DOI: 10.15514/ISPRAS-2019-31(3)-9

Reputation Systems in E-commerce: Comparative Analysis and Perspectives to Model Uncertainty Inherent in Them

M. M. Nosovskiy, ORCID: 0000-0003-4475-3787 <mmnosovskiy@edu.hse.ru>
K. Y. Degtiarev, ORCID: 0000-0001-5519-1033 <kdegtiarev@hse.ru>
National Research University Higher School of Economics (HSE),
Faculty of Computer Science, School of Software Engineering,
3, Kochnovsky Proezd, Moscow, 125319, Russian Federation

Abstract. E-commerce is a runaway activity growing at an unprecedented rate all over the world and drawing millions of people from different spots on the globe. At the same time, e-commerce affords ground for malicious behavior that becomes a subject of principal concern. One way to minimize this threat is to use reputation systems for trust management across users of the network. Most of existing reputation systems are feedback-based, and they work with feedback expressed in the form of numbers (i.e. from 0 to 5 as per integer scale). In general, notions of trust and reputation exemplify uncertain (imprecise) pieces of information (data) that are typical for the field of e-commerce. We suggest using fuzzy logic approach to take into account the inherent vagueness of user's feedback expressing the degree of satisfaction after completion of a regular transaction. Brief comparative analysis of well-known reputation systems, such as EigenTrust, HonestPeer, Absolute Trust, PowerTrust and PeerTrust systems is presented. Based on marked out criteria like convergence speed, robustness, the presence of hyper parameters, the most robust and scalable algorithm is chosen on the basis of carried out sets of computer experiments. The examples of chosen algorithm's (PeerTrust) fuzzy versions (both Type-1 and Interval Type-2 cases) are implemented and analysed.

Keywords: e-commerce; reputation system; peer-to-peer computing; trust management; uncertainty; fuzzy logic; linguistic variable; type-1 fuzzy set; type-2 fuzzy set

For citation: Nosovskiy M.M., Degtiarev K.Y. Reputation Systems in E-commerce: Comparative Analysis and Perspectives to Model Uncertainty Inherent in Them. Trudy ISP RAN/Proc. ISP RAS, vol. 31, issue 3, 2019. pp. 99-122. DOI: 10.15514/ISPRAS-2019-31(3)-9

Репутационные системы в электронной коммерции: Сравнительный анализ и перспективы моделирования присущей им нечеткости

М.М. Носовский, ORCID: 0000-0003-4475-3787 <mmnosovskiy@edu.hse.ru>
К.Ю. Дегтярев, ORCID: 0000-0001-5519-1033 <kdegtiarev@hse.ru>
Национальный исследовательский университет "Высшая школа экономики",
факультет компьютерных наук, департамент программной инженерии,
125319, Россия, г. Москва, Кочновский пр-д, д. 3

Аннотация. В наши дни электронная коммерция (ЭК) показывает беспрецедентные темпы роста во всем мире, вовлекая в эту деятельность миллионы людей на всех континентах. В то же время, ЭК создает почву для злонамеренных действий, что требует особого внимания и контроля. Одним из способов минимизации таких угроз является использование репутационных систем для отслеживания степени доверия в среде пользователей сети. Большинство существующих репутационных систем основаны на сборе отзывов относительно проведенных транзакций, и они, как правило, работают с представленными в виде чисел откликами клиентов (в частности, может использоваться привычная целочисленная шкала 0..5). В целом, понятия доверия и репутации являются примерами неопределенных (неточных) информационных данных, характерных для сферы электронной коммерции. Мы предлагаем использовать аппарат нечеткой логики для формального представления пользовательских отзывов, выражающих степень удовлетворенности результатом совершенных транзакций. В работе представлен краткий сравнительный анализ наиболее известных репутационных систем, таких как EigenTrust, HonestPeer, Absolute Trust, PowerTrust и PeerTrust. С учетом выделенных в результате анализа критериев (скорость сходимости, устойчивость (робастность), наличие гиперпараметров), проведенная серия компьютерных экспериментов позволила эмпирически выделить PeerTrust как наиболее устойчивый и масштабируемый алгоритм из числа рассмотренных. При наличии ограничений в отношении имеющихся данных, подготовлены реализации (Python 3.7) и проанализированы результаты, связанные с особенностями поведения нечетких версий алгоритма PeerTrust на основе нечетких множеств типа-1 (T1FS) и интервальных нечетких множеств второго типа (IT2FS).

Ключевые слова: электронная коммерция; репутационная система; пиринговые вычисления; управление доверием; нечеткость; нечеткая логика; лингвистическая переменная; нечеткое множество 1-го типа; нечеткое множество 2-го типа

Для цитирования: Носовский М.М., Дегтярев К.Ю. Репутационные системы в электронной коммерции: Сравнительный анализ и перспективы моделирования присущей им нечеткости. Труды ИСП РАН, том 31, вып. 3, 2019 г., стр. 99-122 (на английском языке). DOI: 10.15514/ISPRAS-2019-31(3)-9

1. Introduction

E-commerce is a buying-selling runaway activity widening at an unprecedented rate all over the world and inveigling into fascination of various e-stores people of all ages. Ever-growing number of various websites and apps focusing on e-commerce domain makes it simple and alluring to find and to buy immediately almost anything whatever client's heart desires [1].

There is no doubt that e-commerce sales opportunities are rapidly progressing day by day. Owing to Internet, businesses bring their products and services to customers literally in eyewink. The e-commerce share of total retail sales in the United States amounted to 10% in 2018, in expectation of attainment of 12.4% by 2020 with further strengthening its ground [2]. With such perspectives in mind it is easy to realize why e-commerce entrepreneur position becomes so attractive. With an estimated 95% of purchases that will be made online by 2040 and expected growth of year to year sales standing at the level of 15%, the opportunity to find a niche for selling products online has massive indisputable potential [3]. During the last 5 years the amount of retail sales raised from $1.3 billion to $2.8 billion. The latter is expected to nearly double (up to $4.8 billion) by the end of 2021 [4].

One of the most growing types of e-commerce is *online marketplace* that can be defined as a website or app that facilitates shopping from many different sources [5]. Among well-known and successful examples of online marketplaces eBay, Amazon, Rakuten (worldwide) and Avito, Ozon (in Russia) can be mentioned. Online marketplace acts as a platform integrating buyers and sellers. Being a peer-to-peer (P2P) network, it allows buyers to purchase any goods or services offered by sellers through this online platform. Usually, peers (people or businesses) communicating through online marketplace remain in the status 'strangers' with respect to each other. They don't have at their disposal reliable information about alter peer, whether it is a buyer or a seller. Therefore, peers must manage the risk associated with transactions on condition that no prior experience and knowledge concerning mutual reputation of sides exists [6]. This problem can be addressed by means of developing a system on top of the network that should help peers to evaluate their past experience with other peers and to manage trust between them as well as reputation of each peer involved. This kind of systems is called *reputation systems*.

Various implementations of reputation systems exist starting with very simple to more complex ones designed mostly for P2P file-sharing networks [7-10]. Such systems have a positive impact on peer's experience as they help to distinguish trustworthy peers from ill-intentioned and unreliable opponents. For example, in reputation system used by eBay, one of e-commerce leaders, buyers and sellers have a chance to rate each other with numeric scores +1, 0 or -1 after each carried out transaction. The overall reputation of a participant is calculated as a sum of scores earned over last six months [8]. At that all such systems rest upon notions of *trust* and *reputation*. Trust (or, local trust) represents personal experience (attitude) of a user regarding another user, while reputation constitutes an aggregate of these individual trust values on the scale of the whole community. Calculation of local trust and corresponding aggregates underlies implementations of all known reputation systems.

Despite the practical effectiveness of these systems, there is a substantial drawback inherent in them, viz. none of them can handle uncertainty "hidden" in online marketplace's data. The latter means data that relate to all transactions accomplished on the marketplace along with data collected from users after each transaction and metadata concerned with every user in the marketplace.

The primary concern of the paper is to provide the overview of best known reputation systems and to undertake their general comparative analysis on the basis of several key factors (criteria) – they are speed of convergence, complexity of calculations, use of hyper parameters expressing user's preferences, robustness and general system's suitability to handle imprecision and uncertainty of data. In the first place these factors are chosen to convey the requirements of key stakeholders who are owners and developers of a marketplace as well as its users. For the first group of stakeholders general system's effectiveness becomes important, and it is attributed above all to the efficiency of its functioning, computational resources needed to perform the work and ability for customization. Users are mostly interested in reliability of system's output and how well it suits each given user. The last factor mentioned above reflects how naturally specific implementation of the system can be extended to handle data uncertainty and imprecision, since the latter being an inherent part of virtually any system reveals itself in different forms. The recognition of such manifestation forms of uncertainty becomes a task of prime importance to represent appropriately (model) its distinctiveness. Consequently, fuzzy logic is getting one of pivotal theories that captures naturally the phenomenon of imprecision and uncertainty [11].

The rest of the paper is organized as follows: in section 2 notions of trust and reputation, difference between them, are considered. Uncertainty in the marketplace and verbal assessments that are inherent in reputation systems form the contents of section 3. Some basic terms and definitions relating to the field of fuzzy sets and logic are covered in the section 4. Section 5 is devoted to the brief comparison of five well-known reputation systems (EigenTrust, Absolute Trust, PeerTrust, et al.) and stressing their key differences as well as intrinsic similarities. Setup of computer-based experimental part of the work (parameters and their values used) constitutes the material of section 6, whereas the results of carried out experiments are discussed in the section 7. Thereafter, the transition from crisp to type-1 and interval type-2 fuzzy PeerTrust algorithm (analysis of such transition's outcome) is presented in finishing sections 8 and 9. Concluding remarks and observations are drawn in section 10.

