Online Auction Transactions as the Source of Information about the Quality of Sellers'

Artur Strzelecki

The Karol Adamiecki University of Economics in Katowice
Information Science Department
strzelecki@ae.katowice.pl

Abstract. The aim of the paper is to present the results of the research on online auction transactions. The research is based on users' comments issued after committed auctions. The author presents a method for collecting data obtained from the heterogeneous Web, concerning the results of committed auctions. In the analysis of the obtained data the modified versions of the recently proposed sellers' positive and negative reputation measurement were administered. In order to mark both ways of measurement, a graph that presents the strength of the link of auction users must be constructed. The author states that the presented results are the source of valuable information about the quality of sellers' services. All the stated theses are supported by the series of experiments in the extensive data set.

Keywords: We would like to encourage you to list your keywords in this section.

1 Introduction

Online auction transactions are becoming more and more popular in the Internet trade services. The main advantages of online auctions are attractive prices of goods and services, usually lower than market prices of the same goods available in traditional shops. Furthermore, there exists a wide range of choice of articles and services available on auctions. Another factor that fosters this popularity is the existence of certain emotions connected with bidding an article, for example during last - minute bidding and thus outbidding other users.

Each year the percentage of internauts who buy online is becoming larger and larger. Therefore, the number of online auction users is rising even faster than the number of internauts in Poland. According to the data obtained from the NetTrack research, at the end of 2007, 45% of internauts were buying via the Internet [IAB]. The same rise of interest in the net shopping was observed through the results of the research administered through Gemius company [GEM] In 2006, the percentage of internauts who had ever bought online was 55%.

On Internet auctions, there is one seller and one or many buyers who meet. A transaction is finalised at the moment of ending the auction and one user's bidding at the highest price offered, provided that there was one article or service being sold.

The transaction was finalised. The most sensitive stage is the realisation of a finalised transaction. Usually, a buyer is obliged, after the end of an auction, to establish contact with a seller. Proposed forms of contact that confirm the willing for purchasing an article after the end of an auction are phone conversation or/and electronic mail. In this moment, what follows is the exchange of data concerning the form of a payment.

A popular method of a payment is through a bank transfer from a buyer to a seller. Usually, a seller demands that a bid amount due appears on their bank account before an article is sent. This requirement introduces the degree of risk that a buyer agrees to take. The degree of this risk is estimated by the rate of amount due for a bid article. The higher amount a buyer is ready to send to a seller, the greater risk the buyer takes.

The notion of risk on online auctions is the result of fraudulent seller's habits. Fraudulent sellers' habits take place in the situations as follows:

– sending of an article that does not match the description,
– not providing an article after receiving an amount due,
– providing of an article that does not match the description and not informing about important features such as damage,
– offering, through auction services, articles that originate from crime.

It should be emphasized that a buyer is burdened by risk for the reason that they send an amount due first, before an article is sent. A seller, though being the source of risk, does not take risk at all. A question thus arises, how to estimate credibility and reliability of sellers.

The adopted mechanism of both sellers and buyers estimation is by issuing comments after committed auctions. The essence of commenting system is to leave it entirely in the hands of users. It works as a basis for subjective and sovereign opinions on a whole transaction process. An auction service takes no responsibility for the content and quality of issued comments, for which users only are responsible. Interestingly enough, contracting parties are not obliged to issue any comments. Expressing an opinion about a given transaction is voluntary. After a committed transaction, there is usually a possibility of issuing one of the three types of a comment: positive, neutral or negative one, and to supplement it by a description, in which some details of a transaction may be contained. Unfortunately, this kind of estimating sellers' reliability is starting to be unreliable itself. It stems from the fact that, amongst others:

– some sellers who have the largest number of positive comments, which should undeniably prove their reliability, have the largest number of negative comments ,
– inexperienced buyers often do not know that they have a right to issue a comment,
– a buyer, who is not satisfied with all the details of a committed auction, is afraid of issuing a neutral or a negative comment so as not to receive the same type of a comment from a seller in revenge,
– a clique of fraudulent sellers realise fictitious transactions on auctions and they issue positive comments for one another,
– there is a common practice that a buyer should issue a comment first. It is a buyer's obligation to transfer an amount due to a seller as soon as possible and to give details to receive delivery. This is the end of a buyer's role, but it should be emphasised that it is not regulated anyhow, but it is a common practice, whereas sellers demand that buyers issue comments first. They justify it by stating that a

buyer's comment is a signal that an article has been delivered and they can issue a comment afterwards. Nevertheless, this kind of approach is unconvincing, as it makes the commenting system distorted. It enforces a buyer's greater satisfaction with the bid article and it acts as a form of protection of sellers who render low standard services against a negative or neutral comment, for they can always issue a similar one in revenge.

