A social-capital based approach to blockchain-enabled peer-to-peer lending

Janka Hartmann  
Department of Computer Science  
INSA Lyon  
69621 Villeurbanne, France  
hartmannjanka@gmail.com

Omar Hasan  
Department of Computer Science  
INSA Lyon  
69621 Villeurbanne, France  
omar.hasan@insa-lyon.fr

Abstract—There are many peer-to-peer lending platforms that offer users to obtain a loan by committing a collateral or by calculating a “credit score”, which is based on factors such as the users’ credit history. However, the requirements of collateral and credit history are quite burdensome for some users. Nowadays, with more than 55% of the global population using social media [6], there is a lot of public personal data. This data could be used as an alternative risk mitigator for lending. There are many inferences that can be drawn from the users’ social media accounts about their professional behavior and reliability, allowing us to derive the users’ social trustworthiness. We propose to calculate a “social score” based on the social media data of a user. Our contribution is to develop an Ethereum blockchain-enabled fully decentralized lending platform that relies on this score. This platform could give users a chance for a loan even if they do not have a collateral or a sufficient credit score.

Index Terms—blockchain, Ethereum, social capital, lending, finance, P2P

I. INTRODUCTION

Peer-to-peer (P2P) lending allows a borrower to receive a loan directly from a single or multiple individual lenders. The first peer-to-peer lending platform Zopa [8] went online in 2005. The peer-to-peer lending market has been continuously growing since then and is predicted to keep growing in the future [1]. Some other well known P2P lending platforms are CoinLoan [9], Inlock [10], Prosper [11], and Lending Club [12]. Peer-to-peer lending has several advantages in contrast to traditional lending. It dispenses with middlemen. The lending platform itself sets the conditions and enables the transactions. Not having middlemen saves time and money, which often allows the platform to offer better rates.

An additional development in recent years is that peer-to-peer lending platforms are based on blockchain and are using smart contracts. This development brings more trust and transparency. However, since there are no middlemen verifying a potential borrowers’ financial situation, they need to prove to the lenders that they are credit worthy. Loans can be secured or unsecured. They are received by either depositing a collateral or by calculating a credit score to prove one’s creditworthiness.

A. Secured Lending

The term “secured lending” describes a way of lending, where the loan is secured with a collateral. A collateral is a valuable asset (for example, a mortgage on the borrower’s house, investments in cryptocurrencies, etc.) which the borrower has to give as insurance for the loan. After receiving the loan, the borrower has to pay back the money within a certain time. If the borrower is unable to pay, the debt is deducted from his collateral. Secured lending carries an element of risk for the borrower: If he cannot pay back the money, he loses his asset. Moreover, the borrower needs to be in possession of a suitable asset in order to qualify for a loan. The lender on the other hand is promised to get his money back. He is therefore often willing to offer better interest rates, which can be an advantage for the borrower in this kind of lending.

B. Unsecured Lending

Unsecured loans or personal loans work without a collateral. Collateral has two main problems: Firstly, people may not trust in the lending platform enough to deposit their asset. Secondly, people do not always have the money or property required for a collateral. They can have a good income and be a reliable person but without a collateral they might still not be qualified for a secured loan. Instead of taking a collateral as insurance, unsecured loans rely on a creditworthiness system, which is mostly based on a “credit score”. A well-known credit score is the FICO (Fair Isaac Corporation) score [7]. It determines the creditworthiness of a potential borrower by using a fixed formula, which takes into account aspects such as the borrower’s payment history, available credit and age. A penalty has to be paid if the monthly payment is not balanced on time. When the borrower defaults paying back his loan, he loses points of his credit score, but not a collateral. Therefore, from the borrower’s point of view, unsecured loans are more risky, but can be linked to higher interest rates. A common example for unsecured loans are student loans.