2. Trust and reputation. What is the difference between these terms?

Trust and reputation are the main concepts underlying vast majority of reputation systems. In order to clearly recognize the purpose of reputation systems, we need to define what do trust and reputation in terms of online marketplace stand for. Diverse sources give different definitions of the term 'trust'. The basic definition presented in Oxford English Dictionary reads as follows: «Trust is a firm belief in the reliability, truth, or ability of someone or something» [12]. However, such definition cannot lay claim to completeness, since notions of trust and reputation as applied to peculiarities of Internet-based activities must be defined in a more context-specific way. Among other things, Alam & Paul

define trust as «a belief, the trusting agent has in the trusted agent's willingness and capability to deliver the services that they are mutually agreed on in a given context and in a given time slot» [13]. In addition, Wang & Vassileva associate term 'trust' with «a peer's belief in another peer's capabilities, honesty and reliability based on its own direct experiences» [14]. Starting from individual judgments and predictions, Gambetta state that «… trust is a subjective probability that relies on context and reputation, it describes how secure a situation is even though risk is associated with it» [15]. It can be noticed that trust is mainly linked to belief that peers (agents) mentally possess in malicious P2P environment. Thus, trust can be viewed as a soft system's factor that is difficult to express precisely and in complete form. It is tied to distinction of numerous generally inhomogeneous interactions between peers, organization of the network, in which humans play a pivotal role.

For reputation term situation seems resembling, i.e. there is also no conventional definition that most of sources agree on. According to [14], reputation is defined as «peer's belief in another peer's capabilities, honesty and reliability based on recommendations received from other peers». On the other hand, already cited above Alam & Paul propose to consider reputation as «aggregation of all recommendations provided by the third-party recommendation agents about the quality of the trusted agent» [13]. Abdul-Rahman & Hailes define reputation as «an expectation about an agent's behavior based on information about its past behavior» [16]. Kreps & Wilson link reputation to characteristic or attribute «ascribed to one person by another person (or community)» [17]. A complete (at least, voluminous) overview of definitions relating to trust and reputation can be found in [18]. In the present work, we use definitions for terms 'trust' and 'reputation' from [14] since both definitions agree with basic concepts of reputation system and interaction within P2P community.

Even though trust and reputation are very closely related concepts, and many sources simply use them virtually as synonyms, still there is a major difference to emphasize. While trust is subjective in nature, and it expresses local attitude of a peer regarding another agent on basis of his/her own past experience, reputation serves as a global and public perception of a given peer in the midst of other peers. With this point in mind, we may list those important characteristics of trust and reputation that must be taken into consideration when considering reputation systems.

- Context awareness (sensitivity) – trust or reputation of a peer is dependent on what the context of communication is. For instance, a peer can be really trustworthy in delivering books or office supplies, but unreliable in selling electronic accessories,
- Multi-faceted nature (diversity) – even in the same context, peer can evaluate the quality of communication with another peer on the strength of several aspects. In the case of online marketplaces delivery time, quality and price of goods (services) can be mentioned. While the context-sensitivity of trust underlines the fact that the trust in the same agent may vary with reference to different situations, the multi-faceted nature characteristic stands for manifoldness of trust. It definitely plays a substantive role in deciding whether an agent is trustworthy to interact with or not [14],
- Dynamism – apparently, levels of both trust and reputation increase or decrease in view of gaining experience (direct interaction). Such changes may alternate in due course depending on arising situation in the system, with a clear-cut declining tendency observed with time [14],
- Imprecision and uncertainty – it is not very habitual for humans to operate with estimates of trust and reputation in the form of numbers. Definitely, it is not difficult to perform relatively simple calculations even in passing, but explanations and interpretations are usually based on verbal forms (words, phrases and short sentences in natural language). The peer can be classified as *«very trustworthy»*, *«not too trustworthy»* or in some likewise manner. Thus, we express gradations (imprecise estimates) of the extent, to which the peer is reputed as trustworthy or not. The bounds of gradations (verbal granules) are inexact, but nevertheless linguistic forms are easily perceived and processed by specialists and ordinary people in talks, reasoning and decision-making process. We may conclude that trust is a highly subjective

category, and being apparently fuzzy it can be associated with verbal assessment values (granules). The vagueness of the trust is linked outright to uncertainty of reputation as well.

3. Uncertainty in marketplace data. Verbal assessments are very natural in reputation systems

The paper is focused on reputation systems as applied to e-commerce field (and specifically online marketplaces). Because of that it is essential to consider what kind of data concerning peers and their transactions are available, and what sort of data peer's feedback about fulfilled transaction contains. In online marketplaces there are two types of peers – they are sellers and buyers; every transaction implies participation of one seller and one buyer. It is important to distinguish these types of 'players', because they gain trust and reputation that differ by their gist. In the present work we consider three types of marketplace data:

- peer data, i.e. a set of general data pieces that relates to peer itself (personal data, registration date, etc.);
- transaction data – general data about transaction held between seller and buyer (delivery time, payment time, total sum and date of transaction, etc.);
- feedback data refer to data collected from both seller and buyer after completion of each transaction (goods quality, communication quality, shipping service reliability, etc.).

Most of the existing reputation systems work only with peer's feedback [7-10,19]. In general, feedback provides some subjective assessment of experience that a peer has with another peer in the course of transaction's realization. But in certain cases, such experience cannot be thoroughly expressed in terms of integers -1, 0 and +1 as it occurs in eBay system. Why do we think so? Firstly, regarding feedback as a number means neglecting diverse aspects of experience such as those mentioned above. Secondly, and this fact was also underlined earlier, it is more natural for humans to think of evaluating experience in terms of some ordinal scale stretching from «very bad» label to «very good» instead of using «good», «neutral» or «bad» plain marks as linguistic equivalents of -1, 0 and +1. In case of extended scale' use, its grades may overlap with each other, since each label or grade stands not for a single value, but for a range of values instead. For instance, there is no clear-cut border line between values (labels) «very bad» and «bad», but almost all people may differentiate these values mentally while impacting information chunks to others.

Similar situation comes to pass with reference to transaction data. For example, let us consider delivery time of basic electronic accessories from Moscow to Saint-Petersburg. We know that they are normally delivered within N days. Is it «quickly», «slowly» or «neither slowly, nor quickly» for a client? Maybe it is a little bit slowly, but not too much? What can be said about N+2, 2N days or even 5N days? At some point it becomes obvious that delivery time can be associated with label «slowly» or even «very slowly». But what do we mean by that some point? For different people it occurs at different moments, which are not fixed (crisp), and this is when imprecision and uncertainty (fuzziness) of these data reveal themselves.

4. Fuzzy logic theory. Some basic terms and definitions used in the study

Taking into account uncertainty inherent to notions of trust and marketplace data, we need to consider its formal representation for a possible use in trust management and reputation systems (models). The concept of uncertainty is many-sided and rich; furthermore, uncertainty 'accompanies' any interactions of humans with real world [11]. In this connection, reputation systems exemplify active communication of peers based on the exchange of information that is often a matter of human perceptions and interpretations to a various extent. Much depends here on cognition and verbal assessments expressed in the natural language. Such perceived units can be seen as granules with 'soft' bounds rather than exact quantities having unified meaning and interpretation by all parties involved into the process. The theory of fuzzy logic (FL) extends the

ontology of mathematical research in the context of formation of a composite methodology that leverages quality and quantity [20]. It provides ample means to model the "perceived meaning of words/phrases conveying the expert opinions (estimates) in a graded fashion" [21]. The following is a quick glance at main concepts and definitions concerned with fuzzy logic as used in the present study.

Definition 1. *Linguistic variables* (LV) – variables whose values are words, phrases or sentences (linguistic terms) expressed in natural or artificial language [22]. In short, we can state that LV constitute a form of information granulation serving as a base for further transition of those granules to computable counterparts [23]. For example, if we consider the case of delivery time from point A to point B, the label (term) "quickly" is one of possible linguistic values assigned to the variable *Delivery Time*. Its whole term-set can be represented as $\mathbf{T}_{(Delivery\ Time)}$ = «quickly» + «very quickly» + «slowly» + «more or less quickly» + … , where sign '+' denotes the aggregate of linguistic granules rather than arithmetic sum operation.

Linguistic variable *Delivery Time* is defined on the universal set U (realistic range of numeric values representing the delivery time in particular situation), i.e. each element $x \in U$ stands for the time (minutes, hours, etc.) that can be associated as a result of human's perception (judgment) with corresponding terms to various degrees.

Definition 2. Let U be a set of elements (objects) that are denoted generically as x (U={x}); *fuzzy set* $A \subseteq U$ is a set of ordered pairs $\{(x, \mu_A(x))\}$, where mapping $\mu_A : x \to [0,1]$ is a (type-1) membership function of a fuzzy set A. Value $\mu_A(x)$ is a degree (grade) of membership of x in the set A, and U is a problem's domain (universal set). Membership function (fuzzy set) represents possibility distribution of x-values over domain U, and it can be expressed as aggregation $\int_{x \in U} \frac{\mu_A(x)}{x}$ or union $\sum_{x \in U} \frac{\mu_A(x)}{x}$ of pairs $\{(x, \mu(x))\}$, $\mu(x) \in [0,1]$, in continuous and discrete cases, correspondingly.

