Project Vouch

A decentralized identity and reputation system based on public attestation and approval of identity attributes.

Angus Champion de Crespigny

04/07/2017

Incomplete historical document published for public reference only

Revision History

Version	Author	Date	Change Description
1.0	Angus Champion de Crespigny	03/14/2017	Initial draft
1.1	Angus Champion de Crespigny	04/7/2017	Restructured and rewrote paper; removed technical design

Table of Contents

Executive summary

Project Vouch is the project name for an open access, peer attestation-based, self sovereign, secure identity system. In this system, an entity's identity profile resides in a data structure stored on or linked to a decentralized network, or blockchain. This identity profile contains attributes, each of which have been attested for by other entities on the network. Entities are rewarded for making claims on other entity's credentials that are determined to be accurate, and punished for making inaccurate claims. This reward is granted with cryptocurrency distributed through the network.

Introduction – The identity problem

Bitcoin and other cryptocurrencies have achieved an enormous amount of focus over the past few years, spawning hundreds of businesses which cumulatively have attracted billions of dollars in funding. While Bitcoin was developed as an alternative payments mechanism however, there are numerous reasons why it has not been accepted by the traditional financial sector and why its growth as a consumer product has been hampered. While a number of these reasons will improve over time, such as scalability and usability, the primary reason could be seen in the fact that payments, in the current regulatory environment, can be seen as having two components: not just value, but also identity of the parties. While Bitcoin can satisfy the former, regulators require an efficient means to satisfy the latter to prevent and/or detect financial crime. Without regulatory buy-in, legal businesses cannot be established to connect with your everyday consumer, and certainly the traditional financial sector will keep a wide berth.

Solving the identity problem is not easy. In the traditional sector, financial institutions verify identity through a Know-Your-Customer (KYC) process, which involves citing physical identity documents, typically consisting of a passport or drivers licence, and a utility bill. This process is sufficiently expensive that it creates somewhat of a moat around each customer relationship, and makes it so costly to change financial institutions that in some countries people are more likely to get divorced[1]. New Bitcoin startups found similar problems in the early days of the industry, as it became apparent that a significant cost in moving money did not lie in the technical difficulties, but in the regulatory compliance hurdles.

To streamline this process of identity verification and KYC, we could establish digital identities based on digitized versions of these official documents. This has been done in some parts of the world. There are difficulties with this, however. Since digital assets can be duplicated, how do we ensure that the digital identity has not been stolen? In addition, how do we ensure that the identity is in fact authentic, if an authoritative party such as a government has not verified it in the digital realm? Some solutions addressing these issues are coming out, such as EY's identity management platform[2] for the former, and national identities in some countries[3] for the latter. India's Aadhaar project[4], utilizing a biometric key to a national identity, looks to address aspects both.

These, however, are somewhat incomplete to an ultimate future state of self sovereign identity[5]. While in practice they may work, they do not allow the full control, transparency, portability, consent, and minimalization that is required of a full self sovereign solution. On a pragmatic level,

[1] Check
[2] http://www.coindesk.com/big-four-firm-ey-begins-blockchain-id-platform-rollout/
[3] Estonia?
[4] Aadhaar
[5] http://www.coindesk.com/path-self-sovereign-identity/

collecting all identity data by a government is also unlikely to be an ideal solution for the state: such a centralized system, while not only expensive to establish and maintain, would be a significant national security risk.

Therefore, what is the solution? If blockchain technology can be used to move value without any intermediaries, it can be used to confirm identities without any authoritative party in the same way. While Bitcoin works by peers (or miners) validating the accuracy of transactions, this paper outlines how the same technology can be used for peers to calculate and validate the accuracy of claims against people's identity attributes. In doing so, we can establish a self sovereign identity that requires no government, and satisfies all requirements as specified by (ChristopherAllen)

The number one goal of this application is to allow individuals to establish provable, self sovereign identities without requiring a central approving body. These bodies may be costly, corrupt, dangerous, or non-transferrable. Additionally, establishing a recognizable formal identity should be as cost free as possible, and should ultimately be seen as a fundamental human right.