C. Peer-to-Peer Lending

In peer-to-peer lending, there exist secured loans as well as unsecured loans. However, while in traditional lending there is mostly some kind of financial institution participating in the process, peer-to-peer lending offers borrowers and lenders
to connect directly without such an intermediary. This can translate into lower or no fees and there is no longer a single point of failure. However, since there is no borrower creditworthiness evaluation carried out by a third party, the individual lender himself is responsible for determining whether a person can be trusted to pay back their debts. Peer-to-peer lending platforms are online platforms that offer to match people that want to lend money as a form of investment with people who want to borrow money. A “peer” can also be a company or a group that is in need of a loan. An example of a loan to an individual could be a payday loan, whereas companies might need a loan for commercial reasons or to expand their business.

D. Our Approach

The goal of this work is to develop a new approach to calculate the borrowers’ trustworthiness based on their social capital, which does not depend on a collateral or the credit history of the user. This approach can stand on its own or it can be used in addition to traditional concepts such as collateral and credit score. The objective is to minimize the risk that the borrower defaults on the loan. It would give those users a chance for a loan who do not have a high enough credit score or the resources to deposit a sufficiently valuable collateral. This approach is based on the “social capital” theory. In this work, we develop a prototype of a decentralized lending platform built on the Ethereum blockchain. The creditworthiness of the users on this platform is represented by a social score, which is calculated by analyzing the users’ social media accounts.

E. Outline

In section II, an introduction to three successful peer-to-peer lending platforms and their methods to ensure the borrowers’ creditworthiness is given. In section III, we present the fundamentals of the social capital theory. In section IV, the idea developed in this paper based on the calculation of a social score is presented. Sections V and VI comprise of the implementation and evaluation details. This is followed by the conclusion.

II. RELATED WORK

In this section, we describe three existing online peer-to-peer lending platforms. CoinLoan [9] and Inlock [10] both offer secured loans whereas Prosper [11] offers personal loans. The fundamental principle is that one peer lends a loan and another peer borrows a loan. All three platforms include interest fees, which the borrower has to pay to the lender and an origination fee, which the borrower pays to the platform for using the service.

A. CoinLoan

CoinLoan [9] is a peer-to-peer platform founded in 2017. The advantage of borrowing money with CoinLoan is to get a loan right away without having to provide anything except for a collateral in cryptocurrency. The collateral amount, the interest rate and the origination fee are calculated from the user’s inputs. The borrowing money function on CoinLoan is only useful for people owning cryptomoney. Moreover, once a user receives a loan, his collateral is blocked until he has paid off his debts. During this time, the user is not able to sell the deposited crypocurrency, which might want to do in case the cryptocurrency is facing heavy deprecation. If the user does not pay back on time, the owed amount will be taken from his deposit of cryptomoney.

B. Inlock

Inlock [10] is another peer-to-peer lending platform that was founded in 2017. Again, there are no options to prove one’s trustworthiness other than to give a collateral. The collateral has to be paid in the form of cryptomoney. Inlock currently supports only four cryptocurrencies: Bitcoin, Ether, Litecoin and Binance Coin. There is a 110% over-collateralization rate along with a universal collateral termination level. Once the collateral decreases below that level, the debts will automatically be paid off by Inlock using the deposited collateral. Thus, the user has to be careful and keep an eye on falling market values of the crypocurrency.

C. Prosper

Prosper [11] is a peer-to-peer lending platform that was founded in 2005. Unlike Inlock and CoinLoan, it does not offer secured, but only personal and hence unsecured loans. Although the user has to give personal data, applying for a loan is a simple and quick process. Since there is no collateral, users do not need to deposit anything and therefore they do not risk to lose their collateral. On the other hand, penalty fees can rise quickly. For not paying back on time, the borrower has to pay USD 15 or 5% of the outstanding debts. They also lose credit score points. The origination fee is significantly higher than it is on other lending platforms such as CoinLoan. Prosper is also restricted in the kinds of loan they offer. As an example, they exclude student loans and other educational loans.