– the essential part of positive comments issued for sellers who realise the highest number of transactions is posted by buyers after a long period of time, after months or even years. So far, a through analysis of this phenomenon has not been conducted. This situation was observed through a thorough analysis of the dates of the posted comments and comparing it with a chronological number of a transaction.

Taking all these facts into account, there must be some improvement in the mechanism of estimation of online auction sellers. Most of auction services in the Internet use the mechanism of sellers estimation created by eBay company. In this system, credibility of a user is expressed by subtraction of the number of users who ever posted a positive comment and the number of users who ever posted a negative comment. This simple mechanism can be easily manipulated. Comments issued by users are subjective and incomparable. It lacks the context of a transaction. The quantitative approach that has been used so far reveals some serious stumbling blocks arising together with the rise of the number of users and online auction transactions.

2 The Overview of the Used Methods

In Poland the research on auction users' reputation has been conducted for a longer time by Morzy from Poznan University of Technology [MW1],[MW2]. In his papers he publishes the research on exploration of data on the basis of a sample of committed auctions. In his latest paper he presented a measurement of sellers' positive and negative reputation based on the modified PageRank algorithm [PBMW].

The other group conducting the research on sellers credibility are Piasecki et al. [PRZK]. In their work they forecast sellers' practices on the basis of the amplitude of frequency of offered articles on auctions and their prices. With the usage of the method of a decision tree and the C4.5 algorithm they are capable of pointing, with 85% possibility of accuracy, sellers who will act fraudulently during a transaction. Morzy is working on the set of the data obtained directly from the greatest auction service, Allegro, whereas Piasecki et al. obtained data through extracting the essential information directly from Internet auction sites of the same auction service.

3 The Methodology of Research

The methodology of the author is based on both Morzy and Wierzbicki and Piasecki and others' works. The author obtained the research data by constructing a webcrawler. The crawler was to scrutinise all the pages of the registered users of the Allegro service. The first versions of the crawler were prepared to extract and register the whole of html site code for the local volume. It turned out after time that this approach was impractical, for extracted data outdid harddrive capabilities of common Pcs and servers available in the scientific units. The webcrawler, which immediately registered necessary information from the extracted data into the data bases, was modified. The software was created in Perl with the usage of LWP and DBI libraries, that cooperated with the MySQL5 data base.

The most difficult part of creating the webcrawler was preparing a regular expression for extracting information about users. The mechanism was oriented on registering into the data base full information about a transaction, which is a seller's identification, a buyer's identification, a number of a committed auction and a type of an issued comment. During the works, it turned out that the Allegro service had been creating html sites that were incompatible with the web standards, whereas the data were presented semi structurally. The additional inconvenience was allocating whitespaces in the site code, which caused incorrect functioning of the data extraction mechanism.

The mechanism had to be also resistant to non - standard information appearing on sites with users. Invalid comments belong to one of it. This situation is vary rare, it usually happens with the most active users, who protect themselves from fraudulently issued comments this way. The other non - standard situation was the possibility, during the beginning of the service practice, of posting a comment several times on the same auction. The webcrawler would register into the base double information about a transaction at the same time. Usually, the first issued comment was negative, whereas, after explaining a situation, a user issued the new one. It resulted in the inaccuracy of the data, for various types of comments were posted on the same auction, and the primary key was created on the basis of a buyer's identification and a number of an auction. It stemmed from the author's assumption that a buyer can post only one comment on one auction. The situation was solved through creating the independent primary key, and then, during processing of the data, only one, latest comment on an auction was taken into account.

The full data collection lasted for 10 days, it cost 1 TB of the data transfer and it consisted of all the issued comments after committed auctions that had been conducted since the beginning of the service activity. In the base, the following information in Table 1 about the online auction transactions is contained.

Table 1. The Statistics of the Collected Data.