An identity network based on peer attestation

Overview

The following sections outline the general structure of the Project Vouch (or simply Vouch) identity network, focusing on understanding of the general concepts of the required protocol. Implementation considerations and some technical details are addressed later in the paper, and technical design is covered somewhere else entirely. This may be publicized at a later date.

Vouch's backbone is an open source public blockchain, publicly accessible with no fees or barriers to entry. Any individual can join the network by generating a public address, which will function as the link to the individual's identity record. This identity record will be a data structure containing identity attributes such as name, date of birth, and primary address. Other participants on the network will attest to the veracity of these attributes through the network.

This identity record will consist of two components: an attribute hash table (or simply attribute table), and a status table. The attribute table stores a list of attributes that have been validated, along with a score, and is stored encrypted on the network. The status table will be unencrypted.

For ease, this paper may use 'public address' and 'identity record' interchangeably, although technically the former is a reference to the latter.

Attest transactions

To attest to the veracity of attributes of participants on the network, they must know with sufficient confidence that the identity record on the network is associated with the person that they are attesting to – the attestee.

The attestor would do this by composing and sending an Attest transaction. This transaction includes the address of the identity record, the attribute being attested to, the proposed value, and the signature of the attestor. It is then encrypted using the attestee's public key, and submitted to the network. When the attestee observes this transaction, they decrypt it using their private key and add it to their identity record's attribute table. The attribute table is then encrypted, and posted back onto the blockchain as an update.

By encrypting the attestations as well as the aggregation of the attestations in the attribute table, we are able to keep them truly private to only the attestee's identity, however this also means that we rely upon the attestee to update the attribute table themselves. Therefore, to ensure that the attestee is aggregating the attestations correctly and that they are not manipulating them to alter the values of their identity attributes, we use specific cryptographic functions, being Pedersen commitments and ring signatures, in a similar method to confidential transactions as described by (Maxwell). These functions allow us to confirm that the Attest transactions that we have observed on the network have been aggregated correctly into the attribute table that we can also see on the network, despite the fact that both are encrypted.

The attestee, consequently, now has an attribute table showing a list of values for a particular identity attribute, with the number of attestations against that value shown alongside it. This can be demonstrated as follows:

Attribute	**Value**	**Count of attest Transactions**
Name	Angus Champion de Crespigny	55
	Agnus Champion De Crespigny	22
	Angus Descrepney	4
	Samuel L Jackson	1
Date of birth	September 1	40
	December 25	16
	February 14	3

Consequently, we have a majority view that the name is Angus Champion de Crespigny, and the date of birth is September 14.

This functionality, however, is clearly incomplete.

Reputation

Due to the open access nature of the network, it is clear however that some attestations should be valued more highly than others. An attestation from a bank or government department should be considered more reliable than that of an unknown anonymous participant. Consequently, we should incorporate reputation into the network and associate reputation scores of the attestors in Attest transactions, such that an attestee would produce an overall score for an identity attribute by aggregating the reputation scores of all those who have attested to that attribute. This in turn would build the attestee's overall reputation. In essence, an identity that is more broadly recognized by its peers as being associated with a validated person should have a greater reputation, which in turn should be more reliable when attesting to the identity of others.

A reputation score, however, should not be static, and should not be a carte blanche to influence the network. Consequently, there should be a process by which reputation can be negatively impacted due to misbehaviour. This will be implemented via the status table, which the network will incorporate into calculations of reputation score. The status table has entries added to it as they occur.

One obvious form of abuse could come from participants with high reputations intentionally or unintentionally providing attestations of others which are incorrect. The difficulty with this is determining what is incorrect in an environment where there is no one authoritative source. The solution is to, like Bitcoin, run on the assumption that the network will have more good actors than bad actors, and consequently reward the majority and punish the minority. On a regular schedule, the network will consider the attribute value with the highest score from attestations the correct value. The network will reward those who attested to this value, and all others will be punished by adding an Incorrect Attestations entry in the identity account's status table. Consequently, reputation will need to incorporate not only the validity of someone's identity record, but also their performance on the network.

Nonetheless, people make mistakes, and errors in behaviour should not ensure a permanent mark on a participant. For this reason, certain entries in the status table will be purged over time.

The method of reward is discussed later in the paper.

