Incentivizing Honesty in Online Decentralized Markets

Tiago Lucas Pereira Clementino(✉) and José Antão Beltrão Moura

Federal University of Campina Grande, Campina Grande, Paraíba 58425-900, Brazil
tiagolucas@copin.ufcg.edu.br
https://portal.ufcg.edu.br/

Abstract. The Buyers and Sellers Dilemma is an illustration of the risk that every buyer takes when advancing payment for a product or service in a transaction that cannot be verified beforehand. That is the case of transactions involving a buyer who is physically distant from an unknown or anonymous seller as in online, e-commerce scenarios. To increase gains, the seller may be compelled to act dishonestly and not deliver the goods. It is thus of interest to implement solutions that incentivize honesty in said scenarios. For that, this article carried out a literature review in search of decentralized solutions to encourage honesty in non-verifiable transactions. Based on criteria of decentralization, non authentication and non-verifiable transactions tolerance, the solutions found were classified, analyzed and compared based on data from real world operations extracted from the OpenBazzar platform. We ranked the solutions according to effectiveness, although there is still no comprehensive solution that promotes honest behavior under any circumstances.

Keywords: buyers' and sellers' dilemma · decentralization · honesty incentive · decentralized markets · non-verifiable transactions

1 Introduction

A transaction can be seen as a structured protocol composed of a series of operations that facilitate the exchange of values between two parties: an active initiator and a passive responder. In this configuration, the active party (usually, but not necessarily, the buyer) assumes the risk by executing the value transfer (payment) before receiving a response (product or service), leaving room for potential dishonesty on the part of the passive party (seller). Risking transferring values in advance constitutes the Buyers and Sellers Dilemma [3]

Such a dilemma involves a particularly pronounced risk in environments with unverifiable operations, such as in decentralized markets that involve physical currency exchanges. According to Kutera et al [7], the US Federal Trade Commission (FTC, 2021) reported that from October 2020 to March 2021, nearly seven thousand users were victims of digital fraud in the cryptocurrency market due to unfair mediation, causing losses around US$ 80 million.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
O. Gervasi et al. (Eds.): ICCSA 2024, LNCS 14814, pp. 213–229, 2024.
https://doi.org/10.1007/978-3-031-64608-9_14

Previous solutions to the Buyers' and Sellers' Dilemma in decentralized markets incorporate costly preconditions that compromise the applicability of the solution. Examples include security deposits [3], arbitration [1,6,9,19], trust assurance in any of the parties involved [3,8], controllable digital content [11,14,20] and payment protocols aimed at products with digitally verifiable authenticity [13]. Such preconditions ensure the verification of all operations throughout the transaction protocol between buyer and seller.

In decentralized markets, two paths can be followed with the aim of reducing such risk: encouraging honesty or validating all operations [3,8,10,11,14,19,20]. A good example of a decentralized market with non-verifiable operations is the e-commerce platform OpenBazaar, which although discontinued its concept remains quite relevant in the literature [4]. This article presents an analysis of the effectiveness of its honesty incentive models in comparing to other honesty incentive models from the literature. Comparison is carried out here on a dataset of OpenBazaar's transactions (purchases) logs from June 15, 2018 to September 3, 2019 collected by Arps and Christina [2].

The research presented here consisted of three phases: 1) Reviewing the literature to collect major decentralized models whose aim is to encourage honesty in transactions with unverifiable operations; 2) Manually tagging transactions as success or failure in OpenBazaar's historical dataset according to annotated reviews and comments on corresponding purchases; and, 3) Comparing such models based on their predictions of success or failure of transactions we were able to infer from the tagged historical dataset. Metrics used to compare these models were inspired by Machine Learning and they are: precision, recall and F1.

For the comparison (phase 3), three main classes of solutions where collected from the literature: decentralized arbitrator, trust network and guarantee deposits. This work extracted from these solution classes the most promising features according to the criteria of decentralization, non-verifiable transactions tolerance and non authentication.

The related Research Question (RQ) for which this paper brings an answer is: *Which of the classes or combinations of their main features present superior performance?*

The results demonstrate that although there are models capable of promoting honesty better than those used in OpenBazaar, the decentralized buying and selling game remains biased towards dishonesty. Further, we find evidence pointing to the solution composed of trust network and decentralized arbiter as the most effective of all. These findings contribute to the research, design, operations and management of smart contracts, e-commerce platforms and decentralized markets in general.

2 Buyers' and Sellers' Dilemma

The Buyers' and Sellers' Dilemma (BSD), a term originally used in the book Prisoners' Dilemma [15], is a racing problem between dishonest players engaged in establishing a trade based in two operations (payment and transference of

product, for example), both with advantages in opposite directions. Such transaction is a non equilibrated game, as each party tends to look after its own interests only by betraying the other party, since it offers the most advantageous outcome.

