Towards Quality of Experience-based Reputation Models for Future Web Service Provisioning

Tomasz Ciszkowski[1], Wojciech Mazurczyk[1], Zbigniew Kotulski[1,2], Tobias Hoßfeld[3], Markus Fiedler[4], Denis Collange[5]

[1] Institute of Telecommunications, Warsaw University of Technology, Warsaw, Poland, {T. Ciszkowski, W.Mazurczyk}@tele.pw.edu.pl
[2] Polish Academy of Sciences, Institute of Fundamental Technological Research, Warsaw, Poland, zkotulsk@ippt.gov.pl
[3] University of Würzburg, Institute of Computer Science, Würzburg, Germany, hossfeld@informatik.uni-wuerzburg.de
[4] Blekinge Institute of Technology, Department of Telecommunication Systems, Karlskrona, Sweden, markus.fiedler@bth.se
[5] France Télécom R&D, Sophia-Antipolis, France, denis.collange@orange-ftgroup.com

Abstract. This paper concerns the applicability of reputations systems for assessing Quality of Experience (QoE) for web services in Future Internet. Reputation systems provide mechanisms to manage subjective opinions in societies and yield general scoring of a particular behavior. Thus, they are likely to become an important ingredient of Future Internet. Parameters being under evaluation by reputation system may vary greatly and, particularly, may be chosen to assess the users' satisfaction with (composite) web services. Currently, this satisfaction is usually expressed with QoE, which represents subjective users' opinions. The goal is to predict users' satisfaction based on reputation values. This may be beneficial for service providers in terms of checking the fulfillment of SLAs, for retaining QoE on the satisfaction level for other users sharing the same network or service resources, and for providing the users with indications which resource to choose in order to maximize its experience.

Keywords: reputation systems, QoE, web services, Future Internet

1 Introduction

A significant component of contemporary web services is a mutual user interaction which enables to create flexible and customizable web applications. In some extent, especially for social web services provision a middleware part is operated by provider whilst the value added content is delivered by service users. This concept features in Web 2.0 paradigm [32], which introduces a certain level of complexity to the service design and imposes new challenges for Application and Internet Service Providers

(ASPs and ISPs). Distinction between service and its user becomes hazy and may lead to the inevitable lost of manageability on provider side, resulting in service quality degradation or putting user businesses in an unacceptable danger. Widely used countermeasures to the given challenge are *reputation systems* [15], [30], [31]. Apart from the service regulations and its strong security mechanisms the reputation systems perform an online monitoring of user activities and stimulate the associated community for a good behaviour. They also support accountability for malicious attitudes, which is often linked with a lowering the service quality. The reputation systems are enriched with a mechanism of sharing a *reputation scoring* on a particular entity among all interested users, service providers and may be considered in the future for better web service provision. The reputation scoring usually reflects an aggregated subjective opinion on a party but tends to depend on user's activity, time scale and web service context [31]. The variety of party parameters being under evaluation by reputation system may be related to the certain key attributes of the service such as audio-video content and form the user satisfaction which is expressed by a set of QoE (Quality of Experience) metrics of the service [21].

E-communities are dependent on online entertainment, trade and communication which is spread over the Internet services. E-commerce such as Amazon [1], eBay [2], information and social web portals (Wikipedia, Facebook, Google, MySpace, Flickr) incorporate advanced Web 2.0 mechanism for customizable content presentation, share and delivery. A versatile mechanism for usage of applications and computing resources as a service, called *cloud computing* [33] has moved the margin of the computer user experience from the ordinary simple desktop applications to enriched, any-where and any-time, broadly portable web applications. This has been forcing service providers to face challenges in assuring market's demanding attributes such as network-centric Quality of Service (QoS), the application design dependent strong security as well as ergonomics, which fit the human needs. These compose a mix of objective and subjective metrics of the service quality and need appropriate evaluation methods that reputation systems may deal with.