Validate process

Much like the network is unable to verify that the attribute table has been correctly updated due to the encrypted values, the network is also unable to independently verify which attribute value has scored the highest. We once again therefore require the attestee to provide the network with the identities to be punished and rewarded. We do this again using the combination of Pedersen commitments and ring signatures to ensure the integrity of the validation process.

Considering that the validation process requires the interaction of the attestee, and that the attestee is unable to manipulate the process if they do not like the attestations that have been given on their identity, they may choose to delay executing the validation process. As we wish to avoid this to ensure the healthy running of the network, delays will be shown on the identity's status through the visibility of a Last Validated entry which is recorded after each Validate process. After a certain time period has elapsed since the Last Validated entry, further delays will negatively impact the user's reputation. Similar such delays in the attestee updating their attribute table on the network will be recorded in their status table.

Sybil attacks and network membership

While the network design to this point should ensure that all identity accounts associated with individuals perform in the manner intended, there is still a potential weakness in the ability to freely create numerous malicious identity accounts and attack the network – in other words, a Sybil attack. These accounts could be used to attest to the identity attributes of another in a malicious fashion, with no adverse implications of a resulting bad reputation in a failed 'attack'. To prevent this, the network will have two countermeasures. Firstly, participants in the network will need to pass a defined reputation threshold before they become members in the network and can participate in Attest transactions. That is, they will need to have sufficient attestations from known members of the network before they can participate in Attest transactions themselves.

The second countermeasure is a Sanction transaction. This transaction can be sent in the same way as an Attest transaction on an identity account's primary field, likely Name, if it is felt to be a fraudulent account. The Sanction value is aggregated in the same way as other values, and consequently if found to be the majority value, those users would be rewarded and others, who have attested to the identity of a fraudulent account, would be punished.

As it is unlikely an account would trigger the Validate process at all if it was determined to be a sanctioned account, this transaction may not be encrypted and may be evaluated separately.

Reward functionality

For the network to function, there needs to be appropriate rewards and consequences for behaviour. Therefore, affecting reputation scores based on behaviour will need to discourage such behaviour by more than simply a number on a network.

A fully publicly operating network can do this, and encourage network growth, by the distribution of cryptocurrency or tokens. The incentive would perform in the following way.

Correct attributions, as determined by the Validate process, earn tokens in proportion to two factors: the time of the Attest transaction (the earlier it is sent, the greater the reward) and the reputation of the attestor (the higher, the greater).

The time of the transaction is important, as attestors should be rewarded in proportion to the risk involved. An earlier attestation, where no one has attested to the value, is higher risk than a later attestation, which would confirm the value of what many have before, and consequently the attestor should be rewarded as such. This also encourages the rapid correction of invalid attestations, as users can investigate potential misbehaviour and rally other users to correct the record, somewhat like activist investors, albeit in a far less disputable environment.

The reputation of the attestor is relevant for a number of reasons. Firstly, this is a means to encourage a higher reputation score and avoid misbehaviour that may impact the status recorded of the identity account. This way, any misbehaviour directly impacts the attestor’s future earning power. Secondly, an Attest transaction from a high reputation party has significantly higher value, and therefore may be susceptible to bribery or coercion. Consequently, they should be rewarded in a way to discourage such bribery. In this way, the tokens can encourage not only good behaviour on the network, but encourage them to detect and filter out bad behaviour as they are rewarded for doing so.

Fees

Every network needs to manage congestion, particularly blockchain networks that store the complete history of transactions. Transaction fees are traditionally used in blockchain networks to manage congestion, however we do not want to discourage attest transactions which strengthen the network. Consequently we will charge fees for the two other main transactions which use the network – namely, reveal transactions and fund transfers.

Reveal transactions

To this point we have discussed how network participants can develop robust identity profiles. To prove to someone the value of some or all of their identity attributes, the participant must reveal it to another participant through the network in a secure manner. To do this they will send a transaction to the relevant party, along with a relevant fee to the network, encrypted with the public key of the receiving participant. Such a transaction would incorporate a proof tying the value back to the attribute table, and may include logic to validate that an attribute meets a certain criteria, rather than revealing the attribute itself, for example that the participant is over a certain age.