III. SOCIAL CAPITAL

Next, we introduce a concept on which our proposed score is based. According to Rene Dubos in his book “Social Capital: Theory and Research” the “premise behind the notion of social capital is rather simple and straightforward: investment in social relations with expected returns” [4]. He gives four reasons for why “embedded resources in social networks will enhance the outcome of actions” [4]. Firstly, connections can help to get information and information can translate into opportunities. A good example where social connections are often useful is job hunting. Secondly, Dubos claims that having social connections may also have a positive impact on decisions involving the individual, such as discussions about promotions. The decision making process can be strongly influenced by a person putting in a good word for the individual. Thirdly, he claims that companies might value a person’s social capital on top of his personal capital. “The individual can provide “added” resources beyond his / her personal capital, some of which may be useful to the organization” [4]. The fourth
reason Dubos states is that being well connected provides both “emotional support” [4] and “public acknowledgment of one’s claim to certain resources” [4].

Robert Putnam claimed that “economic performance as a whole is better in well-connected societies than in poorly connected ones” [2]. This claim triggered many studies on the topic. A Swedish study on unemployed Swedes resulted in the conclusion that “network size had a considerable positive impact on the likelihood of finding work, far outweighing the official employment agency” [2]. Another study, performed in Germany, revealed “that engagement in a range of social activities is positively linked with job-finding among the unemployed” [2].

In this paper, we aim to use social media network information of a person to determine their social capital and to draw conclusions about possible connections in other areas like financial behavior. “The central idea of social capital is that social networks are a valuable asset” [2]. Being well-connected on social media brings advantages similar to real world connections. It helps the user to stay informed and to find possible opportunities. Further, a social media entry, such as a picture about the individual participating in a certain event, could help the individual to receive recognition. It could be a conversation opener and enable them to establish new contacts. The contacts might be useful and might put in a good word for the individual at some point.

Another widely investigated and supported claim of Putnam is that “higher levels of social capital . . . translate into lower levels of crime” [2]. Further studies have demonstrated that there is higher criminality in neighborhoods where people live rather anonymously and do not maintain contact with their neighbors [2]. Social connections have a large impact on people’s well-being, but there seem to be more benefits. “There appear to be clear and often strong positive links between social capital and educational attainment, economic success, health and freedom from crime.” [2].

IV. OUR DECENTRALIZED LENDING PLATFORM

In this section, we propose a new approach for decentralized lending: unsecured loans based on the users’ social score, instead of their credit score.

A. Creditworthiness Depending on a Social Score

The social score presented in this paper is calculated based on one or several social media accounts of the user. Our algorithm analyzes the user’s accounts to determine his personal social score. The algorithm is based on six hypotheses that estimate the trustworthiness of a user.

a) Hypothesis 1: Users who add their social media account have less to hide: Our social media account says a lot about us: who are our friends, what are our likes and dislikes, what are our ambitions, etc. The social media account may also provide some personal data, such as our age, current location, and our profession. Most people are aware of the fact that their social media account can reveal a lot of information. Therefore, if they have something to hide, for example a bad habit like gambling, they will hesitate to connect their social media to a peer-to-peer lending platform, that calculates a score derived from their social media information. As we saw in the neighborhood example [2] given in the subsection about social capital: Criminality is lower when anonymity is absent and people are part of a community. Based on this theory, the risk to default may be lower for users who add their social media accounts and therefore lose their anonymity. As a result, one may assume that users who disclose their social media account have less to hide and may therefore be considered trustworthy.

b) Hypothesis 2: The more the user is willing to disclose about himself, the more trustworthy he is: There are a lot of social media platforms these days. Currently, three important platforms are Facebook, Instagram and LinkedIn. All three platforms together cover a wide range of information about a user, which include both private and professional information. The more accounts from different platforms a user is willing to disclose, the more information about himself he is ready to provide. By giving these information about himself, the user proves once again his willingness to give up his anonymity. As already mentioned in hypothesis 1, we assume that users who give up their anonymity have less to hide and may therefore be considered trustworthy.