The way of building the reputation distinguishes its subjective and objective nature. For instance in MANETS [16] a locally created reputation [15], [34] reflects generalized opinion on the truthfulness of peers in the network but still remains a subjective measure of the service in a local neighborhood. It is a characteristic property of distributed reputation systems. In contrary a centralized reputation repository performs a global generalization and assuming that parties of subjective opinions are not related the reputation systems, yield an objective scoring [35]. A diversity of contemporary Web services may impose a need to adapt reputation metrics with use of collaborative filtering in order to get accurate and context aware subjective measure: the subjective reputation [36]. Collaborative filtering shapes the reputation in order to emphasize and share the characteristic features of subjective metrics and allows distinguishing the interpretation of a particular reputation for example *user reputation* and *network reputation*. Operators, for example, might be interested in both subjective and objective network reputation. Subjective reputation reflects the individual customer's view on the quality and price-worthiness of a service and is strongly related to the risk for churn, i.e., the risk of leaving the operator because of dissatisfaction. However, the operator has to treat the potential risk from single unhappy customers against the overall service quality, which is typically limited by the margin between income and investments. In this context, even the reputation of the (complaining) user might be of interest; a reasonable user's judgment might be weighted higher than that of a well-known grouch.

The remainder of the paper is structured as follows. Section 2 presents and illustrates potential classification criteria for reputation systems. Section 3 provides an overview of relevant QoE models. Section 4 addresses the use of reputation systems for QoE evaluation. Section 5 concludes the paper.

2 Reputation Systems Classification

In the literature one can find a number of proposed reputation systems applied in different protocols and services. Assuming that the reputation systems are consistent and applied adequately to the requirements of a service or a protocol, they can have a variety of properties that make them more or less useful for other services, also the web services that are the subject of this paper. In the particular case of web services one should take into account the specific constraint that the quality of the service can be considered from at least three points of view. They are: the quality (value) of the website's content, its quality of presentation (represented by, e.g., structure of the page HTML code) and, finally, by the quality of transmission. Therefore, before the application of a reputation system, one must precise the model assumptions to fix what is the real subject of the analysis.

The reputation systems can be constructed using different mathematical techniques. Listing the most popular we have: *probabilistic systems* that describe the reputation as the probability of an expected reaction of a system (party) on a request, *fuzzy theories based systems* where the reputation is a fuzzy number established on a subjective opinion of customers, and *deterministic systems* where the reputation is expressed as an arbitrary number from an assumed range (mainly from the 0…1 interval). Inside of each of the above categories the systems can be classified according to specific mathematical techniques applied for calculating the reputation and, especially, consolidating it in time and using data collected from many sources. Concerning global reliability, the reputation can be classified as *objective* or *subjective*, with possibly some intermediate states, referred as *hybrid*. We denote the reputation system as objective if the reputation is calculated according to knowledge collected independently of the service's customers or is based on the information collected from statistically large group of users.

The next important property of a reliable system is its *sensibility* in time that is how quickly the system reacts on changes in reputation of the service. This is strongly connected with the performance of the system on one hand and its time memory (memory of the reputation history) on the other. This property is especially important when we want to decide how quickly some critical events should affect the reputation and when it should be forgotten. This mechanism properly implemented enables the reputation evolution and rehabilitation of parties with low ratings. Finally, one can consider the *architecture* of the reputation system. Roughly speaking, the reputation system can be *centralized* or *distributed* (decentralized). This can be understood in two ways. Firstly, we can identify if the reputation data is *collected* and transformed in one or in many places. Secondly, the calculation of reputation can be *controlled* by one entity or it can be the result of cooperation of many independent entities.

Analyzing the above properties we can observe that most of them are independent and some of them are in contradiction. The reputation system constructed for a specific service or protocol must have such properties that are the most adequate for its functioning and give the most useful information for its manager. Since the purpose of this paper is construction of the reputation system for supporting QoE for web services, we propose a systematic classification of the reputation systems that can help in appropriate selection of the reputation system.

Reputation systems (RSs) for web services and telecommunication networks can be classified, based on how the reputation is assessed, as objective, subjective or hybrid [3]. Subjective reputation systems (SRSs) create reputation by using measures formed based on subjective opinions and experiences provided by users of such systems. These measures are expressed e.g. in form of rating (or score). Examples of SRSs are eBay [2] or Amazon Auctions [1]. To create reputation measures objective reputation systems (ORSs) rely on ratings that have been assessed based on objective, well-defined, repeatable criteria. The reputation values of the users are created using objective evidence as can be seen by the whole community. Example of such objective reputation system is Amazon book sales. Hybrid reputation systems (HRSs) combine characteristics of both SRSs and ORSs. In most cases these systems rely on ORS but the obtained reputation objective scores are interpreted by using subjective values and motivations. Example of hybrid reputations system may be individual rating of e.g. books apart from their real (objective) sale history.