Definition 3. Let A be a fuzzy set on U, then α-*cut* of A is a crisp (non-fuzzy) set A_α composed of all $x \in U$, whose grades of membership in A are greater or equal to α [22]. Formally, A_α can be expressed as $\{x \mid \mu_A(x) \geq \alpha\}$. A fuzzy set A may be decomposed into and restored from α-cut sets through the resolution identity [24, 25], i.e. $A = \int_0^1 \alpha A_\alpha$, or $A = \sum_\alpha \alpha A_\alpha$, $\alpha \in [0,1]$.

An integral part of any formal modeling approach is closely related to the use of functions. Along with pervasive processing of non-vague objects, fuzzy quantities in last three decades became widespread in algorithms covering enormous circle of application domains. The need to extend the possibility for functions to operate with arguments having the form of fuzzy sets has led to formulation of extension principle [22, 26, 27]. As its name speaks for itself, it is directed to spreading nonfuzzy mathematical concepts to fuzzy ones [28]. It is specifically what is required to handle aspects of uncertainty (fuzziness) with reference to existing reputation systems.

Definition 4. Assume f is a mapping from universal set U to set V, and A is a fuzzy set defined on U (for the sake of simplicity we may consider finite representation of such set, i.e. $A = \mu_A(x_1)/x_1 + \mu_A(x_2)/x_2 + ... + \mu_A(x_n)/x_n$). Relying on the *extension principle*, the image $f(A)$ of A under mapping f is obtained as follows:

$$f(A) = f\left(\mu_A(x_1)/x_1 + \mu_A(x_2)/x_2 + ... + \mu_A(x_n)/x_n\right) =$$
$$= \mu_A(x_1)/f(x_1) + \mu_A(x_2)/f(x_2) + ... + \mu_A(x_n)/f(x_n).$$

In other words, the image of A under f can be deduced from the knowledge of the images $f(x_1), f(x_2), ..., f(x_n)$.

Definition 5. The process of representing initial data (e.g. linguistic values) as membership functions is called *fuzzification*; most of applications require to perform at final stages the opposite translation from fuzzy functional forms to crisp values – this is achieved through *defuzzification procedures* [11, 21, 26, 29].

5. Brief comparison of existing reputation systems – their differences and intrinsic similarities

The number of publications devoted to trust management and reputation systems is pretty imposing, and it is growing from year to year [6, 8, 9, 10, 14, 19, 30]. In the paper, we wittingly touch upon (just brief overview augmented with performance considerations) the most significant systems that proved themselves as effective, robust and applicable to online marketplace reputation management. It is worth mentioning that only those systems that do not use basically fuzzy logic concepts are reviewed in the paper. For instance, systems that utilize fuzzy inference schemes or other fuzzy-logic related notions [20, 31, 32] constitute an interesting research topic, but on level with other relevant cases it is outside of the scope of the present paper.

Results of the conducted analysis of existing sources provided a basis for selection of those criteria that can be classified as crucial from the viewpoint of systems' comparison. They can be described concisely as follows:

- Speed of convergence – iterative algorithms form a core of nearly all reputation systems. Thus, one of important aspects of such algorithms is how fast they converge and produce a result. This feature is covered as a principal one in most of papers related to reputation systems [6,8,9,10],
- Robustness – it is the criterion concerned outright with the main purpose of every reputation system, namely, the prevention of malicious attacks. Therefore, it becomes essential to measure how well a given system is able to held out against malicious peers' activities. Such experiments are covered by S.D. Kamvar, M.T. Schlosser, H. Kurdi, N. Chiluka, N. Andrade, Y. Wang, L. Xiong, et al. in [6,8,9,10,14,19]; however, it is important to mention that papers referenced here cover different types of malicious behavior,
- Hyperparameters – their presence is an important point to consider in the process of system's deployment, since they show the extent, to which the system is customizable. But at the same time, factor of their presence is a 'double-edged sword', inasmuch as, on the one hand, tuning of hyperparameters may lead to better performance of a specific system. On the other hand, it enhances significantly the complexity of deploying the system,
- Handling data imprecision and uncertainty – most of hypothetical or artificial systems do their work in the presence of uncertainty. The latter is often linked to human factor being an integral part of a system in the context of verbally defined and/or interpreted data. The latter are elicited from active discussions with stakeholders and estimations commonly used as inputs in calculations provided for algorithms underlying system's work specifics. Those pieces of information are often 'soft' (imprecise) in their nature, and it opens manifest way for the use of fuzzy logic theory in models of reputation systems. Hence, the comparison of algorithms can be performed in the view of how well system (algorithm) adapts fuzzy logic extension.

5.1 EigenTrust Algorithm

EigenTrust system is originally proposed for a P2P file-sharing network by S.D. Kamvar, M. Schlosser and H. Garcia-Molina in the paper that became one of the most cited papers on reputation systems [8]. EigenTrust calculates a global trust value for each peer based on his/her past behavior by incorporating opinions of all peers in the system [19]. Opinions concerning a particular peer are

represented as a local trust value. After each communication peer assesses his/her experience by the value from restricted set comprised by integers -1, 0 or 1. The local trust value is an aggregation of all communication experience assessments. It was shown how to normalize local trust values in a way that leads to elegant probabilistic interpretation similar to the Random Surfer model and efficient algorithm to aggregate these values [8,33]. Pre-trusted peers that can be seen here as a hyperparameter (it must be chosen in advance for the whole system to operate) are used to guarantee convergence and breaking up malicious collectives. The choice of pre-trusted peers is important, and it can compromise the quality of system to a marked degree [8]. As also shown in [8], for a network with 1,000 peers the algorithm converges after completion of less than 10 iterations. Theoretical base for fast convergence of EigenTrust algorithm is discussed by T.H. Haveliwala and S.D. Kamvar in [34]. Robustness of the system is evaluated under several threat models, and the system shows good overall performance in all cases [8]. For both Individual malicious peers and Malicious collectives threat models, EigenTrust system outperforms non-trust-based systems showing five to eight times better results with fraction of inauthentic downloads (FID) less than 0.2 for every setting. For Malicious collectives with camouflage case (model 3), system shows less impressive results, but still Malicious Spies threat model (fourth model) tends to be the best strategy for malicious peers to attack trust-based network [8].

As already mentioned earlier, EigenTrust system uses aggregated local trust values that must be normalized beforehand to avoid system's 'demolition' due to assignment of very high and very low local trust values [8]. Normalized local trust value c_{ij} can be calculated as $c_{ij} = \max(s_{ij}, 0)/\sum_j \max(s_{ij}, 0)$, where s_{ij} is a local trust value. The shown way of normalization isn't free of drawbacks. For one thing, normalized values don't draw a distinction between a peer with whom a given peer *i* did not interact and a peer with respect to whom peer *i* has had a poor (negative) experience. Secondly, c_{ij} values are relative, and they cannot be interpreted easily in the absolute sense [8]. Thus, an attempt can be made to extend EigenTrust algorithm with fuzzy logic notions to obtain transparent and interpretable modification of the original computational scheme. Particularly, calculation of local trust value may be altered to accumulate different types of marketplace data, but further study that concerns the impact of fuzzification on probabilistic interpretation of EigenTrust algorithm is required.

5.2 HonestPeer Algorithm

HonestPeer as an enhanced version of EigenTrust algorithm is discussed by H. Kurdi [9]. The algorithm endeavors to address one of the major problems with EigenTrust system, viz. pre-trusted peers. HonestPeer minimizes the dependency on that pre-trusted set of peers by choosing one honest peer dynamically for every computation step of global trust value (GTV). This honest peer, i.e. the peer having the highest global trust value, plays a crucial role in further computations of GTV. The speed of HonestPeer's convergence is almost the same as for EigenTrust algorithm, despite the need to perform additional calculations. Following [9], two benchmarks are considered – they are EigenTrust algorithm and no algorithm. Performance of HonestPeer algorithm is estimated under different experimental settings embracing variable number of users and files as well as number of pre-trusted peers and with the examination of percentage of inauthentic file downloads by good peers and success rate of good peers (success rate of good peers equals the ratio of *#valid files received by good peers* to *#transactions attempted by good peers*).

HonestPeer algorithm surpasses EigenTrust in effectiveness and capability to 'help' good peers to download valid safe files. This fact can be attributed to the ability of HonestPeer to choose honest peers dynamically after each round, while in case of EigenTrust pre-trusted peers are chosen statically irrelative of their performance [9]. Since HonestPeer is basically an enhancement of EigenTrust algorithm, the use of fuzzy logic may be appropriate and explicable as a practical matter to address those forms of uncertainty that are typical for system under consideration.

5.3 PowerTrust Algorithm

The reputation system PowerTrust, which is based on power-law distribution of peer feedbacks discovered after examination of 10,000+ eBay users' transaction traces, is covered by R. Zhou and K. Hwang [10, 35]. In PowerTrust system a few power nodes are selected dynamically according to their reputation. These nodes can be dynamically replaced, if they become less active or demonstrate unacceptable behavior. Good reputation of power nodes is accumulated from the operation history of the system – functional modules of PowerTrust system as well as flow scheme that relates to collection of local trust scores and global reputation aggregation are visually demonstrated in [10]. Without going into particulars, it should be mentioned that raw data input for PowerTrust is treated as local trust scores, which are then aggregated to obtain global reputation score of each peer. The Regular Random Walk module supports the initial reputation aggregation, while Look-ahead Random Walk (LRW) module is used to update the reputation score periodically. LRW also works with Distributed Ranking Module to identify power nodes. The system leverages power nodes to update Global Reputation Scores (vector V) [10].

The experimental performance of PowerTrust in terms of reputation convergence overhead to measure aggregation speed, ranking discrepancy to measure the accuracy, and root-mean-square (RMS) aggregation error to quantify system's robustness to malicious peers shows that PowerTrust algorithm outperforms EigenTrust by more than factor 1.5 in case of convergence speed [10]. Under all settings PowerTrust exhibits its robustness against collusive peer groups of various sizes.