Consider a buyer and a seller, or the active player and the passive player, respectively. The active player is the one who proposes the transaction and assumes the risk of delivering its value first. The passive player is the one who receives the value from the active player and for a moment has in hand its original value (the product or service it is supposed to deliver) and the value of the other. At this point s/he must decide whether to act honestly and deliver her/his consideration or not. This dilemma is based on the assumption that the passive player will always prefer to act dishonestly, and the reason is described below.

One of game theory principles describes the concept of Nash Equilibrium [12], which refers to games in which there is a strategy to be followed by a player in order to always ensure the best possible result no matter the strategies of the other player. This 'safe' strategy is the Nash Equilibrium point.

Definition 1 (Buyers' and Sellers' Dilemma – BSD). *Consider the game G, where:*

1. *P consists of the set of exchanged values/products, and $x, y \in P$;*
2. *Q corresponds to the set of players negotiating (generally 2, but it will use $|Q|$ for generalization), and $B, S \in Q$;*
3. *$S^i = \{s^i_1, ..., s^i_{|S^i|}\}$ corresponds to a set of strategies or actions of the player i;*

In the simple game described above and represented in Fig. 1, concluding the negotiation is impossible, as it is advantageous for the buyer (B) never to pay (transfer x) for the item (y), leading the seller (S) to not deliver it. If the buyer does not pay for the product, s/he may retain only the payment value or the payment and the product/service (if the seller delivers the product/service even though s/he has not yet been paid). If the buyers pays for the product/service in advance, s/he may still receive the seller's product, but s/he runs the risk of being left with nothing. The sellers, if s/he does not sent the product (transfer y), can keep just the product/service or the product/service and the received payment; if s/he sends, s/he can still keep the received payment, but runs the risk of being left with nothing. In the just described possibilities, it is always more advantageous to be dishonest.

In despite of the BSD, mutual responsibility and the presence of intermediaries like eBay and Amazon, that for a premium, reimburse the buyer in case of non-delivery of paid for goods, guarantee security in centralized markets. However, legal repercussions or centralized trust intermediaries depend on the verification of all operations and the centralization of authority, conditions that solutions from literature attempt to circumvent.

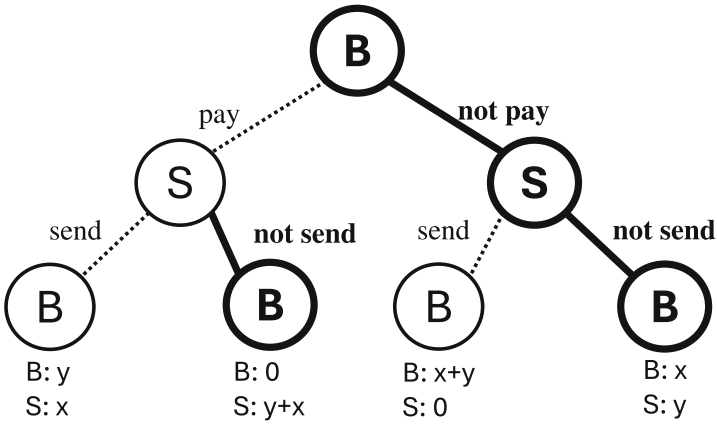


Fig. 1. Representation of the decentralized BSD as an extensive game. The dominant paths are highlighted. It is observed that the dominant strategy always represents better results for the player, buyer (B) or seller (S)

3 Literature Review – Phase 1

This section brings a literature review on solutions solutions applied to BSD. The review was carried out from December, 2023 to January, 2024. The results are divided into three main classes of solutions: trust networks, arbitration and guarantee payments protocols. These three classes are collectively and generally called Honesty Incentive Models (HIMs). Here we also look into combinations of models from more than one class. Our comparative experiment explores the major features of each model class.

Zindros [21] presents the real world OpenBazaar platform, that implements a specific approach of two classes, arbitration and trust network. In OpenBazaar any user can act as a human arbitrator in transactions and this must resolve disputes when a problem occurs in exchange for a monetary reward that is only paid if the referee is called upon. The trust network is based on feedback and the trust in an arbitrator or player depends on such feedback and a text file CV that almost always exposes the arbitrator's identity and the arbitrator is chosen only by the seller.

3.1 Trust Network

The approach introduced by Duong-Trung et al [6], besides automatic conflict resolution, penalizes couriers or carriers who breach contracts based on a reputation system. In the literature there are other works that address decentralized solutions for Cash on Delivery (CoD) and address the BSD in a similar way [1,3,9], although all of them utilize a trust network associated to other models.

3.2 Arbitrator

Son et al [17] employ blockchain and smart contracts to enhance the Cash on Delivery (CoD) protocol for e-commerce. Their approach ensures secure transactions and resolves the BSD by imposing specific rules on players and utilizing smart contracts as intermediaries. Tsabary et al [19] introduced a similar solution called MAD-HTLC, which leverages blockchain miners as contract arbitrators.

Duong-Trung et al [6] introduce a comprehensive CoD model that relies on smart contracts for automatic conflict resolution, eliminating the need for human intermediaries who, according to them, would consume more assets and time for the players involved. However, this approach includes an additional conflict resolution element, penalizing couriers or carriers who breach contracts based on a reputation system. According to the authors, the use of a Smart Contract as an arbiter associated with the reputation system, in addition to the infrastructure provided by decentralized technologies that implement the Smart Contract, is sufficient.