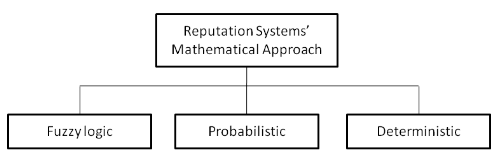

Fig. 1. Classification of RSs based on mathematical approach

Moreover, reputation systems can be divided based on the mathematical approach to reputation creation (Fig. 1). The rest of this section is devoted to describe reputation systems based on provided classification.

2.1 SRs based on probabilistic approach

For reputation systems based on probabilistic approach there are mainly two possibilities to consider: Bayesian networks and Subjective Logic.

Bayesian networks

Bayesian networks in reputation systems enable a theoretically sound basis for computing reputation scores. Bayesian reputation systems are based on computing reputation scores by statistical updating of probability density functions (PDF). Two types of Bayesian RS are proposed – binomial [4], which is based on Beta PDF, and multinomial [5], which is based on Dirichlet PDF. The updated reputation score is calculated by combining historical reputation score with the new rating.

Serious limitation of binomial reputation systems is that they accept as inputs only binary ratings (positive or negative) and cannot reflect polarized ratings. Multinomial reputation systems overcome this drawback and can have any number of rating levels, and represent polarized ratings.

For binomial reputation systems input parameters are provided as scalars. Historical data is involved in form of *longevity factor* and positive and negative evidence sum in reputation score calculation. For multinomial reputation systems input parameters are provided as vectors. Historical data is also involved in form of longevity factor and accumulated evidence in a given period of time in reputation score calculation.

Bayesian reputation systems collect ratings about users or service providers from members in a community. Then they are send to the central location i.e. reputation centre, where the reputation scores are computed and published. After the ratings about particular users are received, these users' reputation value will change accordingly. That is why Bayesian reputation systems are objective.

Subjective logic

Subjective logic [8] is a part of probabilistic logic, which includes in calculations uncertainty, belief and disbelief. Inputs and outputs in Subjective Logic are subjective opinions about states in a state space. This approach may be utilized to form a reputation system like presented e.g. in [7]. In this paper opinions are denoted as ω^{A}_{x} and:

- $\omega^{A}_{x} = (b, d, u, a)$ for binomial opinions which expresses party *A*'s belief in the truth of statement *x*. Scalar *b* represents belief, *d* represents disbelief, *u* represents uncertainty and *a* is the base rate (a parameter which is then used to compute an opinion's probability expectation value).
- $\omega^{A}_{X} = (\vec{b}, u, \vec{a})$ for multinomial opinions which expresses the relying user's *A* belief over a state space *X* and $\vec{a}, \vec{b}, u \in [0, 1]$. Vector $\vec{b}$ represents belief masses over the states of *X*, scalar *u* represent uncertainty mass ($\sum \vec{b} + u = 1$) and $\vec{a}$ represents the base rates over *X*, and $\vec{a}$ is used for computing the probability expectation value of a state *x* ($E(x) = \vec{b}(x) + \vec{a}(x)\,u$, so the vector $\vec{a}$ determines how uncertainty contributes in this equation).

When reputation values are expressed as subjective opinions, each transitive reputation path may be computed with the *discounting* operator. Moreover, paths can be combined using the cumulative or averaging *fusion* operator. These operators form part of Subjective Logic.

Reputation systems based on Subjective Logic are used to derive local and subjective reputation scores so they are applicable mainly to distributed systems.

Bayesian reputation systems are compatible with reputation systems based on Subjective Logic. Combining these two mathematical approaches provides a powerful basis for assessing the quality of online services, in particular, web services. Such reputation system is proposed in [6]. This compatibility between Bayesian reputation systems and Subjective Logic creates a flexible tool and allows creating reputation values that consist of both subjective and objective ratings. This may be the most suitable solution for modeling QoE from reputation systems based on probabilistic approach presented earlier.

2.2 SRs based on fuzzy logic approach

Fuzzy logic is an attempt of rigorous mathematical treatment in situations where a model should reflect individual's opinions but one cannot collect sufficiently large statistical data (experience). Such cases are typical when we try to classify rare events and therefore fuzzy methods found their place in reputation models. Here we present several fuzzy logic based methods of calculation of reputation and identify their properties useful for the web services case. All of them use fuzzy measures to express trust and reputation. They differ in how the individual trust of one entity to another is expressed, what kinds of reputation are considered and how individual and social reputations are aggregated to obtain the effective reputation used in decision-making.