The user may also choose to encapsulate a reveal transaction within a smart contract, with an appropriate fee attached that the user wishing to confirm the identity must deposit before the transaction is sent. This fee, however, would be sent to the user whose identity is being revealed. In this way, users can charge for their identity information.

To ensure congestion is appropriately managed, the greater complexity of the logic, the greater a fee the network would charge.

Fund transfers

The final transaction defined in this paper is the traditional fund transfer, for allowing participants to send their Vouch tokens to others. These, as stated above, will require fees to be processed. Fund transfer is secondary in priority for the network, but is obviously necessary.

Identity account examples

Following are examples of identity accounts and their reputation calculations.

Example 1: First World responsible citizen with a financial relationship

Table 1 Attribute table

Attribute	**Value**	**Score**
Name	Angus Champion de Crespigny	5500
	Agnus Champion De Crespigny	15
	Angus Descrepney	4
	Samuel L Jackson	1
Date of birth	September 1	5400
	December 25	20
	February 14	3

Table 2 Status table

Value	**Date**
Incorrect Attestation	January 15, 2017
Incorrect Attestation	February 1, 2017
Last Validated	April 4, 2017

In this example, the user has a high score against their name and date of birth, as it has been verified by a financial institution: a participant with a high reputation score. There are some other entries, however these have little credibility. The user made two incorrect attestations, however these are aging and will likely soon be purged. He validated his account recently.

A reputation calculation may look like the following:

$$Reputation = 5500 \times C_{name} + 5400 \times C_{dob} - \big(Now - (April\ 4\ 2017)\big) \times C_{valid} - \frac{C_{attest}}{Now - (January\ 15\ 2017)} - \frac{C_{attest}}{Now - (February\ 1\ 2017)}$$

That is, the following produces the reputation score:

1. The scores of each attribute are multiplied by specific coefficients before being added
2. The time elapsed since the last Validate process is multiplied by a coefficient and subtracted from the reputation score

3. A final coefficient is divided by the time elapsed since an incorrect attestation for each Incorrect Attestation record. Each of these calculations is subtracted from the overall reputation score.

Consequently, a generalized reputation calculation would look as follows:

$$Reputation = Name\ Score_{max} \times C_{name} + DOB\ Score_{max}\ \times\ C_{dob} - (Now - Last\ Validated)\ \times\ C_{valid} - \sum_{i}^{n} \frac{C_{attest}}{Now - Incorrect\ Attestation_i}$$

Note that attribute scores are going to be far higher than reputation scores. Additionally, the coefficients may not actually be coefficients or rational numbers, but may be formulas themselves.

Example 2: Financial institution

Table 3 Attribute table

Attribute	**Value**	**Score**
Name	ABC Bank	1000000
	ABC	15
	Ellemeno Bank	4
Date of birth	VOID	1000000
	January 1	4
	February 1	3

Table 4 Status table

Value	**Date**
Last Validated	April 7, 2017

The financial institution has a very high score for its name due to multiple attestations from financial institutions with high reputations, and from individuals who can easily validate the institution due to its public nature. It has some false entries however these have no realistic chance to replace the primary value. The data of birth is irrelevant in this case, although in the final implementation this may be used for date of incorporation.

Due to their stringent KYC processes and automated systems, they would be unlikely to have many, if any, incorrect attestations, and would perform the Validate process on an automated basis.

Example 3: Third World citizen without formal identity

Table 5 Attribute table

Attribute	**Value**	**Score**
Name	Charles Barkley	130
	Chuck Barkley	40
	Charles Barroom	10

Date of birth	May 15	100
	May 12	80

Table 6 Status table

Value	Date
Incorrect Attestation	January 15, 2017
Incorrect Attestation	February 1, 2017
Incorrect Attestation	February 3, 2017
Incorrect Attestation	February 8, 2017
Last Validated	March 20, 2017

This participant does not have as high attribute scores as other participants, as they have had to gather attestations from friends, family, and significant people in their community. Consequently the attribute scores are closer, although only the top score for each attribute will be relevant. Additionally, the participant has more incorrect attestations and hence their reputation will be affected. This may be the case if easy to use customer interfaces are not developed to help standardize the name format that people use.