3.3 Payment Guarantee

Le et al [8] introduce a CoD protocol for commercial delivery where drivers are required to mortgage a sum of money for guarantee. More over, an authentication system is used for drivers and a hash code identifies the products.

Tsabary et al [18] introduce the concept of LedgerHedger, a two-party mechanism that ensures timely confirmation of transactions in smart contract protocols. LedgerHedger has the issuing player pay for a transaction in advance, and the other player agrees to pay a required fee, even if it exceeds the value of the transaction. This ensures that the transaction will be confirmed within a defined time frame. Asgaonkar and Krishnamachari [3] also present a solution based on double guarantees. Such double guarantees, although effective, can make transactions unfeasible, especially when they are greater than the negotiated value [18].

Finally, the most complete model of payment guarantee protocols was provided by Schwartzbach [16], where they present two main conclusions: it is not viable to provide a guarantee deposit model without any type of information about player behavior, and it is unlikely that a non-adaptable guarantee deposits protocol based on empirical rules will work. The mathematical approach presented seems the most complete.

3.4 Overall

The difficulty in virtualizing real-world services, cash and products in a decentralized manner is the core problem of non-verifiable transactions and leads most solutions to the BSD to resort to preconditions that limit the application domain or violate the principle of decentralization [3,8,10,11,13,14,19,20]. Table 1 presents all solutions collected from the literature and identifies the preconditions applied by each one to avoid dishonesty. Such preconditions are diverse, but can be classified into four categories, they are:

- Some level of centralization: Models appeal to centralization to verify the transaction in some level, overcoming the difficult of ensure trust without complete information;
- Authentication: Identifies users to intimidate the dishonest ones with legal repercussions;
- Digital verification of delivered products: In a decentralized environment it is applicable only to products that may be completely digitalized;
- Feedback: It is used to build trust, but is more susceptible to fails because may be forged.

Features used for avoiding such preconditions are used to compose the models compared here and described in more details in Sect. 5. As one can see from Table 1, none of the solutions implements exclusively trust network, although without such trust network, some other kind of behavior inference must be provided (as smart contract oracles, for example).

Table 1. Solutions for BSD from the literature

Reference	Model	Centralization	Authentication	Digital Verification	Feedback
[8]	Guarantee deposits	✓	✓	✓	
[18]	Guarantee deposits			✓	
[3]	Guarantee deposits, Trust network			✓	✓
[16]	Guarantee deposits				
[17]	Arbitrator			✓	
[19]	Arbitrator			✓	
[6]	Arbitrator, Trust network	✓	✓	✓	✓
[9]	Arbitrator, Trust network	✓	✓		✓
[1]	Arbitrator, Trust network			✓	

4 Tagging OpenBazaar Historical Data – Phase 2

4.1 OpenBazaar

OpenBazaar, a decentralized e-commerce marketplace, received significant attention from its initial release in 2016 until its closure in 2020. When deciding to use a decentralized platform like OpenBazaar, you should be aware that you are on your own in several aspects. Although the interface sends the same consumer message as centralized markets, in OpenBazaar payment is made by cryptocurrency directly to the seller and there is no refund, nor responsibility on the part of the platform.

It is important to note that OpenBazaar already implements two of the classes of encouraging honesty analyzed here. However, the arbitration model applied is biased, with the seller (BSD's passive player in this case) being solely responsible for choosing a list of possible arbitrators, a contradiction. The reputation model is based on feedback and scoring, without a trust network. Thus, one

can compare the effectiveness of the honesty incentive solution applied (arbitration and reputation based on scoring and feedback) with the approaches selected from the literature: decentralized arbitration, trust network without feedback and payments guarantee.

4.2 Dataset

Using multiple daily traces of the OpenBazaar network over approximately 14 months (June 25, 2018 to September 3, 2019), Arps and Christin [2] observed its evolution over time. A total of 6,651 different participants were observed. More than half of all users (3521) were only observed on a single day, and on average only approximately 80 users were active simultaneously on a given day seeking for a total of 4379 listed items. 6292 text messages associated with purchases were also observed.

4.3 Manually Tagging Transactions

From the 6292 text messages, the success or failure of their corresponding transactions was manually inferred and tagged, message by message, yielding a total of 2517 successful transactions and 347 unsuccessful transactions that sums up to a total of 3034 manually tagged transactions; the success or failure of 170 transactions was not possible to infer.

5 Honesty Incentive Models – HIMs

This section describes the strategies behind each honesty incentive model (HIM) we compare. Here it is presented how each model was implemented in order to evaluate its effectiveness in encouraging honesty. Since it is not possible to implement incentive to honesty without any behavior inference strategy [16], decentralized arbitrator and payment guarantee were implemented together with trust network in order to use trust network as a souce of behavior inference.