In much the same way as for both EigenTrust and HonestPeer, the point of fuzzy logic's application to PowerTrust algorithm is a local trust value. Several linguistic terms may be defined on [0,1] interval to be used consequently for computation of global reputation. It is of definite interest to research whether the use of fuzzy logic may affect properties of PowerTrust algorithm or not.

5.4 Absolute Trust Algorithm

The algorithm for aggregation of trust among peers in P2P networks (Absolute Trust algorithm) was presented by S.K. Awasthi and Y.N. Singh [30]. Most of reputation systems are built upon scenario when all peers evaluate other peers by way of assigning foregoing local trust values that are a subject for further aggregation aimed at obtaining peers global reputation scores. In general, three different types of evaluation scenarios (*one-to-many*, i.e. one person is evaluating many persons, *many-to-one* scheme, under which many persons are evaluating one person, and *one-to-one* case, which implies that one person is evaluating another person) can be identified. In an effort to strengthen feedback's reliability in many-to-one evaluation scheme, any evaluation provided must allow for the competence of evaluator (evaluating party in the system) in computations via proportional weight's factor. Global trust of *j*-th peer can be used by way of weight in aggregation of local trust scores in calculation of any given *i*-th peer's global trust. Thus, a set of peers communicating (providing services) to the *i*-th peer can be reduced to just one virtual representative. It results in obtaining one-to-one evaluation scheme, and the trust of a set will be dominated by peers having higher global trust [30].

The existence and uniqueness of global trust vector as an outcome of aggregation approach is proven in [30]. The closed-form peer's global trust expression lays a basis for direct comparison of global trust values calculated for any two peers in the system (network with N nodes). There is no theoretical explanation of fast algorithm's convergence, but experiments show that it converges fast (about 7 iterations for 100 peers in the network) [30].

Robustness of the algorithm is evaluated regarding behavior of EigenTrust and PowerTrust systems. Several network configurations are considered in [30] such as the ones under the presence of pure malicious peers, peers with unpredictable behavior as well as malicious collectives (groups of peers whose familiarity positively affects their own reputation values diminishing corresponding values of persons outside such groups). It was shown that for first two configurations the performance of Absolute Trust improves significantly as compared to counterparts (by appr. 2% to 4% of authentic

transactions that relate to exchanging files between peers, respectively). As concerns malicious collectives, performances of algorithms are almost identical, with marginal superiority of Absolute Trust over its aforesaid rivals.

The local trust metric in this algorithm can be defined in many ways, and it forms prospects to develop a fuzzy local trust metric. It is worth mentioning that aggregation procedure used in the algorithm can be practically retained. The customizable local trust metric allows to use fuzzy logic approach to extend the algorithm in relatively easy and natural way.

5.5 PeerTrust Algorithm

PeerTrust is another example of P2P reputation system designed specifically for e-commerce communities that are characterized by distinctive problems and risks [6]. L. Xiong and L. Liu identify five important factors that relate to evaluation of peer's trustworthiness as regards supplying other peers with corresponding services. These factors are feedback obtained by a peer, feedback scope (e.g. total number of transactions occurred between peers), credibility of feedback source, transaction context aimed at drawing distinction between extremely crucial and less important or uncritical transactions, and community context to address community-wide characteristics and vulnerabilities. Based on formalization of these parameters, the authors proposed a peer's *j* trust value (metric) $T(j)$ consisting of two parts [6]. The first one is a weighted multiplicative combination of amount of satisfaction peer obtained after realization of each transaction, adaptive transaction context for *i*-th transaction of a peer and credibility of the feedback received from peers. Community context factor constituting the second part of $T(j)$'s expression increases or decreases the impart of the first part to trust value owing to allowance of distinctive community's features. The proposed metric $T(\cdot)$ should be considered as a general form, in which corresponding parts can be 'tuned' in terms of parameters and factors used [6]. Every part of the metric can be implemented differently – alternatives of possible credibility measure metrics (trust value/TVM, personalized similarity/PSM) are presented by L. Xiong and L. Liu in their paper.

Speed of convergence and complexity of PeerTrust algorithm appreciably depend on metrics definitions and specific implementation strategies. In general, the performance of system under PSM metric is a bit worse than in case of TVM, but on the other hand, the former provides better results as the number of peers in the network is increasing. System's robustness is assessed on the grounds of effectiveness against malicious behavior of peers comparing to conventional algorithm, in which the average of the ratings is used to measure the trustworthiness of a peer without taking into account the credibility factor. The trust computation error as a root-mean-square error (RMSE) of the computed trust value of all peers and the actual likelihood of peers performing a satisfactory transaction are computed to evaluate the algorithm's performance. PeerTrust with PSM metric ensures striking results as calculated RMSE does not exceed the value of 0.05, and transaction success rate attains virtually unity.

It must be admitted that PeerTrust system is very flexible over the existing possibility to choose local trust metric. Therefore, it seems that the practical application of fuzzy logic approach to handle naturally nascent uncertainty (vagueness) of certain parameters and characteristics in the algorithm looks justified enough. The system also possesses a great potential to incorporate all types of marketplace data, especially through transaction and community context parts of the general metric $T(\cdot)$ that afford means of broad coverage of manifold system's peculiarities.

6. Experimental part – setup stage. General comments

For the experimental part of the study, we implemented a simulator (in Python 3.7), and the section describes the general simulation setup, including the community setting, peer behavior pattern, and trust computation.

We assume that hypothetical (simulated) community consists of N peers, for which two peer types are defined, namely, they are *honest* and *strategic*, or malicious, peers [36]. The first one embraces those commitment long-run players focused on cooperation, since the latter maximizes player's lifetime payoffs, if the player consistently sticks to action in long-range outlook. In contrast, opportunistic player who cheats whenever the occasion is beneficial for him is bound to a strategic type [6]. The percentage of malicious peers in the community we denote by K. It is reasonable that behavior pattern of good peers is to always cooperate and provide honest feedback after each transaction. However, a correct modeling of malicious peers behavior is a bit challenging task that may require certain simplifications. In particular, we may consider that malicious peers always cheat during transactions and give dishonest ratings to other peers, i.e. they rate negatively a peer who cooperates and provides good rating to a peer who cheats. In case of EigenTrust and HonestPeer algorithms there are also pre-trusted peers that play an important role from the standpoint of algorithm's consistency. Respective PRE_TRUSTED parameter stands for the percentage of pre-trusted peers that relate to good peers only. In general, *behavior pattern of peers* is a topic on a slippery ground, i.e. it can be placed among those aspects of models of reputation systems that require close scrutiny. Why? Potentially, the above cited pattern is definitely not unique, so in order to make models viable other feasible options must be addressed hereafter with great care.

We may also assume that community has CAT categories of services that are provided by peers. From amongst these categories each peer is interested only in a specific subset having the cardinality not less than S. Each category is associated with at least P percents of peers in the community. When a peer queries a service of a specific category, only peers associated with this category can respond to such query. At that, two transaction settings are simulated – they are random and trusted. Random (or, *simple*) setting means that a peer, which responds to the query, is selected randomly (uniform distribution is used) from a set of all peers that can provide queried category of service. In trusted setting the responder is also selected randomly from all peers that can respond to the query, but it is done with respect to their reputation, i.e. a peer with higher reputation has better chances to be chosen. If there are peers with zero reputation, then there is a 10% chance that the responding peer will be chosen uniformly from those peers. It efforts the opportunity for new peers to start building up their reputation.

Binary feedback system is used to evaluate peer after each completed transaction. It means that values 0 and 1 are practiced for PeerTrust and Absolute Trust algorithms, -1 and 1 are used in cases of both EigenTrust and HonestPeer approaches. Local trust and reputation computation steps as such depend on the algorithm in use. Some algorithms have their own hyperparameters that must be specified. Default values of parameters are listed in Table 1. Simulation session (cycle) consists of SIM_NUM transactions. Global reputation is updated after every UPDATE_NUM transactions. Experimental results are averaged by 5 cycles of simulation. Although we simulate online marketplace community – usually it is big enough, dozens to hundreds of thousands of peers – experiments are performed under the presence of modest number of peers. It may be considered as a perceptible limitation, however, the main aim of simulation is to obtain those *prior* results that lay down the ground for further analysis of weak/strong points of models considered here in terms of deeper understanding of their potential to incorporate formal representation of uncertainty (imprecision) factors into these models. In real-life environment it seems highly unlikely that the major part of marketplace peers is malicious as it was defined earlier. Therefore, we don't consider in simulation a malicious peers share exceeding 35%.

7. Experimental part – results and their comparison

We introduce a metric that shows the effectiveness of the reputation system as a rate of unsuccessful transactions (RUT). The unsuccess of transactions is bound up with the outcome of those transactions, in which responding peer happens to be malicious. It is obvious that the less value of the metric is the better. Besides, for the time being we *do not* consider PowerTrust algorithm in the empirical study, since it requires more close inspection and implementation cycle.

7.1 Effectiveness against malicious behavior

The objective of conducted experiments is to evaluate the robustness of the reputation systems against peers with malicious behavior. In the first experiment we alter the percentage of malicious peers in hypothetical community from 10% to 35% with other parameters keeping their default values (Table 1).