The number of transactions	102 153 732
Positive comments	100 912 797
Neutral comments	476 569
Negative comments	764 366
The number of sellers	1 575 199
The number of buyers	3 368 274

4 The Basic Definitions

The weakness of auction sellers' credibility measurement systems lies in the fact that the opinion of a given user may be strongly influenced by other users' behaviour, who can post false comments. This influence appears when frauds cooperate to manipulate the opinion of a given user. Usually, manipulation happens when false identities of auction users are used to issue false positive comments, which has never been posted on an auction.

Therefore, the opinion of buyers who buy from the seller *s* may be false in favour of the seller *s*. The effective solution is a question addressed to other sellers about what kind of an opinion on the seller *s* their clients express. This way, issuing the opinion about the seller *s* is done by recommendation of buyers who buy also from other sellers. Comments of buyers who are experienced in buying and posting comments are also taken into account.

Given a set of sellers $S=\{s_1,s_2,...,s_n\}$. The sellers s_i and s_j are linked if there exists at least *min_k* of buyers who realised transactions with both the seller s_i and the seller s_j. For the given sellers s_i and s_j a set of shared buyers is denoted by K_{ij}. The neighbourhood $N(s_i)$ of the seller s_i contains all the sellers s_j that are linked to the seller s_i for the assigned by a user minimal threshold *min_k*. The minimal threshold *min_k* requires an analysis of these sellers only whose number of conducted transactions is enough. A buyer who realised transactions both with the seller s_i and s_j, indirectly expresses an opinion about both the sellers.

Given $G=(S,E)$ the directed graph with a set of vertices S and a set of edges E. The set S is identical with the set of the sellers. In the set E, between the edges s_i and s_j there exists an edge only if the s_i and s_j are linked for the value of the minimal threshold *min_k* assigned for a user. Denoting $f\__{ij}$ as a comment issued by the buyer k_i from the seller s_i. It is assumed that a given buyer realises only one transaction with a given seller. If there are more transactions, latest posted comment is taken into account. Each comment can be positive, neutral or negative. For each type there was a numerical value assigned:

$$f_{ij}^{+} = \begin{cases} 0.2 & if \quad f_{ij} = "neutral" \\ 0.8 & if \quad f_{ij} = "positive" \end{cases} \tag{1}$$

$$f_{ij}^{-} = \begin{cases} 0,2 & if \quad f_{ij} = "neutral" \\ 0,8 & if \quad f_{ij} = "negative" \end{cases} \tag{2}$$

On the basis of the values shown above, there are weights assigned to each edge of the S - graph. Each edge e_{ij} between the junctions s_i and s_j has two weights that represent the amount of the positive and negative information that appears between the sellers. The positive weight is expressed by the formulation:

$$w_{ij}^{+} = \sum\nolimits_{k_n \in K_{ij}} f_{ij}^{+} \quad (3)$$

The positive weight represents the positive comments posted by the contracting parties of the s_i for the seller s_j. Naturally, the negative weight is expressed by the formulation:

$$w_{ij}^{-} = \sum\nolimits_{k_n \in K_{ij}} f_{ij}^{-} \quad (4)$$

The negative weight represents dissatisfaction of buyers from the given seller s_j, who were also dissatisfied with the transaction with the seller s_i.

Each edge e_{ij} in the S - graph represents the estimation of the services rendered by the seller s_j, issued by the buyers who were approved by the seller s_i. The positive weight w^{+}_{ij} represents trust declared by the seller s_i towards the seller s_j. The negative weight w^{-}_{ij} represents lack of trust expressed by the seller s_i towards the seller s_j.

The algorithm that uses these relations is modeled on the PageRank algorithm [PBMW]. Similarly to the PageRank algorithm, the proposed solution refers to random moving along the graph. A random user simply progresses along the edges of the graph. If a real web user got into a small loop between two sites, it is improbable that they would continue moving in it endlessly. Instead, they would jump onto another site. This behaviour is denoted as temporary boredom and jumping onto another edge. The vector F, that was introduced into the PageRank formulation, acts as counterbalance for the impact of unnatural rise of the ranking during the iteration in closed loops. During the research, it turned out that the best value to take for is that proposed by Page et al. [PBMW], which is: $d = 0.85$

$$F = (1 - d) \quad (5)$$

In the algorithm presented below (6), all the edges are directed and diverted. The F vector increases the value of the ranking. It is responsible for assigning a value to the sellers, to whom a random user will jump when in the loop. The value of the F vector illustrates what the probability of a random user jumping onto another random edge is. This is an equal approach for all the sellers, to whom the additional value is assigned, only because they exist and that they met the basic condition *min_k*.