5.1 Trust Network – T

The reputation network implemented in our comparison describes a more decentralized concept, Web-of-Trust, where there is no feedback (less susceptible to fails) and transactions between users are public, giving each user the freedom to make their own conclusions. Originally, Web-of-Trust deals with a p2p network to establish the authenticity of the connection between a public key and its owner. Over time, users accumulate keys from others they want to designate as trusted introducers, and a trusted network is gradually formed without a centralized certification authority. Users legitimize each other by accumulating and redistributing a collection of third-party certificates. The public key network formed is a flexible, decentralized and fault-tolerant scheme, verified through consensus among users [5].

For transactions with non-verifiable operations, the Web-of-Trust model is capable of providing trust between buyers and sellers who have never had direct interactions based on the experience of those with whom they interacted separately.

Two players $a, b \in A$ evaluate the possibility of carrying out a transaction T. To this end, the network's trust in a given player b is ϕ_b, according to Eq. 1.

$$\phi_b = \frac{\ln\left(\frac{\sum_{a \in A} s_t(b,a)}{P_t}\right)}{\ln\left(\text{MAX}_i\left(\frac{\sum_{a \in A} s_t(i,a)}{P_t}\right)\right)} \tag{1}$$

Where $s_t(b, a)$ refers to the historical balance of transactions between b and a up to the current moment t. Here, balance is understood as the sum of the values in products and currency of the other party in each transaction – thus, dishonest operations by the player contribute negatively to the balance. Finally, P_t represents the total balance of transactions among the entire active population of OpenBazaar up to time t. Here, 'time' means an arbitrary time interval at the end of which trust is recalculated for all players.

Thus, ϕ_a grows as the total balance of a transactions grows in relation to the total balance of the population as a whole. For normalization purposes, it is taken the best reputation among all players $\phi = 1.0$. This best reputation ϕ is represented in Eq. 1 by MAX_i playing the role of normalizer, preserving $\phi \in [0.0, 1.0]$. As the best reputation among all players increases, the others fall in a logarithmic proportion.

Also considering an acceptable trust threshold to carry out a δ transaction, observe Eq. 2.

$$T_{s,b} = \begin{cases} 1 & \text{if } \delta_t < \phi_s \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Thus, for transaction T to occur, ϕ must be greater than a threshold δ for passive player s (seller). This factor is proportional to the total historical balance of a given player s and inversely proportional to the total balance of the population. Here, trust in the buyer is irrelevant because OpenBazaar's business model leaves no room for dishonesty on the part of buyers as they always transfer their values first.

5.2 Trust Network + Arbitrator – TA

Arbitration is based on the selection of the arbitrator through consensus between the parties and was implemented making the Web-of-Trust proposed in the literature be the source of trust inference. This is because for there to be consensus in the selection of the arbitrator, both parties, even separately, must trust the arbitrator.

Web-of-Trust model works well with the moderation of transactions using reliable decentralized arbitrators, since, given a population Q with a few

transactions $S' << Q$, it is already possible to estimate the honesty of players involved in a transaction (buyer, seller and arbiter).

Consider Eq. 3. Now three players a, b seeking the transaction, and c candidate for arbitration, where $\gamma_{a,b}(c)$ is the function that defines whether c is trustworthy enough to be the arbiter of transaction $T_{a,b}$.

$$\gamma_{a,b}(c) = \begin{cases} 1 & \text{if } \delta_t < \frac{s_t(c,a)}{\sum_{a \in A} s_t(c,a)} \text{ and } \delta_t < \frac{s_t(c,b)}{\sum_{a \in A} s_t(c,a)} \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Moreover, to define among the entire population who is the best referee for the operation, see Eq. 4:

$$\gamma_{a,b} = \begin{cases} c & \text{if } \forall_{c \in A} \delta_t < \text{MAX}(\frac{s_t(c,a)}{\sum_{a \in A} s_t(c,a)}) \text{ and} \\ & \delta_t < \text{MAX}(\frac{s_t(c,b)}{\sum_{a \in A} s_t(c,a)}) \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Which describes an operation of asymptotic complexity $O(n \log n)$, to discover the best arbiter for all operations, with n being the total number of new transactions at each complete time frame t.

5.3 Trust Network + Payment Guarantee – TP

For the collateral deposit protocol, we borrow from the work of Schwartzbach [16], in which the author assumes a decentralized market based on non-verifiable transactions. To this end, Schwartzbach [16] demonstrates the need for some trust inference mechanism between the parties, such as oracles (smart contracts with external access) used in distributed ledgers. Such need was fullfiled by an association to the Trust Network model, as with Decentralized arbitrator Model (see Sect. 5.2

Using the Trust Network to play the role of a source of trust inference as oracles do in Distributed ledgers, consider that two players a, b, buyer and seller, wish to exchange the values x, y, cash and product. According to Schwartzbach [16], the security deposits necessary to guarantee the transaction between a and b are x', y', are given Eqs. 5 and 6.

$$x' = \frac{2\phi_a}{2\phi_a - 1} \, x \tag{5}$$

$$y' = \frac{\phi_b}{2\phi_b - 1} \, y \tag{6}$$

This approach was chosen because it is satisfactorily decentralized and capable of dealing with the uncertainty of an incomplete information environment. Here incomplete information may be considered as the absence of information about the success or failure of transactions (non-verifiable transactions).