Table 1. Parameters and their values used in experiments

Affiliation with…	Parameter	Description	Default value
Community setting	N	number of peers	1000
	K	percentage of malicious peers	15
	CAT	number of categories	10
	S	minimal number of categories for each peer	3
	P	minimal percentage of peers associated with each category	5
Simulation setting	SIM_NUM	number of queries in a simulation	10000
	UPDATE_NUM	number of transactions in reputation update cycle	100
EigenTrust & HonestPeer	PRE_TRUSTED	percentage of pre-trusted peers	5
Absolute Trust	GOOD_W	weight of good transactions in local trust	10
	BAD_W	weight of bad transactions in local trust	1

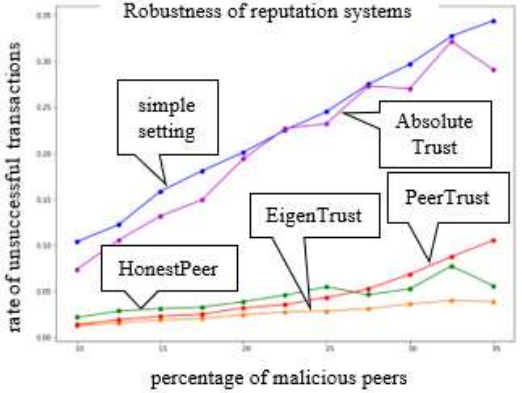

Fig. 1. The growth of rate of unsuccessful transactions depending on the increase of malicious peers' percentage (from 10% to 35%) for different algorithms

As is easy to see in fig. 1, the rate of unsuccessful transactions grows almost linearly with the increase of values (axis x) for simple setting; trusted settings show better results though. EigenTrust, HonestPeer and PeerTrust algorithms show extremely moderate growth of RUT with the increase of malicious peers' percentage. Absolute Trust algorithm demonstrates quite disappointing results characterized by negligible gain (within appr. 2.1% on average) as compared to simple (random) system's case.

7.2 Speed of convergence and scalability

In this set of experiments, we take aim at evaluating the general speed of algorithms convergence and their scalability with regard to the increase of number of peers (fig. 2). As will readily be observed, algorithms PeerTrust and Absolute Trust generally need not more than 2 iterations to converge, while EigenTrust and HonestPeer need to go through 4+ iterations. More than twofold difference on very small values practically equalizes rivals under the conditions of experiment. Thus, all algorithms seem to be quite scalable concerning the number of iterations needed to converge, since the latter does not grow substantially with the increase (from 1,000 to 3,500) of number of peers in the community.

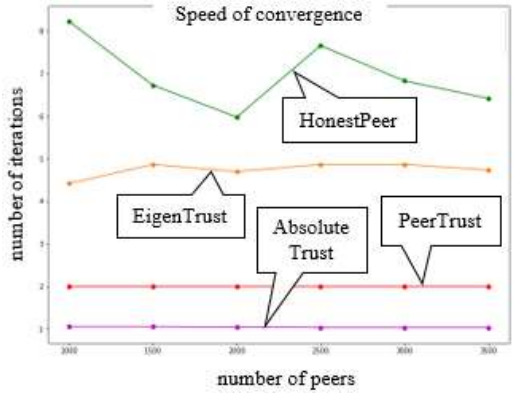

Fig. 2. The speed of convergence (number of iterations needed) of algorithms depending on the number of peers (in the range from 1,000 to 3,500)

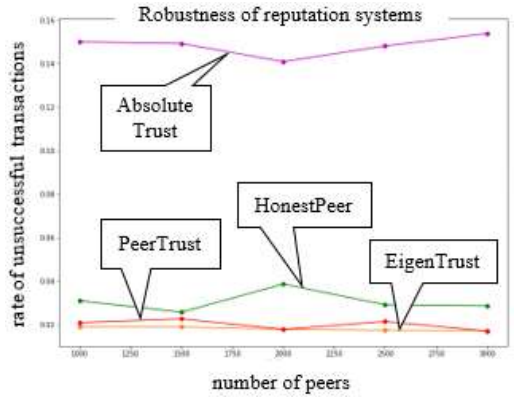

Fig. 3. The speed of convergence (number of iterations) of algorithms depending on the number of peers (in the range from 1,000 to 3,500)

We also evaluate how consistent corresponding systems are against the background of increasing number of peers under «freezing» of other parameters (fig. 3). It can be seen that situation remains almost indistinguishable be it small or bigger community – the rate of unsuccessful transactions mostly remains unchanged in the context of the same malicious peers' percentage.

7.3 Choice of the «best» (most feasible) system

According to the results of experiments summarized above as well as constraints and assumptions put forward, PeerTrust model appears the most robust and effective reputation system among alternatives. It is quite stable regarding the growth of percentage of malicious peers in the community and scalable enough to handle evenly larger number of peers. What is more, local trust metric in PeerTrust system is highly customizable, and this fact simplifies the possibility to extend it with fuzzy factors inhere in reputation systems. In a wide sense we can talk about marketplace data uncertainty that requires much attention in further development of the topic and elaboration of formal aspects of models. Thus, in this instance we opt for PeerTrust system with the object of its modification on the basis of Zadeh's extension principle.

8. Transition from crisp PeerTrust system to Fuzzy PeerTrust system. Is it worthy of notice?

In order to implement fuzzy reputation system, we need to understand above all what data will be represented by fuzzy sets (numbers). In non-fuzzy version of PeerTrust algorithm binary feedback system is used. We suggest utilizing a broader scale to express degrees of satisfaction concerning transaction. It naturally arises from peculiarities of human's perception of information (comments, judgments) – it is not a very convenient and alluring way for humans to think in terms of zeros and ones (or, any other numbers). For the human mind such terms as «bad», «normal» and other resembling options look more understandable and well-suited for interpretation and processing. Being guided by this observation, the new algorithm's feedback can be represented by five verbal degrees of satisfaction, namely, they are *«very bad»*, *«bad»*, *«normal»*, *«good»* and *«very good»*. More fine granulation does not look preferable here, because it may lead to certain confusion in view of human perception of satisfaction's shades – the 'magic' number 7±2 and the seminal paper (1956) by American psychologist George A. Miller on limits on our capacity for processing information straight away cross our mind.

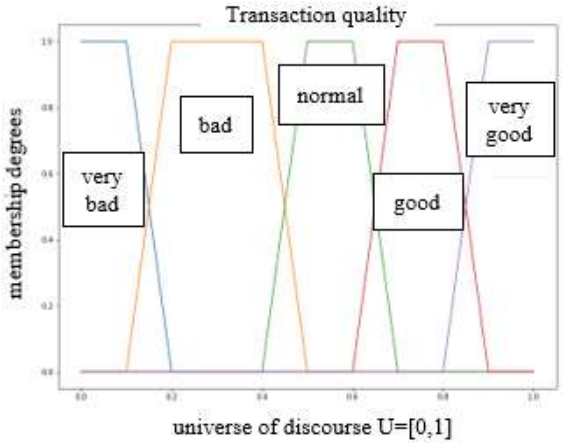

Fig. 4. Linguistic values (trapezoidal membership functions) of the variable 'Transaction quality' (universe of discourse U=[0,1])

Such verbal terms are treated as linguistic values of the variable *«degree of satisfaction»* or *«transaction quality»*; each value can be formally represented by trapezoidal membership function on universe of discourse U=[0, 1] as shown in fig. 4. The type (e.g. Gaussian, bell-shaped, etc.) and the location of fuzzy sets on U may vary noticeably depending on estimates provided by expert

group with reference to characteristic features and implicit shades of model under consideration [22]. The rest of the algorithm remains unchanged, and all specified operations are carried out with fuzzy numbers (intervals) instead of crisp values till the attainment of the defuzzification stage. Defuzzified reputation values are used to choose the responding peer exactly in the same way as described above. In the paper centroid method (COA) is used to obtain those values, but effectiveness and performance of the algorithm may depend distinctly on the chosen defuzzification approach [21, 26].

Here, special attention should be paid to the following: in the paper we consciously consider only one type of data falling under fuzzification, viz. the feedback regarding a buyer. Primarily it is connected with the amount of required modifications and scope of computational experiments to be covered by the text of the limited size. But we are aware that other foregoing types must be addressed thoroughly in the course of the ongoing empirical study.

In conditions of maintenance of community and simulation settings (see the details of conducted experiments described above), but under the imprecision (vagueness) taken into account in the feedback system, we compare the experimental components of Fuzzy PeerTrust with original PeerTrust and EigenTrust algorithms.

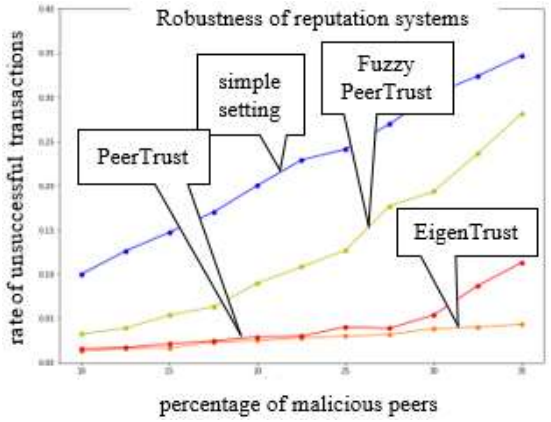

Fig. 5. The growth of rate of unsuccessful transactions depending on the increase of malicious peers' percentage (from 10% to 35%) for EigenTrust, PeerTrust and Fuzzy PeerTrust algorithms

8.1 Effectiveness against malicious behavior (Type-1 fuzzy case)

In the first place, we want to evaluate the robustness of fuzzy modification of PeerTrust system. Experiment settings are retained, the percentage of malicious peers is changing within the range from 10% to 35%. The results as shown in fig. 5 lead to the conclusion that Fuzzy PeerTrust algorithm is definitely more robust in comparison with Simple system. Under small percentage values (appr. interval [10%,18%]) of malicious peers, the performance's characteristic of Fuzzy PeerTrust is close enough to original PeerTrust and EigenTrust. However, it demonstrates worse results than crisp algorithms over the whole range of x-axis values concerned.