c) Hypothesis 3: Trustworthy users have authentic social media profiles: To avoid that users simply disclose some accounts they just created or some fake account for the pure purpose of improving their social score, the authenticity needs to be checked. Their are many indicators when it comes to identifying fake accounts. These include posting original pictures, having a significant number of mutual friends and followers, and having a non-recent date of creation of the account. For example, on Instagram, fake accounts do not have a lot of mutual friends and typically follow more people than they are followed by. According to [3], fake profiles have around 30 times as many friends as followers.

d) Hypothesis 4: The bigger the social network and activity, the more credit worthy is the person: The Swedish study concluded that “network size had a considerable positive impact on the likelihood of finding work” [2]. Therefore, users with more social contacts and thus more followers and friends on social media, are less likely to get stuck in unemployment. Also, the well-being of people strongly depends on their social network and on how connected they are. The more social activity a user has, the higher is his social capital. “There appear to be clear and often strong positive links between social capital and educational attainment, economic success, health and freedom from crime.” [2]. Therefore, more friends and connections as well as posts lead to a higher social score.

e) Hypothesis 5: People who make truthful statements are trustworthy: “Interpersonal trust is fundamental for the effective functioning of social interactions as well as of society as a whole. It has been found to be related to many societal outcomes such as lower corruption perception” [5]. Trust is fundamental in peer-to-peer lending as well. A lender needs to trust in the borrower’s good will to pay back the loan. To
be trusted to get a loan, a user’s honesty is tested. Therefore, before connecting with the social media accounts, the user will be asked to give three personal information: the full name, date of birth, and email address. When connecting the social media accounts, these information will be compared. If the information are corroborated by multiple social media accounts, this is taken as an indication of the user’s honesty and openness.

f) Hypothesis 6: Consistency is a sign for stability: If a user agrees to connect multiple social media accounts, the data of these accounts can be checked for consistency. Being friends with the same people and showing a similar profile on different social media platforms indicates that these accounts represent one and the same person.

B. Social Scoring in Peer-to-Peer Lending

The mentioned hypotheses need to be converted into variables and formulas that we can calculate our social score with. Each of those variables will have an impact on the final social score. The final social score will be received by calculating the average of the individual social media platforms accounts plus a bonus for disclosing more social media accounts.

\[ DISC_x = OPEN_x \times AUTH_x \]

(1)

The first variable is \( DISC_x \), which stands for disclosure of the user’s account from the platform \( x \). Here, \( x \) could be for Facebook, Instagram, or LinkedIn. The variable can vary between zero and one, depending on the user’s openness to disclose his social media account \( x \) and that account’s authenticity. The \( OPEN_x \) variable can only take the value one or zero. If the account \( x \) is disclosed by the user, \( OPEN_x \) equals 1, whereas if the user doesn’t disclose his account \( x \), the \( OPEN_x \) variable is zero and therefore \( DISC_x \) equals 0. The \( AUTH_x \) variable describes the authenticity of the disclosed account \( x \). It varies between 0 and 1. It is zero, if an account contains no information at all or if it is classified as fake. It is one when an account is authentic. The variation between 0 and 1 is incremental according to a value pre-defined by the platform operator. For example, one could add 0.1 for the first 10 followers / friends and another 0.1 for the next 20, and so on. The maximum value of \( DISC_x \) is therefore one if the user discloses his account \( x \) (\( OPEN_x = 1 \)) and the account is classified as authentic (\( AUTH_x = 1 \)). In the following we present a function to calculate the Social Score of a person \( x \), which is abbreviated by \( SC_x \).

\[ SC_x = DISC_x \times HON_x \]

(2)

\( HON_x \) is short for honesty. The entered user data on our peer-to-peer lending platform consists of the name, the email address and the date of birth. This data is compared to the data available on the social media platform \( x \). If none of the information match, \( HON_x \) equals zero. If only the email address matches, the honesty-value is 20. The same applies if only the name matches. A matching birth date adds 10 to \( HON_x \). The maximum value for \( HON_x \) is 50 when all three information match.