6 Experiment Results – Phase 3

Although it achieved some success, OpenBazaar was short-lived when compared to centralized models like Amazon or eBay. Such a fate may have been anticipated due to OpenBazaar's difficulty in encouraging honesty on the system, falling into BSD. In total, of the 3034 transactions observed, approximately 12% were unsuccessful, that is, the seller received payment but did not deliver the products/services, the only form of dishonesty observed. There are still 170 unclassified transactions left, so this proportion could be even more serious. This rate is very high when compared to non-anonymous centralized models, supported by laws and payment intermediary models that make dishonesty within the site very rare.

However, those responsible for OpenBazaar did not ignore the problem and implemented two solutions to verify transactions and encourage honesty: arbitration and score-based reputation. In this section we use the most promising decentralized solutions according to the revised literature, here named T (Trust Network), TA (Trust Network plus Decentralized Arbitrator) and TP (Trust Network plus Payment Guarantee), to predict which transactions would be successful and which would fail, and then checking against the OpenBazaar's manually tagged historical data. With this, we obtained conclusions regarding the evaluated models and the solutions actually implemented in OpenBazaar[1].

6.1 Trust Network only – T

The trust network model based on transaction history without feedback was the one that obtained the best results in avoiding BSD in decentralized environments, according to the literature among the reputation models [1,3,9]. For this experiment, this model also serves as a source for inference to try to estimate the trust of each player, in addition to being compared to the others models.

Notice in Fig. 2a, the total number of unsuccessful transactions (lighter red line) decreases as the δ threshold increases. This is because a bad seller is likely to have a bad reputation, and increasing the threshold predicts the failure of transactions involving sellers with a bad reputation. With the threshold δ very low, most transactions are classified as safe; if this threshold is too high, the model points out almost all transactions as insecure. If the δ threshold is too high, successful transactions may be affected. This is depicted by the rising blue line. For a $\delta >> 1.0$, most transactions are classified as insecure.

The optimal threshold point δ is shown in Fig. 3a, where the precision and recall metrics cross F1, at approximately $\delta = 0.7$. The green, red and blue lines in Fig. 3a represent the f1, precision and recall metrics, respectively.

[1] All results presented here were obtained from the implementation in R language of the proposed models. All source code is openly available at https://github.com/tiago-clementino/honesty_in_openbazaar.

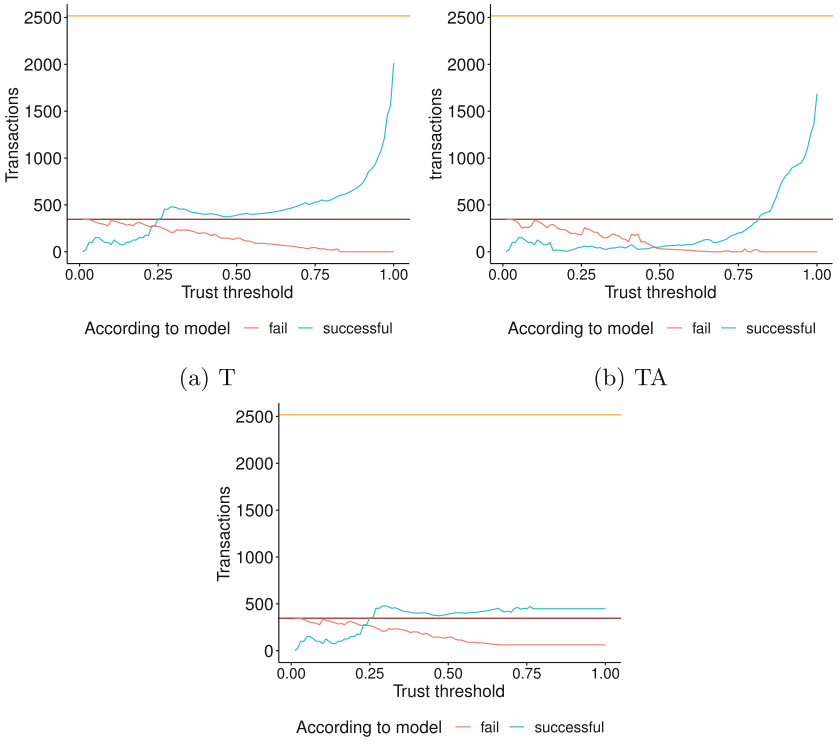
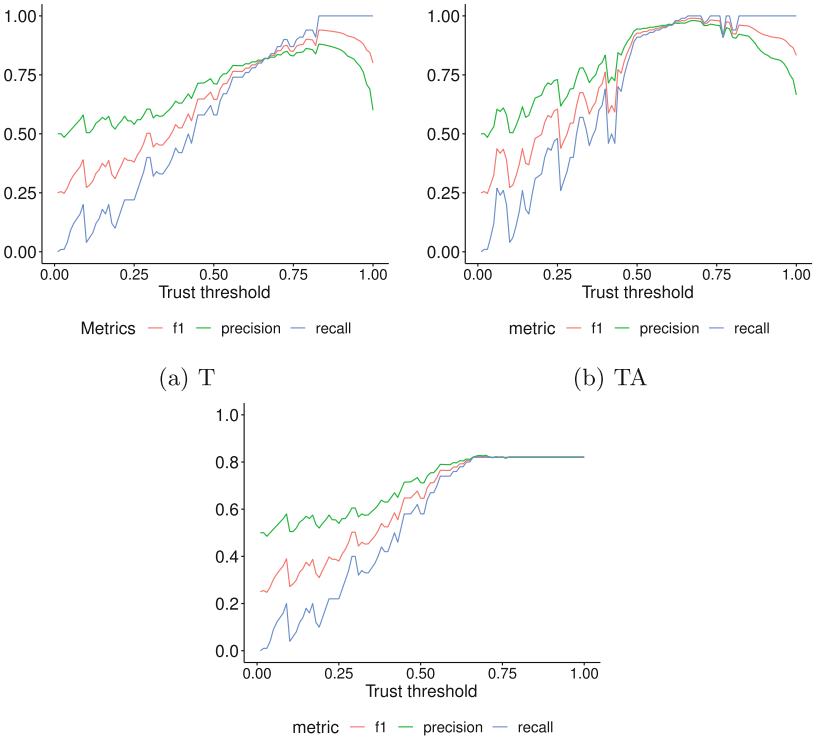