$$P(s_i) = (1 - d) + d \sum_{s_j \in N(s_j)} \frac{P(s_j) w_{ji}^{+}}{\sum_{s_k \in N(s_j)} w_{jk}^{+}} \quad (6)$$

The presented algorithm (6) differs from Morzy's proposition in two aspects. Firstly, the vector F is not divided by the number of the sellers who take part in the calculations. Each seller is assigned to a value that at least equals the F vector already in the first iteration. Secondly, the ranking $P(s_j)$ of the seller s_j, whose part is transmitted to the ranking $P(s_i)$ of the seller s_i in every subsequent iteration, is weighed through the proportion of the weight w_{ij}, which the seller s_j turns over to the seller s_i, to the sum of all weights w_{jk}, which the seller s_j turns over to other sellers in the neighbourhood.

As a result of all the changes introduced to the algorithm, the sum of the rankings of all the sellers taken into account, equals their number, therefore the average is 1. The original hypothesis of the authors of the PageRank algorithm states that the average value of all the PageRank websites is 1. This leads to a conclusion that if each page has a link, the sum of all the PageRank values in the system equals the number of pages. After achieving convergence, the values of the rankings are very diversified, beginning with the values close to the vector F, to values thousands times as much. The results were modified to a readable form through a logarithmic function.

5 The Experiment

To conduct the research, the data previously obtained were used. The set was divided into two subsets, with one in which there were positive and neutral comments only, and the other with neutral and negative comments. In the research the missing comments were not taken into account. Morzy states that a missing comment is the one which "was purposely omitted, eg. in the result of a transaction, whose result is hard to be considered satisfactory and which does not deserve recommendation, but which does not qualify as a negative one either." A separate article was devoted by Morzy and Wierzbicki [MW2] to this matter, in which they studied 890 876 comments issued for 656 376 committed transactions, for 10 000 buyers. In the considered work, the missing comments were not included for two reasons: for it is impossible to extract them from the auction service website, because they do not exist, therefore it is not known how the transaction which was bid by a buyer ended, furthermore, the author knows, through empirical experiments, that not every missing comment qualifies for the definition stated above. Once a transaction is finalised, a comment may not be issued for the reasons such as: lack of time for posting it, buyer's forgetting about a committed auction, not issuing a comment purposely because of bad experience with other sellers (many sellers issue comments for buyers with a great delay) or lack of knowledge about the obligation to issue such a comment.

In the first part of the experiment, the author conducted studies on the sellers' negative reputation. It resulted from the smaller amount of calculations needed. On PC class Dell Latitude equipped with Centrino 1.73Ghz processor and 512MB RAM, calculation of the SS - graph and the negative reputation lasted 48 hours. The code of the program and the data base were optimised to the calculations through previously conducted series of tests on a smaller set of data. The code of the program was read by the interpreter of Perl language, thus, one of the possibility of optimisation was to prepare queries for the base before entering into calculative iterations, so as not to repeat them unnecessarily. Generally, the optimisation was to minimise the amount of necessary loops indispensable for the program that calculated the value of the ranking for one seller. The size of the S - graph resulted in 267 216 vertices and 2 065 932 edges. The whole convergence of the algorithm calculation was achieved in 80 iterations. A whole convergence is a moment, in which each subsequent iteration of the calculation of all rankings does not change its value anymore.

On the Figure 1, there is the illustration of the values of the ranking for all the sellers. The ranking met the criterion of *min_k = 1*, thus, there exists at least one buyer, who issued a negative or neutral comment for both the sellers s_i and s_j. The bar chart shows the results obtained after the transformation through a logarithmic function. On the chart, there is the total amount of all the 267 216 sellers, divided according to the obtained ranking. The rest of the sellers, which is 1 307 983, has the negative ranking equaled 0. The difference between the ranking equaled 4 and 5 is clear. To the value of 4, the ranking constantly rises, whereas after reaching the value of 5 it falls rapidly, and with the following values it quickly comes closer to the minimum. The conclusion arises that there exists the vast group of the sellers who often commit fraud during transactions being realised. It refers especially to the sellers with the ranking from 2 to 4. The Figure 2 presents the convergence. It seems that the convergence achieves the borderline values by twentieth iteration, nevertheless, the average value of the ranking reached only $P(s) = 0.986$. After the thirty - first iteration, the average value was nearing $P(s) = 0.999$. The highest ranking, before the transformation through the logarithmic function, for one seller was 389.06. The Table 2 illustrates the results of the first ten of the worst sellers.