\[ bonus = n \times 2 \times CONS \]

(3)

A bonus is given on the final social score depending on the number of disclosed accounts \( n \). The maximal number of disclosed accounts is 5. The variable \( CONS \) stands for consistency. It varies between 0 and 5 and conveys the concordance between the different disclosed accounts. When the same username is used among the social media accounts of the different portals, \( CONS \) increases by 2. We increase the bonus by 2 again if the email address matches. For the same birthday information, the bonus is increased by 1 for consistency. The maximal number of points reached by the bonus is therefore 50.

\[ SC = \frac{(SC_1 + SC_2 + \ldots + SC_n)}{n} + bonus \]

(4)

The final equation is composed of the sum of the single social score’s average and the bonus. The complete formula would look like this:

\[ SC = \frac{(O_1 \times A_1 \times H_1 + \ldots + O_n \times A_n \times H_n)}{n} + n \times 2 \times C \]

(5)

In this equation, \( O \) is short for open [\( OPEN \)], \( A \) stands for authenticity [\( AUTH \)], \( H \) for honesty [\( HON \)] and \( C \) for consistency [\( CONS \)]. If the user does not disclose any accounts, every \( OPEN \) variable and \( n \) is zero. No social media based observation can be made as no data can be accessed. Since all the partial terms \( SC_1, SC_2, \ldots, SC_n \) contain a multiplication by zero, they all end up having the value 0. For \( n = 0 \) the bonus equals zero, and as a consequence, the final social score (\( SC \)) will add up to zero as well.

The highest social score that can be achieved this way is 100 and the lowest is 0. Note that a low social score in this system does not mean that the user is guaranteed to have bad intentions or cannot be trusted. It could also mean that the user is not very active on social media. In any case, it means that the user did not reveal much about himself and that his social media accounts do not give us sufficient reason to trust him.

In this situation, the user may obtain loans through traditional lending mechanisms such as using a collateral or his credit score.

V. IMPLEMENTATION

The information that is entered by the user and the calculated social score are saved on the blockchain. This information includes the user’s name, email address, date of birth, loan amount, and the social score. The clear disadvantage of saving this information on a public blockchain is lack of the user’s privacy. Future work should address this problem and provide a solution, for example, by using encryption and zero-knowledge proofs. For this prototype, we decided to make the data publicly accessible because we want the users to give up their anonymity. We also want the whole process to be as transparent as possible. In a future iteration of the
prototype, we may store only the less sensitive information on the blockchain, for example, only the loan amount and the social score. The changes required in the code would be minimal since `setInfos` is the only function that would need to be adapted.

A. Tools and Technologies

We decided to develop this implementation on an Ethereum test network to which we connect using Ganache (www.trufflesuite.com/ganache). The smart contracts for this project are written in the programming language Solidity on the Ethereum IDE Remix (remix.ethereum.org). To connect our smart contract with the Ganache blockchain, we use meta-mask (metamask.io). For the frontend of the implementation, we use HTML and JavaScript. The frontend connects to the backend and therefore we test our blockchain by using the library web3.js (web3js.readthedocs.io).

B. Smart Contract Implementation and Ethereum Gas Usage

Once the smart contract is deployed, one can use the functions within this smart contract while interacting with the blockchain. By calling a function in the smart contract that writes on the blockchain, for example the “setInfos” function, a block is mined and therefore gas must be paid. However, when one of the view-functions is called, no gas needs to be paid and no new block is mined on the blockchain. The “setInfos” functions is called only once when a person registers for the first time.

One main advantage of a smart contract is that the calculations performed are transparent and everybody can have trust in the calculated score. The smart contract calculates the score and then stores it in the blockchain by itself, without the intervention of any outside code. This way, everybody can trust in the process, including the user himself, and nobody can manipulate it. Figure 1 shows the listing of the functions in the smart contract implemented in Solidity.