Fig. 2. Transactions per trust threshold. Where horizontal dark red line presents the total failed transactions (347 products not delivered), while the horizontal orange line represents the total successful transactions (2517 products delivered) (Color figure online)

6.2 Trust Network + Decentralized Arbitrator – TA

The mechanism behind the decentralized arbiter plus trust network (TA), instead of using buyer-seller mutual trust determines that buyer and seller reach consensus on a third player they both trust. The buyer then transfers payment to the consensual chosen third player who withelds payment to the seller until s/he delivers the product/service. Based on that, this model achieves the best results according to Fig. 3b, where the intersection of f1, recall and precision occurs at 0.965, the best value, and in a trust threshold δ of 0.607, a feasible value.



(a) T

(b) TA

(c) TP

Fig. 3. Horizontal axis presents the trust threshold, and vertical axis represents the magnitude of each metric (precision, recall and f1)

The largest results of all model compared are numerically described in Table 2, where the lines represents the metrics and the columns represents the models. The first metric 'Insecurity fail transactions' represents all transactions where the product delivery did not occurs as expected because some dishonest behaviour of the seller. The larger this metric is, the better the model in preventing sellers dishonesty. The second metric is the 'Insecurity successful transactions' which means those successful transactions were wrongly predicted as insecurity by the model. The lower this metric is, the better the model. The third row represents the intersection of the metrics f1, precision and recall, pointing to the best trust threshold of each model. Finally, the last row represents the trust threshold δ where the intersection of metrics f1, recall and precision occurs.

Table 2. Better metrics of each solution compared (T, TA and TB)

Metrics	T	TA	TP
Insecurity fail transactions	287(82.71%)	336(96.83%)	287(82.71%)
Insecurity successful transactions	461(18.31%)	79(3.14%)	445(17.68%)
F1 = recall = precision	0.821	0.965	0.823
Threshold δ	0.666	0.607	0.663

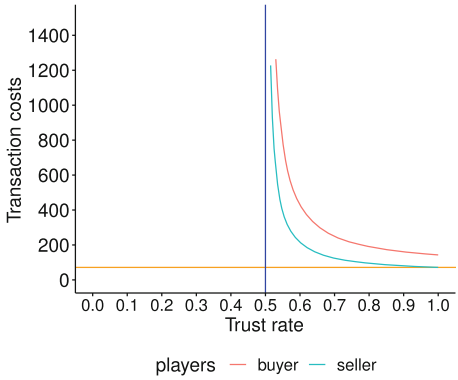


Fig. 4. Cost of payment guarantee (in USD) according to trust rate of seller. The horizontal orange line represents the average transaction amount. The vertical dark blue line represents the trust rate $\phi = 0.5$ (Color figure online)

Looking again at Table 2, you can see numerically that the largest f1, recall and precision intersection is 0.965 for the TA model. For all other metrics, the TA model is significantly better than its competitors.

6.3 Trust Network + Payment Guarantee – TP

Its is important to note that this approach was replicated from the work of Schwartzbach [16]. Schwartzbach concluded that its difficult to provide a viable and effective payment guarantees protocol based on empirical rules, and that design such a protocol without any information about the behavior of players is an intractable problem. This last conclusion motivated us to propose the use of the trust network model as the source of information about the players behavior, as in Arbitration model (TA).