8.2 Speed of convergence and scalability

Another set of experiments was aimed at estimation of the speed of convergence of Fuzzy PeerTrust and its scalability in view of the community's growth. As expected, the speed of convergence remains the same as for original PeerTrust with two iterations on average to converge, and it differs essentially from corresponding characteristic (appr. 4.61 on average) of EigenTrust algorithm (fig. 6). In terms of robustness Fuzzy PeerTrust can also be pronounced scalable, since it does not show

significant decrease in quality with the growth of the number of peers in the community (fig. 7). We observe smooth fluctuations of RUT at the level of 0.064. It is worth mentioning that all properties of crisp algorithm remain intact in comparison with its fuzzy counterpart.

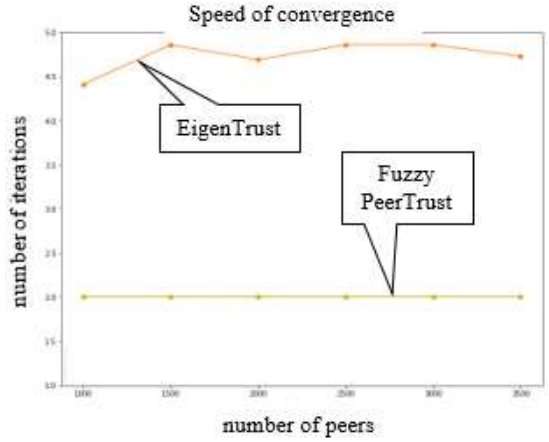

Fig. 6. The speed of convergence (number of iterations needed) of EigenTrust and Fuzzy PeerTrust algorithms depending on the number of peers (in the range from 1,000 to 3,500)

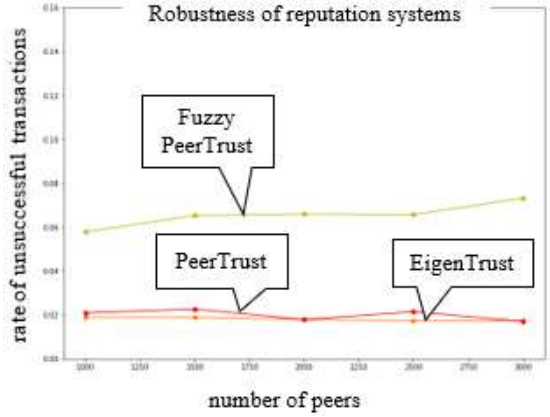

Fig. 7. The growth of rate of unsuccessful transactions depending on the number of peers (in the range from 1,000 to 3,500) for EigenTrust, PeerTrust and Fuzzy PeerTrust algorithms

Computations are initiated with trust vector $t^0 = \left(t_1^0, t_2^0, ..., t_N^0\right)^T = \left(\frac{1}{N}, \frac{1}{N}, ..., \frac{1}{N}\right)^T$ (default case), where N is the number of peers in the community, t_v^0 is a default trust value of a peer v, $v = \overline{1,N}$ [6]. Reputation of a peer v in the form of fuzzy set (number) is denoted as $fuzzy\left(t_v^{i+1}\right)$; $fuzzy\left(\tilde{S}(v,j)\right)$ stands for a feedback (fuzzy number) of peer v concerning j-th transaction; $defuzz(\square)$ signifies the reduction of a fuzzy argument to crisp value (deffuzzification step). To calculate the product of fuzzy number $fuzzy\left(\tilde{S}(v,j)\right)$ and crisp number $t_j^i \Big/ \sum_{k=1}^{I(v)} t_k^i$ as well as the sum (1) of thus obtained fuzzy

numbers, Zadeh's extension principle is used [28,29,37]. As a result, steps to be performed (Algorithm 1 / **case F1**) can be expressed as follows:

Result : t – vector of global trust values

$$t^0 = \left(\frac{1}{N},...,\frac{1}{N}\right)^T,\ i = 0$$

repeat

for $v \leftarrow 1$ **to** N **do**

$$fuzzy\left(t_v^{i+1}\right) = \sum_{j=1}^{I(v)} fuzzy\left(S(v,j)\right) \cdot \frac{t_j^i}{\sum_{k=1}^{I(v)} t_k^i} \qquad (1)$$

$$t_v^{i+1} = defuzz\left(fuzzy\left(t_v^{i+1}\right)\right) \qquad (2)$$

end

$$\sigma = \left\| t^{i+1} - t^i \right\|$$

$$i = i + 1$$

until $\sigma < \varepsilon$;

As already mentioned, it is important to put emphasis on the choice of defuzzification method to use in (2). In general, the step of defuzzification relates to the conversion of a fuzzy quantity expressed in the form of membership function to a crisp number. In this case, we can talk about a diverse group of "fuzzy-to-crisp" data transformation methods, including, in particular, Center of Gravity (COG or centroid), Bisector of Area (BOA), Mean of Maximum (MOM), Smallest of Maximum (SOM) and Largest of Maximum (LOM) standard computational schemes as some of the most commonly used approaches. A rigorous and detailed discussion of defuzzification strategies can be found in [38, 39].

The results of the conducted experiments with Fuzzy PeerTrust under default values of parameters (Table 1) for different defuzzification methods shows that SOM scheme performs significantly better in the presence of smaller standard deviation as compared to other strategies. Intuitively SOM provides better results for the case in hand, because reputation system is punishing malicious peers more 'harshly', and it leads to better isolation of such peers from good peers. At the same time, changing defuzzification method in experimental settings does not affect scalability of the algorithm itself, since as the number of peers increases, the rate of unsuccessful transactions remains unchanged at insignificant fluctuations observed. Overall, we consciously avoid generalizations here, since the competitive advantage of SOM in the given algorithm should be confirmed empirically in the future.

At the same time, an important point of the algorithm shown above is that certain aforesaid attributes of trust and reputation like context-awareness (sensitivity), decrease (of the level) with time, their multifaceted nature (diversity) are not taken into account. We may regard this version of the algorithm as *basic* one (or, *F-basic* if we consider factor of fuzziness in its core); it paves a 'wide' way for algorithm's further revision, amendment and improvement.

9. Switching from using Type-1 fuzzy sets to Interval Type-2 fuzzy sets in reputation systems – the way to deal with uncertainty of expert's assessments

It can be noticed that the shift towards application of type-1 fuzzy sets in algorithms leaves us anyway within the scope of *crisp* real values of membership functions, which are associated with elements from a problem's domain (or, universal set) U. Despite active use of type-1 fuzzy sets in research works and industrial projects for almost forty years, existing publications specifically note that such sets exhibit very limited capabilities for modeling uncertainty, because of $\mu_A(x)$ crispness ($\forall x \in U$) mentioned above [40, 41]. In case of type-2 fuzzy sets, their membership functions are getting fuzzy, i.e. each specific $\mu_A(x)$ becomes associated with more than one value unlike their type-1 counterparts.

The latter allows for representation of vagueness inherent in natural language constructs (words, phrases) that express the assessments made by experts. Following explanations done by German philosopher F.G. Frege, the notion of vagueness relates to so-called «boundary line»; it can be expanded to the case of absence of clear truth conditions that is observed in most practical cases involving human judgments [42].

We represent linguistic values of the variable *«degree of satisfaction»* or *«transaction quality»* by interval type-2 trapezoidal membership functions defined on the universe of discourse U=[0, 1] as shown in Fig. 8. Types of membership functions as well as their location on the universal set U may vary noticeably depending on estimates provided by members of expert group [22]. It is important to note that type-2 fuzzy sets may appear due to natural slight differences in expert assessments and aggregation methods applied to them. As we can see, each of five functions shown in fig. 8 is bounded by two type-1 functions called upper (UMF) and lower (LMF) membership functions. Each function's 'thickened' boundary (footprint of uncertainty, FOU) is formed by *primary* interval-shaped memberships $\mu_A(x) \subset [0,1]$ ($\forall x \in U$) that can be seen as a collection of vertical slices of original type-2 function.

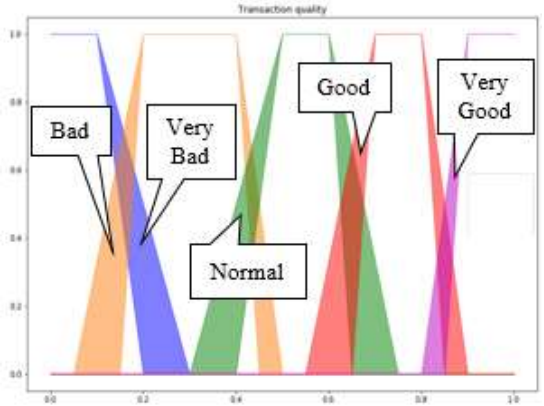

Fig. 8. Linguistic values of the variable 'transaction quality' (universe of discourse U=[0,1]) represented in the form of interval type-2 trapezoidal membership functions

Corresponding *secondary* function is connected to each interval $\mu_A(x)$, i.e. secondary membership functions are defined on the whole set of $\mu_A(x)$ for each type-2 function under consideration. The usual two-dimensional portrayal of type-2 MFs implies their 3D-view, which is determined by the presence of secondary grades. In the present study, the focus is restricted to interval type-2 fuzzy

sets (IT2FS), for which all secondary grades are equal to one. Being a convenient uncertainty modeling tool to capture representation of heterogeneous verbal responses formed within the group of domain experts, such functions are actively used when solving various practical problems due to their well-developed theoretical foundations and sound computational efficiency. If preceded by shown trapezoidal membership functions, the thicknesses of areas bounded by their pieces (LMF and UMF) convey degrees of uncertainty as a result of aggregation of converted to numeric form expert opinions concerning perception of values of the linguistic variable *«transaction quality»*.