Five of the functions in the smart contract write to the blockchain and therefore cost gas. The other functions are read- or view-only functions. The smart contract is deployed and written on the blockchain, which costs ether. The gas cost is 2141686. However, the smart contract only has to be deployed once. The functions within the smart contract, which cost ether because they write on the blockchain, are also called only once. This is done during the registration process. Not all of the functions are necessarily called. If the user does not connect his Facebook account, the function “calculateFBScore” is not called. The same is true for connecting the Instagram account and the LinkedIn account. If only one account is connected, there is no bonus added on top of the social score and as a consequence, the “calculateBonus” function is not called. The only gas costing function that is always called, is the “setInfos” function.

We also employ a helper function to compare strings. Further, we have one calculation function for each social media network and one function to calculate the bonus. None of the functions has a return value. They all work directly on the globally saved “socialScore” variable.

VI. Evaluation

In this section, we evaluate the execution of the social score function on the “Social circles: Facebook” real user dataset. The objective is to determine whether the analysis of real social datasets can help the operator set the parameters of a platform in production.

A. Setup of the Test Environment

To evaluate the function itself, we use real user data from the Stanford Network Analysis Project (SNAP). The dataset is called “Social circles: Facebook” [13]. To interpret the dataset and to make calculations based on it, we work with the Anaconda Prompt (docs.anaconda.com) and the python environment Jupyter Notebook (jupyter-notebook.readthedocs.io). Within jupyter notebook, we imported the libraries pandas (pandas.pydata.org) and networkx (networkx.org) as well as matplotlib (matplotlib.org).

B. Results and Observations

We used the 107.edges file from the SNAP dataset [13]. It contains 53498 edges (signifying friend relationships) and the corresponding nodes. The dataset is anonymized. As discussed earlier, the number of friends influences the authenticity of a person. Evaluating the data shows that more than a third of all users have more than 50 friends and therefore have a chance to get the best social score if we set the threshold to this value. On the other hand, there are also almost 15% users who do not have more than ten friends and consequently get the worst result in this category. The results are shown in figure 2. In our evaluation, we only differentiate between...
three steps concerning the number of friends: 10, 30 and 50. Since more than a third of the users have enough Facebook friends to get the best result possible in the “amount of friends” category, we evaluate if a higher limit would be more suitable. For testing reasons we will adapt these limits. We will start with 200 and then go down to 150, 100 and lastly to 50. The amount of users tested in this experiment is 1034. When we set the threshold to 200, which means that the user needs more than 200 Facebook friends to get the best result in this category, only 12 of the 1034 users qualify for the best result. Subsequently, we obtain the result of 50 users qualified for a threshold of 150, 163 users for a threshold of 100, and finally 389 users for a threshold of 50. The results of this last experiment are plotted in the figure 3. In this experiment, we see that by analyzing real datasets, we can set the thresholds for the platform in order to correspond to the desired rate of users who should qualify. The number of friends is only a small part of the final social score. Other factors include the number of accounts connected, number of pictures posted, account creation date, the bonus for connecting 5 accounts, etc. These factors could not be considered in this experiment due to the absence of this information in the dataset.

VII. CONCLUSION

In this paper, we have presented a new approach to calculate the users’ trustworthiness on peer-to-peer lending platforms. This approach is neither collateral nor credit score based. The formula that we use to calculate a user’s social score relies on the social capital theory and consequently the information retrieved from the user’s social network accounts. It considers how well connected a user is, since connections may help the user achieve professional and personal success. The social score is implemented in a smart contract running on the Ethereum blockchain. The whole lending process is automated and does not need any input from a middleman. The presented approach offers a new method to verify users’ trustworthiness regarding lending, where even users with a non-sufficient credit score and no valuable assets as collateral could get a chance on a loan. We quantified the amount of gas that is consumed by the deployment of the smart contract. Moreover, we also evaluated the execution of the function on a real social network dataset. This experiment demonstrated how we can use analysis of social network data to determine optimal thresholds for a platform in production. In terms of future work, we consider user privacy a concern that needs to be addressed in subsequent iterations of the platform.

REFERENCES