Table 2. The Sellers' Negative Ranking, the Results after the Transformation through the logarithmic Function.

No	Name	Id	Negative ranking	Position/sale	Suspended
1	APTEL	1331662	389,060	2 - 128895	No
2	Marjot2	1193028	315,028	5 - 94366	No
3	Multi_kom	3816249	303,040	6 - 88146	No
4	elektroland	537684	230,825	7 - 84355	No
5	Telfor	1792668	212,198	38 - 41081	No
6	akcesoria_do_gsm	3397317	196,292	29 - 47104	Yes
7	_Mischell	1067984	164,376	45 - 39238	Yes
8	aporad1	4882231	162,077	853 - 9579	Yes
9	sklep.tsenger.pl	243168	148,475	104 - 29155	Yes
10	-MARTEL-	2376544	133,555	97 - 30121	Yes

For the calculations through *min_k=1* criterion, only 70% of all the sellers who ever received a negative or neutral comment were taken. The rest 30% of the sellers received a negative or neutral comment from a buyer, but the same buyer did not issue a similar comment for anyone anymore.

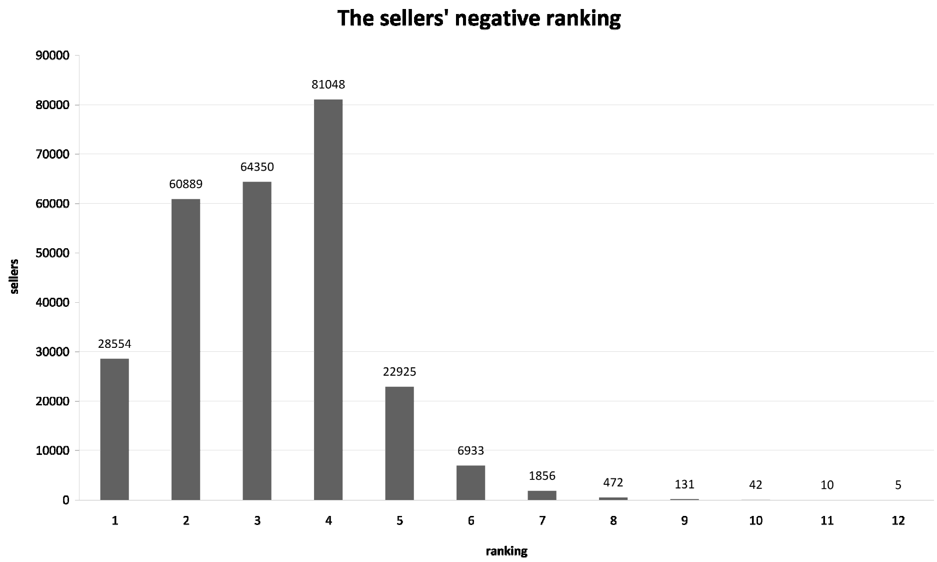

Fig. 1. The Layout of the Sellers' Negative Ranking.

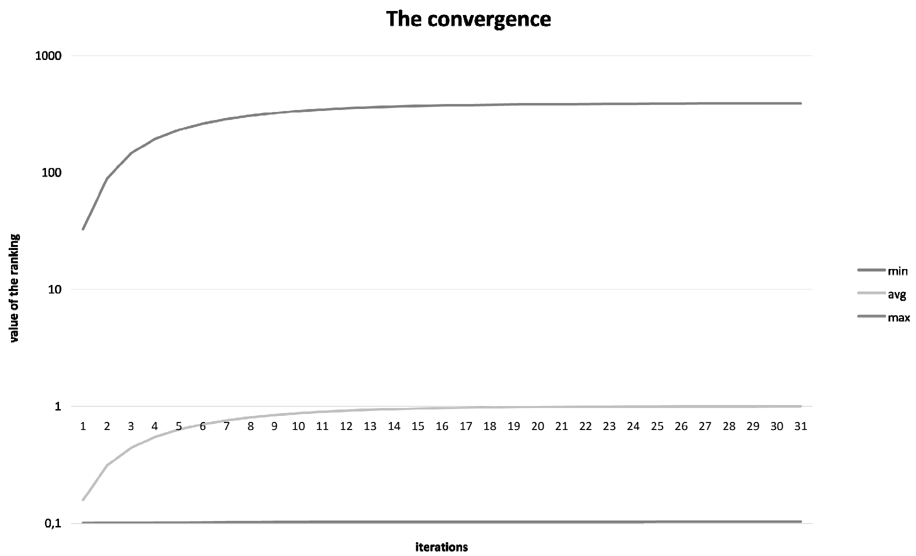

Fig. 2. The Convergence of the Sellers' Negative Rankings in the Subsequent Iterations.