Switch to using these functions in models leads to modification of the aforesaid *F-basic* algorithm, in which all operations are performed on interval type-2 fuzzy sets until defuzzification stage is reached. Steps to be performed can be expressed now (Algorithm 2 / **case F2**) as follows:

Result : t – vector of global trust values

$$t^0 = \left(\frac{1}{\mathrm{N}},...,\frac{1}{\mathrm{N}}\right)^{\mathrm{T}},\ \mathrm{i}=0$$

repeat

for $v \leftarrow 1$ **to** N **do**

$$fuzzy\left(t_v^{i+1}\right) = \sum_{j=1}^{I(v)} fuzzy\left(S(v,j)\right) \cdot \frac{t_j^i}{\sum_{k=1}^{I(v)} t_k^i} \qquad (1)$$

$$fuzzy\left(t_v^{i+1}\right) = reduce\left(fuzzy\left(t_v^{i+1}\right)\right) \qquad (1')$$

$$t_v^{i+1} = defuzz\left(fuzzy\left(t_v^{i+1}\right)\right) \qquad (2)$$

end

$\sigma = \left\| t^{i+1} - t^i \right\|$

$\mathrm{i} = \mathrm{i} + 1$

until $\sigma < \varepsilon$;

It is proposed to implement modifications by means of «type-2 to type-1» type reduction (1′) to obtain the averaged trapezoidal membership function (resultant type-reduced set); afterwards, the latter is defuzzified. It is noteworthy that type reduction algorithms are the topical area of research, so extra experiments related to realization of defuzzification are an essential component of further extension of the work.

Results of comparing fuzzy Type-2 PeerTrust with other algorithms are shown in fig. 9 and 10. Just as expected, they're comparable to the performance of Fuzzy PeerTrust algorithm. Better results as compared to Simple case are quite predictable; there is a close enough resemblance to the original PeerTrust and EigenTrust systems, especially for percentage of malicious peers in the range from 10% to 20%. In average, fuzzy Type-2 PeerTrust shows slightly worse rates than crisp systems, although it is possible to find an intuitive explanation for that.

Consecutive application of type reduction and defuzzification procedures may lead to certain "displacement" of calculated values in comparison to original crisp models. It should not be considered as a shortcoming of the system; it is a fact that must be taken into consideration when using IT2FS. Potentially, it makes sense to use several type reduction and defuzzification procedures in every case in question. Calculations will obviously become more time-consuming but will allow

to take account of existing uncertainty factor in a more complete manner, leading to obtaining interval results rather than exact points. As shown in Fig. 10, the adoption of IT2FS in models does not affect the scalability of system in the context of experimental conditions.

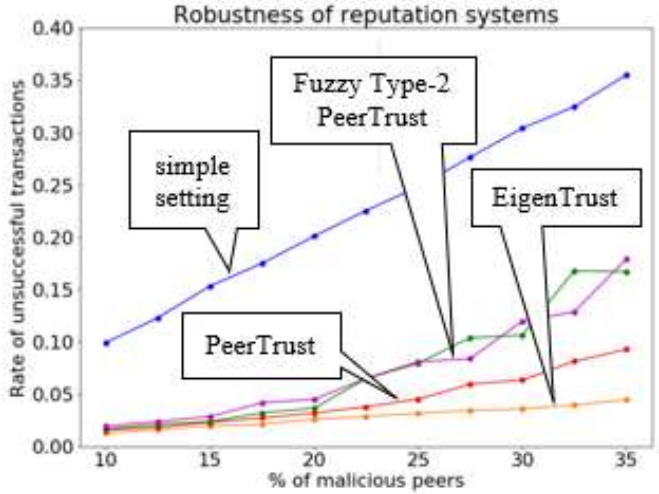

Fig. 9. The rate of unsuccessful transactions depending on the increase of malicious peers percentage (from 10% to 35%) for EigenTrust, PeerTrust and Fuzzy Type-2 PeerTrust algorithms

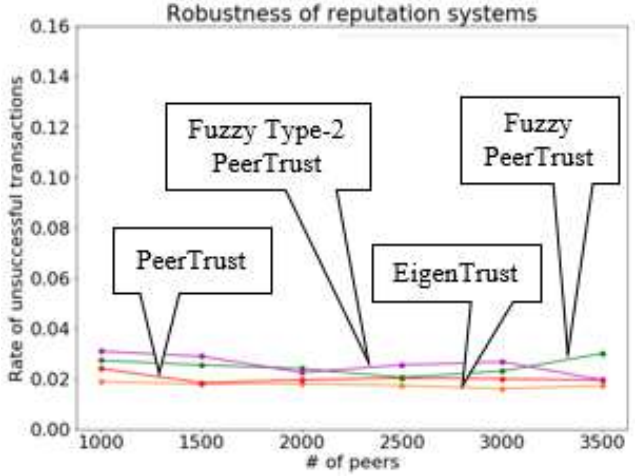

Fig. 10. The rate of unsuccessful transactions depending on the number of peers (in the range from 1,000 to 3,500) for EigenTrust, PeerTrust, Fuzzy PeerTrust and Fuzzy Type-2 PeerTrust algorithms

The results attained enable to speak decidedly about existing perspectives of fuzzy logic approach's application in reputation systems (corresponding algorithms), whether type-1 or interval type-2 fuzzy set is considered. Even despite somewhat higher computational costs compared to original crisp algorithms, greater transparency, better perceptibility by humans, interpretability and flexibility from a viewpoint of verbal expression and formalization of the scores provide a basis for

further studying of the topic. The present paper can be considered as a mere first step in this direction.

10. Conclusion

E-commerce is a fast-growing market that implies continual and utterly active communication between users being 'strangers' to each other on numerous occasions. Because of that it is essential to establish *reputation systems* to better handle available online information with the object to more accurately discern trustworthy and non-trustworthy players in systems creating grounds for peers to be more careful about their reputation. By the far-famed example of eBay reputation system, even relatively simple ones show themselves as very helpful from the viewpoint of malicious behavior's limitation and trustability increase. Online marketplaces that became immensely popular in the last 10-15 years as sites offering wide enough range of reasonably priced goods from various sources can be also considered as P2P networks. There is a vast range of reputation systems developed for P2P networks (mostly aimed at file-sharing) that can be adapted to e-commerce.

The main problem that is covered in the paper relates to the fact that none of these systems work with uncertainty (blurriness) of marketplace data and vagueness typical for notions of trust and reputation. Uncertainty in different forms of its manifestation is definitely inherent in reputation systems, and some of those forms can be addressed by fuzzy logic. This very inhesion, but not disconfirmed artificial desire, has served as an impellent factor to start this study.

Most likely, it can be argued that it is difficult to identify on the basis of several singled out key criteria unconditional leader among analyzed systems (algorithms EigenTrust, Absolute Trust, HonestPeer, PowerTrust and PeerTrust), since each of them has positive aspects as well as drawbacks. All algorithms, except PowerTrust, were implemented (Python 3.7) under the same conditions discussed in detail in the paper with the purpose of comparing fairly their relative performance. For reasons partially covered in the paper, Absolute Trust and PeerTrust systems were prudently regarded from the standpoint of their robustness and scalability as front-runners, i.e. candidates for reasoned fuzzification. Besides undertaking comparative analysis of those five significant and most popular reputation systems, the paper makes a mark for transition from crisp system (by the example of PeerTrust algorithm) to its fuzzy counterparts. The latter provided for an approach based on the use of type-1 (T1FS) and interval type-2 fuzzy sets (IT2FS).

Corresponding fuzzy models (we call them provisionally *F-basic* and modified *F-basic* algorithms – **cases F1** and **F2**, correspondingly, as they are denoted above) consider now only one characteristic of trust and reputation, namely, it is transaction quality or degree of peer's satisfaction. Other important attributes like context-awareness (sensitivity) or decrease of trust's level with time were not scrutinized yet. Nevertheless, initial experimental results attained in line with the fact of constant presence and active use of verbal assessments in reputation systems confirm the need to continue research in this field. Verbal forms are both habitual and convenient for human's perception despite of intrinsic vagueness and uncertainty. That is why, fuzzy logic approach, to the opinion of authors, has good prospects for both close examination and use in reputation systems.

At the same time, it should be mentioned that there are immediate tasks related to fuzzy models that require primary attention. The choice of shapes of membership functions and their fine tuning (location on the universe of discourse) on the basis of either existing data or expert assessments, a more detailed study of the potential of models using IT2FS as well as the use of different type reduction and defuzzification strategies are amongst the most topical ones.

References

[1]. The Next Scoop, 2018. E-Commerce is Growing at an Unprecedented Rate All over the Globe - The Next Scoop, web-resource: https://thenextscoop.com/e-commerce-is-growing-at-an-unprecedented-rate-all-over-the-globe/ (access date 17.02.19).