Although such HIM appears to be an effective strategy since the collateral values are always at least equal to the values gained through dishonesty. But there are two drawbacks to consider: 1) First, the protocol is biased towards sellers who can pay as collateral in some situations the same values as the values of their products, but have the chance to end up with their products and the buyer's payment, according to OpenBazaar's business model. 2) As you can see

in Fig. 4, which shows the growth of transaction costs (payment guarantee costs in USD) by the trust rate of buyers and sellers, the payments protocol is not even calculable if the trust rate confidence is less than $\phi <= 0.5$ (dark blue vertical line). Otherwise, for higher trust rates, but close to 0.5 (like 0.55), the collateral values are so large that they make such a payment impossible (see red and light blue lines for buyers and sellers respectively).

Therefore, we restrict guarantee payments according to a hybrid model: as long as the trust threshold δ be lower than the reputation/trust rate necessary for the seller to pay more than double its value[2], it is assumed that payment guarantees are not viable and the model falls into the trust network model. This means that if the seller needs to pay more than twice the value of its products as collateral, such collateral will be too expensive to continue the transaction. Since the buyer does not have the opportunity to behave dishonestly, according to the OpenBazaar business model, it is not necessary to demand payment guarantees from him/her.

Going back to Fig. 2c, you can see that the payments guarantee protocol avoids the explosion of fake positives for very large values of δ, where the majority of true transactions are marked as insecurity (light blue line), but it happens after its intersection between f1, precision and recall point (see Fig. 3c). So that, even with serious questions of viability, the payment guarantees protocol plus trust network remains almost as effective as the raw reputation network, as one can see in Fig. 3c and Table 2, where the intersection point of F1, recall and precision are almost the same in both. Because of that, the greatest results (fail transactions marked as insecurity greater and success transactions marked as insecurity lower, according to Table 2) are also almost the same in both models.

Looking again at Fig. 3, we can see that even the least efficient models (T and TP) manage to predict the majority of unsuccessful transactions with sufficient large threshold (light red line below the dark red line). Note also that the HIMs actually applied on the OpenBazaar platform (dark red line in Fig. 1) – i.e., unilateral arbitration and reputation based on feedback, proved to be less effective than all the solutions tested here.

6.4 Answering the RQ and Threats to Validation

From the results in Subsects. 6.1, 6.2 and 6.3, the overall superior performance in terms of honesty incentives and also considering f1, precision and recall is TA. That's the short and simple answer to our Research Question. For a more detailed answer however, one has to take into account situations where T or TP appear superior or even more feasible – e.g., that is the case when honesty is the predominant behavior and any costly mediation strategy is needed – in this case, T appears more applicable; or, on the other hand, when honestly presents "average, typical behavior" and the transactions are critical and must be ensured – in this case, TP is a better choice.

[2] seller's + buyer's values assuming, for analysis' tractability, that the values of the buyer and the seller are approximately equal.

Because our results were instrumental to indicate superior HIMs than those applied on the OpenBazaar platform, they should be considered together with an important threat to validity. The analysis presented here is based on historical data and as such, it cannot be used as an impact gauge of the compared solutions for they were not in actual usage.

7 Conclusions and Future Work

7.1 Conclusions

This study compared honesty incentive models (HIMs) elicited from a literature review that was guided by the requirements of decentralization, tolerance for non-verifiable transactions (incomplete information) and non-authentication (anonymity), based on metrics of precision, recall and F1. The elicited HIMs were: trust network without feedback (T), decentralized arbitration (TA, associated to trust network) and payment guarantee (TP, also associated to trust network). The trust network without feedback plays a special role, serving as a source of information regarding the behavior of players. Said role is commonly performed by decentralized oracles in distributed ledger technologies (blockchains, for example).

The comparison was made using the elicited models to predict unsuccessful transactions (products not delivered or delivered with problems) on a manually labeled dataset extracted from the decentralized e-commerce platform OpenBazaar. Such predictions are based on how much each model is able to classify the security of each transaction, and based on an ideal confidence threshold θ to predict the actual successful and unsuccessful transactions. Such an ideal threshold is one where the precision, recall and F1 metrics intersect. In addition to the comparisons presented here, the advantages and disadvantages of each model are described, as shown in Table 3.

Table 3. Advantages and disadvantages of each model

Models	Advantages	Disadvantages
T	By itself is a good solution, but it may be improved if associated to other solutions	Provide information about behavior of players but not ensures the transaction
TA	The solution that outperformed the others if associated to behavior inference system such as a trust network	It does not work by itself
TP	When viable it is almost infallible	It does not work by itself and most of the time it is not viable because it requires a high trust rating to do so

In conclusion, it was observed that the application of the trust network without feedback associated with decentralized arbitration (TA, see Table 2) showed significantly superior performance than the others based on metrics of recall, precision and F1. The OpenBazaar's original HIMs were found to perform more poorly when compared to the elicited HIMs. TA, the solution that outperformed the others, reached $F1 = 0.965$ (see Table 2), which is quite high.