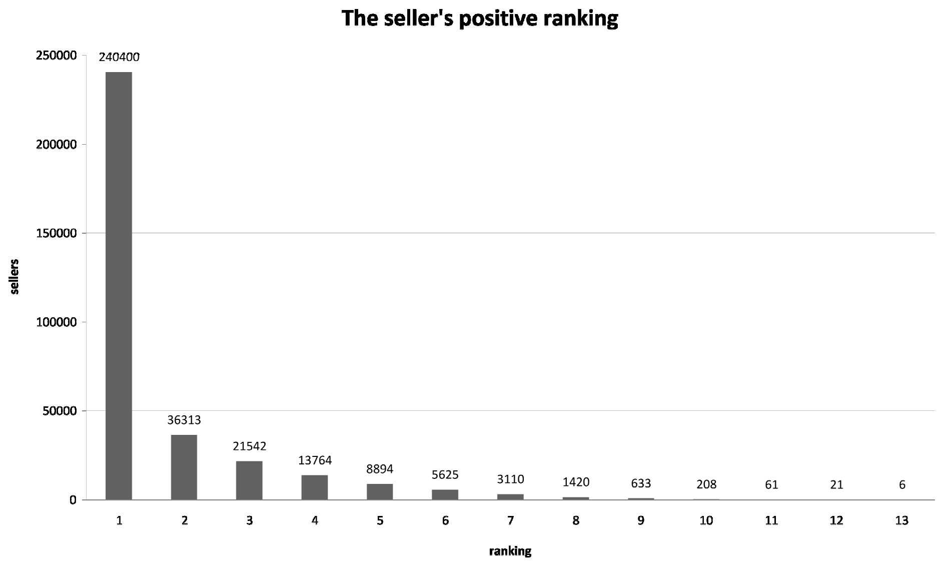

Fig. 3. The Layout of the Sellers' Positive Ranking.

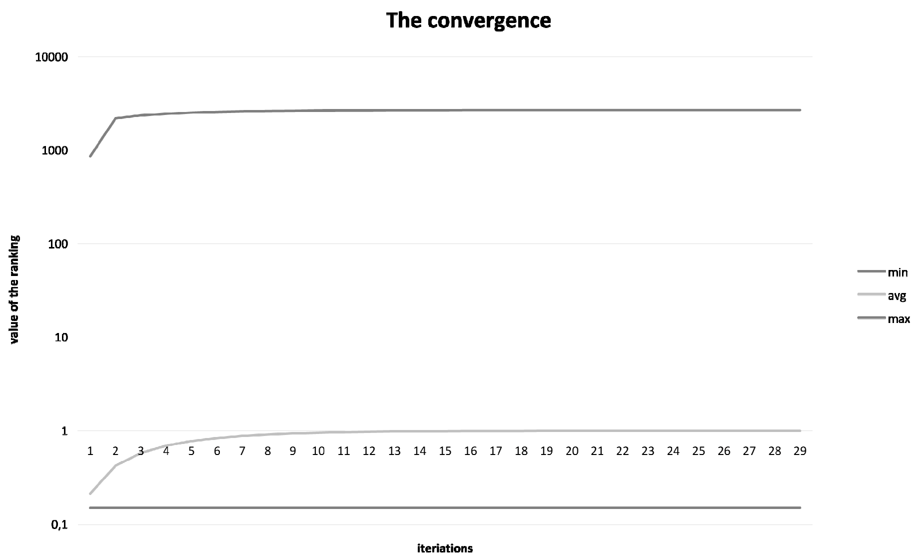

Fig. 4. The Convergence of the Sellers' Positive Rankings in the Subsequent Iterations

The conducted studies show that the algorithm which marks a seller's reputation is not correlated with the number of conducted auction transactions at all. It follows from the data presented in the column *position/sale* in the Table 2. The first ten of the negative sellers is not correlated with the volume of the sale at all. Therefore, there exists no menace that as a result of the sale rise, the sellers with the highest number of realised transactions will be naturally high in both the rankings. Thus, they should not be merged.