[2]. Statista, 2018. E-commerce Share of Total Retail Sales in United States from 2013 to 2021, web-resource: https://www.statista.com/statistics/379112/e-commerce-share-of-retail-sales-in-us/ (access date 26.02.19).
[3]. The Next Scoop, 2018. 2019 E-commerce Trends, Statistics and Metrics, web-resource: https://thenextscoop.com/ecommerce-trends-statistics-and-metrics-2019/ (access date 14.02.19).
[4]. Statista, 2018. Global Retail E-commerce Market Size 2014-2021, web-resource: https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/ (access date 08.02.19).
[5]. Forbes.com, 2017. What Are Online Marketplaces and What Is Their Future? web-resource: https://www.forbes.com/sites/richardkestenbaum/2017/04/26/what-are-online-marketplaces-and-what-is-their-future/#704431c13284 (access date 06.02.19).
[6]. Xiong L., Liu L. PeerTrust: Supporting Reputation-Based Trust for Peer-to-Peer Electronic Communities. IEEE Transactions on Knowledge and Data Engineering, vol. 16, no. 7, 2004, pp. 843-857.
[7]. eBay, 2019. web-resource: www.ebay.com (access date 15.03.2019).
[8]. Kamvar S., Schlosser M., Garcia-Molina H. The EigenTrust Algorithm for Reputation Management in P2P Networks. In Proc. of the 12th International Conference on World Wide Web, 2003, 640-651.
[9]. Kurdi H. HonestPeer: An Enhanced EigenTrust Algorithm for Reputation Management in P2P Systems. Journal of King Saud University - Computer and Information Sciences, vol. 27, no. 3, 2015, 315-322.
[10]. Zhou R., Hwang K. PowerTrust: A Robust and Scalable Reputation System for Trusted Peer-to-Peer Computing. Proc. IEEE Transactions on Parallel and Distributed Systems, vol. 18, no. 4, 2007, 460-473.
[11]. Çelikyılmaz A., Türkşen I.B. Modeling Uncertainty with Fuzzy Logic. With Recent Theory and Applications. Studies in Fuzziness and Soft Computing, vol. 240, 2009, 311 p.
[12]. English Oxford Living Dictionaries, web-resource: https://en.oxforddictionaries.com/ (access date 04.03.19).
[13]. Alam F., Paul A. A Computational Model for Trust and Reputation Relationship in Social Network. In Proc. of the 5th International Conference on Recent Trends in Information Technology, 2016, pp. 1-6.
[14]. Wang Y., Vassileva J. Bayesian Network Trust Model in Peer-to-Peer. Lecture Notes in Computer Science, vol. 2872, 2003, pp. 23-34.
[15]. Gambetta D. Can We Trust Trust? Chapter - Trust: Making and Breaking Co-operative Relations, Dept. of Sociology, University of Oxford, 1980, pp. 213–237.
[16]. Alfarez A.-R., Hailes S. Supporting Trust in Virtual Communities. In Proc. of the 33rd Annual Hawaii International Conference on System Sciences, 2000, pp. 1-9.
[17]. Kreps D.M., Wilson R., 1982. Reputation and Imperfect Information. Journal of Economic Theory, vol. 27, 253-279.
[18]. Hussain J.K., Hussain O.K., Chang E. An Overview of the Interpretations of Trust and Reputation. In Proc. of the IEEE Conference on Emerging Technologies and Factory Automation (EFTA-2007), 2007, pp. 826-830.
[19]. Chiluka N., Andrade N., Gkorou D., Pouwelse J., 2012. Personalizing EigenTrust in the Face of Communities and Centrality Attack. Proc. IEEE 26th Int. Conference on Advanced Information Networking and Applications, 503-510.
[20]. Zhang J. Trust Management Based on Fuzzy Sets Theory for P2P Networks. In Proc. of the WRI World Congress on Software Engineering, 2009, pp. 461-465.
[21]. Semenkovich S., Kolekonova O., Degtiarev K. A Modified Scrum Story Points Estimation Method Based on Fuzzy Logic Approach. Trudy ISP RAN/Proc. ISP RAS, vol. 29, issue. 5, 2017, pp. 19-38.
[22]. Zadeh L. The Concept of a Linguistic Variable and Its Application to Approximate Reasoning – I. Information Sciences, vol. 8, no. 3, 1975, pp. 199-249.
[23]. Zadeh L. A. Fuzzy Logic, Neural Networks and Soft Computing. Communications of the ACM, vol. 37, no. 3, 1994, pp. 77-84.
[24]. Zadeh L.A. Fuzzy Languages and Their Relation to Human and Machine Intelligence. In Proc. of the International Conference on Man and Computer, 1972, pp.130-165.
[25]. Zadeh L.A. Similarity Relations and Fuzzy Orderings. Information Sciences, vol. 3, no. 2, 1971, pp. 177-200.
[26]. Zimmermann H.-J. Fuzzy Set Theory – and Its Applications, 4th ed., Springer Science+Business Media, LLC, 2001.
[27]. Zadeh L.A. Fuzzy Sets. Information and Control, vol. 8, no. 3, 1965, pp. 338-353.
[28]. de Barros L.C., Bassanezi R.C., Lodwick W.A. The Extension Principle of Zadeh and Fuzzy Numbers. In A First Course in Fuzzy Logic, Fuzzy Dynamical Systems, and Biomathematics, Studies in Fuzziness and Soft Computing, vol. 347, 2017, pp 23-41.

[29]. Ross T.J. Fuzzy Logic with Engineering Applications, 3rd ed., John Wiley & Sons, 2010, 595 p.
[30]. Awasthi S.K., Singh Y.N. Absolute Trust: Algorithm for Aggregation of Trust in Peer-to-peer Networks, 2016, web-resource: http://arxiv.org/abs/1601.01419 (access date 17.03.2019).
[31]. Rao S., Wang Y., Tao X. The Comprehensive Trust Model in P2P Based on Improved EigenTrust Algorithm. In Proc. of the International Conference on Measuring Technology and Mechatronics Automation, 2010, pp. 822-825.
[32]. Song S., Hwang K., Zhou R., Kwok Y.-K. Trusted P2P Transactions with Fuzzy Reputation Aggregation. IEEE Internet Computing, vol. 9, no. 6, 2005, pp. 24-34.
[33]. Page L., Brin S., Motwani R., Winograd T. The PageRank Citation Ranking: Bringing Order to the Web, Technical Report, Stanford Digital Library Technologies Project, 1998.
[34]. Haveliwala T.H., Kamvar S.D. The Second Eigenvalue of the Google Matrix, Technical Report, Stanford University, 2003.
[35]. Faloutsos M., Faloutsos P., Faloutsos C. On Power-Law Relationship of the Internet Technology. In Proc. of the ACM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication (SIGCOMM-1999), 1999, pp. 251-262.
[36]. Dellarocas C. The Digitization of Word-of-Mouth: Promise and Challenges of Online Reputation Mechanism. Management Science (Special Issue on E-Business and Management Science), vol. 49, no. 10, 2003, pp. 1407-1424.
[37]. Klir G., Yuan B. Fuzzy Sets and Fuzzy Logic Theory and Applications, Prentice-Hall/Upper Saddle River, 1995, 592 p.
[38]. Tóth-Laufer E., Takács M. The Effect of Aggregation and Defuzzification Method Selection on the Risk Level Calculation. In Proc. of the EEE 10th Jubilee International Symposium on Applied Machine Intelligence and Informatics (SAMI), 2012, pp. 131-136.
[39]. Roychowdhury S., Pedrycz W. A Survey of Defuzzification Strategies. International Journal of Intelligent Systems, vol. 16, no. 6, 2001, pp. 679-695.
[40]. Mendel J.M. Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions, 1st ed. Prentice Hall, 2001, 674 p.
[41]. Mendel J.M., John R.I. Type-2 Fuzzy Sets Made Simple. IEEE Transactions on Fuzzy Systems, vol. 10, no. 2, 2002, pp, 117-127.
[42]. Dubois D. Fuzziness, Uncertainty and Vagueness: Toward a Less Blurry Picture (Is a Fuzzy Set Vague?), 2008, web-resource: https://www.logic.at/lomorevi/LoMoReVI/transvague.pdf (access date 14.06.2019).

Информация об авторах / Information about authors

Mikhail Mikchailovitch NOSOVSKIY is a student of the Bachelor's degree program «Software Engineering» at the National Research University Higher School of Economics (HSE), Moscow, Russia. His research interests include fuzzy modeling, data analysis and identification of fraud activity.

Михаил Михайлович НОСОВСКИЙ является студентом образовательной программы бакалавриата «Программная инженерия» в Национальном исследовательском университете «Высшая школа экономики» (НИУ ВШЭ), Москва, Россия. Его исследовательские интересы включают в себя нечеткое моделирование, анализ данных и выявление фродовой (мошеннической) активности.

Konstantin Yurievitch DEGTIAREV earned his M.S. degree in applied mathematics ('engineer-mathematician' qualification) from Moscow Institute of Electronic Engineering, Russia, and his Ph.D. degree in engineering sciences from Moscow Academy of Instrument Engineering and Informatics, Russia. He is currently an Associate Professor at the Software Engineering Department of the Faculty of Computer Science at the National Research University Higher School of Economics (HSE) in Moscow. His research interests include fuzzy logic/soft computing in system analysis, verbal computing, perceptions and representations in system analysis, use of fuzzy time series in forecasting. He is a Member of the IEEE (Systems, Man and Cybernetics Society).

Константин Юрьевич ДЕГТЯРЕВ получил степень магистра (специалитет) прикладной математики в Московском институте электронного машиностроения и степень кандидата

технических наук в Московской академии приборостроения и информатики, Россия. В настоящее время он является доцентом кафедры программной инженерии факультета компьютерных наук Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ) в Москве. Его исследовательские интересы включают нечеткую логику/мягкие вычисления в системном анализе, вербальные вычисления, восприятие и представление в системном анализе, применение нечетких временных рядов в прогнозировании. Он является членом IEEE (общество 'Systems, Man and Cybernetics').