The result of this is that this participant will have a lower overall reputation score and lower attribute scores, but can still present some level of confidence that his attributes are correct.

In reality, this particular account is likely to be typical of all identity accounts when the network is first deployed.

Example 4: Rogue user

Table 7 Attribute table

Attribute	**Value**	**Score**
Name	Cheryl Tunt	40
	Carol Tunt	20
	SANCTION	10
Date of birth	October 31	40

Table 8 Status table

Value	Date
Incorrect Attestation	January 15, 2017
Last Validated	January 16, 2017
Incorrect Attestation	March 16, 2017
Incorrect Attestation	March 16, 2017
Incorrect Attestation	March 17, 2017
Incorrect Attestation	March 18, 2017
Incorrect Attestation	March 19, 2017
Incorrect Attestation	March 22, 2017

In this example, the participant has not validated their account in some time, and has done a significant number of incorrect attestations quite recently. They also have low attribute scores, although they are present. There may be numerous reasons why this is happening, however at least some network participants believe this is a fraudulent account, and hence have submitted Sanction transactions.

A few things may happen to this user. The incorrect attestations will significantly impact their reputation score, reducing their earning potential on the network. With this number, it may even reduce the reputation score below the required Member threshold, consequently preventing them from submitting any more attestations. The number of Sanction transactions against the user could increase to become the highest scored value for the Name attribute, at which point the account would lose its membership. In other words, this account could be removed from network participation in multiple ways.

If this account is in fact a valid user, then they will need to remediate their behaviour and likely seek additional attestations for their attributes. Over time the incorrect attestations will be purged from the status table, and their reputation can improve, including regaining membership to the network.

Implementation considerations

Public addresses and identity accounts

It is unlikely that there will be a one-to-one association between public addresses and identity accounts, as in doing so raises two issues: firstly, it creates a digital fingerprint that, while a network participant's identity attributes may not be visible, their activity would be. Secondly, it provides a fixed point of access to the participant's identity, which may not be practical for long term security. Consequently, the public address used in transactions will likely be linked cryptographically to the participant's identity account, but is unlikely to be the same.

Identity accounts and attribute tables

To minimize the amount of data stored on the blockchain, attribute tables will likely be stored off-chain, with only the cryptographic fingerprints required to perform the network's functions stored on the chain.

Impossibility of a perfect solution

No system will be perfect. The goal of this system is not to be a perfect, all encompassing identity solution, but to provide a building block that will satisfy some of the most pressing issues in a standard, global identity. While all attempts will be made to incorporate as much flexibility into the system's design to allow future enhancements and allow the development of all desired features, the primary priority will be to establish a stable network. Regardless of the strength of the development team, building an autonomous economy is high risk, and we cannot be certain of the unintended consequences. Therefore, the initial protocol would be launched with a reduced set of features. Then, as the network's integrity and scalability are demonstrated, future features will be incorporated.

Predefined attributes

The accepted attributes will be predefined to a subset of fixed values, such as name, date of birth, and country of birth. Future iterations will incorporate flexible values such as citizenship and home address. The structure should ultimately enable the addition of other identity attributes, however a method to ensure that additional identity attributes are appropriately awarded is unclear. That is, validating someone’s name is of high value; validating that of their pet’s is not, and should not be awarded equally. While the significance of new attributes could be determined by a form of crowd validation, it would be safer to address this in a later upgrade. Additionally, it is likely that additional attributes beyond the fundamental attributes would require a fee.

Fixed coefficients

Initial coefficients will be fixed and updated by a central administrator. Future iterations will include a voting mechanism. Note, this voting mechanism can be performed off chain in the initial launch.

Obfuscation of Reputation, and parties in Attest operations

While technically these values can be obfuscated, a benefit of the transparency of these in the network is to address potential weaknesses in the network at intial stages. That is, to identify networks of malicious participants, or systematically weak reputations due to an unusual number of Validate operations from low reputation accounts. The trade off is visibility of overall reputations and Validate networks, however as all identity information and the scores associated with attributes is obscured, little information is leaked. Future enhancements may include complete obfuscation.

Alignment with Self Sovereign Identity requirements

Alignment with Verified Credentials standards