Preparing the *S* - graph for the sellers' positive reputation marking on the basis of the positive and neutral comments lasted 12 weeks. The calculations were conducted parallelily on three computers class IBM 445 server. The average speed of marking the graph was approximately 65 000 000 edges daily, whereas Dell Latitude 610 computer was able of calculating 6 000 000 ones a day. After reaching approximately 2 500 000 000 edges in the graph being constructed, the author took a decision to exclude the sellers whose strength of the link met the condition $w^{+}<4$. Thanks to it, the data set remained relatively small and the size of the database did not slower the calculations.

The created graph meeting the condition $w^{+}>=4$ for each seller had 331 998 vertices and 172 051 708 edges. Marking of the ranking for one iteration lasted 72 hours on average. All the calculations for 30 iterations lasted about 3 months.

Figure 3 presents the sellers' positive ranking. The sellers met the condition $min\_k=5$, thus, there exist at least 5 buyers who bought from both the seller s_i and s_j. Above 72% of the sellers meeting this condition were set at the first ranking value. The rest 28% of the sellers created a decreasing value distribution reaching 13. The convergence of reaching the following iteration results presented on Figure 4 is similar to the negative ranking. Figures 1 and 3, despite the same model of calculations, show diversified habits of online auction users.

6 The Conclusions

The conducted research shows the effectiveness of the proposed solution. The measurements do not represent the simple quantitative approach in calculating of the comments, they are based on the buyers' experience. The author hopes that the change of the calculative platform or enhancing calculative power will allow presenting a full measurement of positive and negative reputation of all Polish online auction sellers in the future.

At present, according to the Allegro auction service regulations, users have the possibility to express an opinion about the process of a transaction. An article to which an auction refers is not taken into consideration. It is worth considering a solution to separate the estimation of article quality and service quality. A user faces a dilemma when a transaction itself was conducted properly, but an article is damaged. Having the possibility to express an opinion on both a transaction process and an article, the commenting system would become fuller, more credible.

During analising the data empirically, the author discovered a very disturbing situation. Alongside the suspended accounts of the users from the first ten of the worst sellers, positive comments constantly appear. The chronological number of an auction reveals that these are the auctions that had been conducted from several months to years before suspending a seller's account. In all likelihood, the essential part of comments is issued by users who have false identities, using them to manipulate sellers' opinion. Similarly, alongside the accounts of the users who are not suspended by the service, but who have a vast number of realised transactions, the considerable part of the comments is issued for the auctions that had been conducted in the period of months or years before. The conducted additional research reveals a large number of transaction comments issued with at least 6 month delay with respect to the actual transaction end. The number of such comments is 1 748 061. It is above 1.5% of the number of all the transactions. It is disturbing that this behaviour refers to the sellers with the greatest volume of sales.

References

1. Gemius S.A. Raport Internet 2006,
2. Morzy M., Wierzbicki A.: Eksploracja aukcji internetowych w poszukiwaniu pozytywnej i negatywnej reputacji sprzedawców, W: Technologie Przetwarzania Danych: II krajowa konferencja naukowa, pod red. Tadeusza Morzego, Marcina Gorawskiego, Roberta Wrembla, Poznań, Wydawnictwo Politechniki Poznańskiej, 2007.
3. Morzy M., Wierzbicki A.: The Sound of Silence: Mining Implicit Feedbacks to Compute reputation, P.Spirakis et al. (Eds.): WINE 2006, LNCS 4286, 2006.
4. Page L., Brin S., Motwani R., Winograd T.: The PageRank Citation Ranking: Bringing Order to the Web, Technical Report, Stanford University 1998.
5. Piasecki J., Roczniak W., Zygmunt A., Koźlak J.: Analiza wzorców nielegalnych zachowań w aukcjach internetowych, W: Technologie Przetwarzania Danych: II krajowa konferencja naukowa, pod red. Tadeusza Morzego, Marcina Gorawskiego, Roberta Wrembla, Poznań, Wydawnictwo Politechniki Poznańskiej, 2007.
6. Raport strategiczny IAB Polska, Internet 2006, Polska, Europa i Świat