7.2 Challenges for Future Work

Despite the consistency of the results, there is still no solution in the literature that actually solves the Buyers' and Sellers' Dilemma (BSD) in an anonymous, decentralized environment involving non-verifiable transactions. This finding emphasizes the ongoing nature of the problem and the need for further research and innovative solutions in this domain.

Albeit the above gap in research results, the results reported here contribute to Blockchain and DLTs applications specifically in the field of fraud prevention protocols for online decentralized and anonymous markets dealing with non-verifiable transactions. The main contribution of this study is the provision of a synthesized basis, forming a body of knowledge, for future reference, development and research on the treatment of the BSD in full non-verified, anonymous and decentralized markets.

References

1. AlTawy, R., ElSheikh, M., Youssef, A.M., Gong, G.: Lelantos: a blockchain-based anonymous physical delivery system. In: 2017 15th Annual Conference on Privacy, Security and Trust (PST), pp. 15–1509. IEEE (2017)
2. Arps, J.E., Christin, N.: Open market or ghost town? The curious case of OpenBazaar. In: Bonneau, J., Heninger, N. (eds.) Financial Cryptography and Data Security: 24th International Conference, FC 2020 , Kota Kinabalu, Malaysia, February 10–14, 2020 Revised Selected Papers, pp. 561–577. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-51280-4_30
3. Asgaonkar, A., Krishnamachari, B.: Solving the buyer and seller's dilemma: A dual-deposit escrow smart contract for provably cheat-proof delivery and payment for a digital good without a trusted mediator. In: 2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), pp. 262–267. IEEE (2019)
4. Bhawna, Gupta, P., Rai, P., Chauhan, A.: Blockchain application in consumer services: a review and future research agenda. Int. J. Consum. Stud. **47**(6), 2417–2450 (2023)
5. Caronni, G.: Walking the web of trust. In: Proceedings IEEE 9th International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WET ICE 2000), pp. 153–158. IEEE (2000)
6. Duong-Trung, N., et al.: Multi-sessions mechanism for decentralized cash on delivery system. Int. J. Adv. Comput. Sci. Appl **10**(9) (2019)
7. Kutera, M., et al.: Cryptocurrencies as a subject of financial fraud. J. Entrepreneurship, Manage. Innov. **18**(4), 45–77 (2022)

8. Le, H.T., Le, N.T.T., Phien, N.N., Duong-Trung, N., Son, H.X., Huynh, T.T.: Introducing multi shippers mechanism for decentralized cash on delivery system. Int. J. Adv. Comput. Sci. Appl. **10**(6) (2019)
9. Le, N.N.T., et al.: Assuring non-fraudulent transactions in cash on delivery by introducing double smart contracts. Int. J. Adv. Comput. Sci. Appl. **10**(5), 677–684 (2019)
10. Mamageishvili, A., Schlegel, J.C.: Optimal smart contracts with costly verification. In: 2020 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), pp. 1–8. IEEE (2020)
11. Müller, M., Janczura, J.A., Ruppel, P.: Decoco: blockchain-based decentralized compensation of digital content purchases. In: 2020 2nd Conference on Blockchain Research and Applications for Innovative Networks and Services (BRAINS), pp. 152–159. IEEE (2020)
12. Nash, J.F., Jr.: Equilibrium points in n-person games. Proc. Natl. Acad. Sci. **36**(1), 48–49 (1950)
13. Pande, S.S., Mandolikar, S., Shitole, S.: Bitland-a decentralized commercial real estate platform. In: 2022 IEEE Bombay Section Signature Conference (IBSSC), pp. 1–6. IEEE (2022)
14. Radhakrishnan, R., Ramachandran, G.S., Krishnamachari, B.: Sdpp: Streaming data payment protocol for data economy. In: 2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), pp. 17–18. IEEE (2019)
15. Rapoport, A., Chammah, A.M.: Prisoner's dilemma: A study in conflict and cooperation, vol. 165. University of Michigan press (1965)
16. Schwartzbach, N.I.: Payment schemes from limited information with applications in distributed computing. In: Proceedings of the 23rd ACM Conference on Economics and Computation, pp. 129–149 (2022)
17. Son, H.X., et al.: Towards a mechanism for protecting seller's interest of cash on delivery by using smart contract in hyperledger. Int. J. Adv. Comput. Sci. Appl. **10**(4) (2019)
18. Tsabary, I., Manuskin, A., Eyal, I.: Ledgerhedger: Gas reservation for smart-contract security. Cryptology ePrint Archive (2022)
19. Tsabary, I., Yechieli, M., Manuskin, A., Eyal, I.: Mad-htlc: because htlc is crazy-cheap to attack. In: 2021 IEEE Symposium on Security and Privacy (SP), pp. 1230–1248. IEEE (2021)
20. Zhang, P., Wei, J., Liu, Y., Liu, H.: Proxy re-encryption based fair trade protocol for digital goods transactions via smart contracts. arXiv preprint arXiv:2306.01299 (2023)
21. Zindros, D.: Trust in decentralized anonymous marketplaces. National Technical University of Athens (2016)