necessary to compromise the natural storage location of data to exploit "cloud" economies of scale, this could happen as local as possible to the natural circle of intimacy. Such a data structure is an interesting direction for future research.

7 Conclusion

In this paper we have proposed a paradigm for identity verification based on insights from the sociology of identity: namely that identity is redundant, social and intersectional. This paradigm may overlap with much other work that we are not aware of and is only one step towards implementation. However, we believe that its tight correspondence to the social reality of pre-formal identity makes it a natural candidate for further research and embellishments.

In particular, an attractive feature of this approach to identity is the effective power

structure it embodies. Centralized data structures create single points of failure not just for "hacking" but also for a small group, possibly associated with a nation state or profit seeking corporation, to exploit access to and control over a wide range of data for purposes that may not be transparent or democratically accountable. Purely individualistic data structures, however, are unsustainable and largely incoherent. The redundant, social, intersectional structures we describe, on the other hand, naturally diffuse power. In such a setting, just as in pre-formal societies, those attempting to gain access to sufficient information to conduct and investigation, carry out large scale social engineering or other potentially authoritarian actions will typically require the at-least-grudging assent and knowledge of a wide range of communities. Such communities can then, if such actions are seem as illegitimate, intervene to stop them, providing a natural technological check on authoritarian actions that goes beyond the formal legal protections against them.

To build such a system, however, will obviously take great elaborations on our framework. We now briefly describe some areas we would particularly like to see further developed.

The baseline formalism we develop above is provided at a high level and omits many details that need to be worked out for implementation. Should trust be conditional on knowledge domain? How should confidence and the degree of trust be related to each other? How can trust networks be reliably maintained without relying in an excessively burdensome way on input from the user? How should those supporting the verification of an individual limit the social stake they are willing to put on the line for that individual?

Beyond the basic protocol, many of our incentive mechanisms are even hazier. Over what time period should liens be repaid? What guarantees do these offer on the amount that can be earned through coordinated attacks? How much trust should be augmented and after what events, or how much time, if trust turns out to be earned? When trust is betrayed, should all implicated links be fully burned? Or exactly how should blame be distributed?

Perhaps the greatest challenges come around more direct user interface and intrastructure costs. What is the best way to integrate security around access to one's data store into such a system? Should such personal data stores be locally retained (say on a smartphone), stored remotely but in a highly personalized manner or be managed by a socially local "cloud" provider? How can the entry of the trust graph and recording of the knowledge graph be both secure and impose minimal burden on the user?

Despite these challenges, we look forward to experimentation with identity solutions that exploit redundant, social, intersectional structure.

References

Casella, Alessandra, "Storable votes," *Games and Economic Behavior*, 2006, *52* (2), 391–419.

Chango, Mawaki, "Becoming artifacts: Medieval seals, passports and the future of digital identity," *Syracuse University, School of Information Studies - Dissertations, Paper 74*, 2012.

Crenshaw, Kimberle, "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics," *University of Chicago Legal Forum*, 1989, *1* (8), 139–167.

Economist, The, "Establishing identity is a vital, risky and changing business," *The Economist*, 2018, *49:9123, December 22*, 63–68.

Hitzig, Zoë and E. Glen Weyl, "A Crossroads, not an Island: A Response to Hanoch Dagan," *University of Michigan Law Review Online*, Forthcoming.

Jackson, Matthew O., *The Human Network: How Your Social Position Determines Your Power, Beliefs, and Behaviors*, Pantheon Books: New York, 2019.

_ , Tomas R. Rodriguez-Barraquer, and Xu Tan, "Social Capital and Social Quilts: Network Patterns of Favor Exchange," *American Economic Review*, 2012, *102* (5), 1857–1897.

Karlan, Dean, Markus Mobius, Tanya Rosenblat, and Adam Szeidl, "Trust and social collateral," *The Quarterly Journal of Economics*, 2009, *124* (3), 1307–1361.

Kivelä, Mikko, Alex Arenas, Marc Barthelemy, James P. Gleeson, Yamir Moreno, and Mason A. Porter, "Multilayer Networks," *Journal of Complex Networks*, 2014, *2* (3), 203–271.

Lalley, Steven P. and E. Glen Weyl, "Quadratic Voting: How Mechanism Design Can Radicalize Democracy," *American Economic Association Papers and Proceedings*, 2018, *108*, 33–37.

Mobius, Markus and Adam Szeidl, "Trust and social collateral," *National Bureau of Economic Research working paper 13126*, 2007.

Nakomoto, Satoshi, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2018. http://wfc-knowledgecentre.com/wp-content/uploads/2016/07/Bitcoin-A-Peer-to-Peer-electronic-Cash-System.pdf.

Schwartz, David, Noah Youngs, and Arthur Britto, "The Ripple Protocol Consensus Algorithm," *available at https://www.cryptiaexchange.com /Whitepaper_Ripple.pdf*, 2014.

Scott, James C., *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*, New Haven, CT: Yale University Press, 1998.

Simmel, Georg, *Soziologie: Untersuchungen über die Formen der Vergesellschaftung*, Berlin: Suhrkamp Verlag GmbH, 1908.

Stoetzler, Marcel, "Intersectional Individuality: Georg Simmel's Concept of "The Intersection of Social Circles" and the Emancipation of Women," *Sociological Inquiry*, 2016, *86* (2), 216–240.

Young, Kaliya, "Domains of Identity."

Zelizer, Viviana A., *The Purchase of Intimacy*, Princeton: Princeton University Press, 2005.