Good example of the fuzzy logic based reputation system is presented in Sabater and Sierra [9] where three kinds of the reputation are considered: individual, social and ontological. The individual reputation takes two values -1 and 1 and is based on individual decision. The social reputation one inherits from a group it belongs to, while the ontological reputation is consolidated value of the two other reputations. Moreover, the calculated reputation value naturally decreases in time.

The authors in [10] proposed a system where a party can play several roles, each in certain proportion. The global reputation value of the party is a weighted aggregation of the reputations in each of the roles (quantified according to defined measures).

In [11] a site assigns the party one of three linguistic trust levels (-1, 0, 1) after each interaction and cumulates the experience of contacts. When the number of contacts is sufficiently big, the reputation is calculated according to own experience. Otherwise, the site uses reputation of a party obtained from other sites.

Song et al. [12] define a system where the site's reputation is formed based on parties' own aggregated experience (using four factors: prior success rate, cumulative site utilization, its job slowdown ratio and job turnaround time) and the site self-defence capability (taking into account four security factors). Except of calculating reputation, some models propose also mechanisms of cheating detection (see, e.g., [13]) that help reducing false decision-making.

2.3 SRs based on deterministic approach

One may find a set of reputation systems that incorporates a deterministic approach to realize a mathematical evaluation engine of reputation systems. These groups of systems are usually optimized for real applications and take the opportunity of heuristic reputation modelling. For example Google's PageRank [14] scores a Web page according to how many other pages are linked with the scored one. For such a hyperlinked network of pages the reputation of referring site has an influence to the scoring of pages that are referred to. System has a centralized nature and in order to avoid illegal positioning the additional mechanisms are used, for instance domain name costs or frequency of page updates.

Similar approach to PageRank may be found in Liu et al. [15], where the reputation system is proposed for Mobile Ad-hoc Networks (MANET). The scoring is built according to the own experience of nodes and the shared reputation of close neighbourhood.

Liu's model assumes that management of subjective opinions is realized in the decentralized environment. The system has ability to reflect a history of collected opinions and evolve with changing dynamics. The input parameter vector is composed of a weighted list of attributes, which are shaped respecting the importance of evaluated features. The opinions are mostly connected to the trust of network nodes but may be extend to parameters reflected in QoE metrics. Proposed extension of Liu's proposal may be found in [16], [17] where basic system was applied for an anonymous communication and real time communication system SecMon. These extensions point out that Liu's reputation system performance and sensibility stay in close relation with the amount of input data. This means that reputation provide less reliable output results, especially in a initial phases of building reputation or limited activity. To cope with this drawback in [16], [37] there was proposed a *virtual time quantum* keeping the reputation evolution on sustainable level.

Liu's reputation system has a native ability of scoring QoE-related metrics and reflecting a context dimension of application. It makes the reputation system a suitable and interesting candidate for the reputation building in Web services.

3 QoE Models Classification

Quality of Experience combines user perception, experience and expectations with non-technical and technical parameters such as application- and network-level QoS. There is, however, still a lack of quantitative descriptions or exact definitions of QoE. One particular difficulty consists in matching subjective quality perception to objective, measurable QoS parameters for various applications. Reputation may be an appropriate mean to overcome this and to obtain a QoE value without explicitly knowing a direct relationship between QoE and QoS parameters. In this section, we introduce a classification of existing QoE metrics and how to measure them.

There exist two basic measurement options, which are subjective testing and objective testing. Usually, subjective quality tests form the basis for perceptual objective test methods. The subjective tests are carried out by test panels of (real) users. While many (possibly even diverging) views on the quality of the outcome can be taken into account leading to accurate results as well as a good understanding of the QoE and its sensitivity, this type of test can be both time-consuming and costly, since the tests have to be conducted by a large number of users for statistically relevant results.

Objective tests are carried out by an algorithm on behalf of a real user, trying to imitate (or predict) user perception based on key properties of the reference and/or the product. Objective tests can follow psychophysical approaches and engineering approaches, a detailed description of which is found in [21]. For VoIP, the PESQ (Perceptual Evaluation of Speech Quality) standard [22] objectively evaluates and quantifies voice quality of voice-band 300 – 3400 Hz speech codecs. It uses psycho-acoustic and cognitive model to analyze and compare the reference and the outcome. PESQ allows for repeatable and automated measurement processes, which is necessary for obtaining statistical significant results.

Depending on the available information for subjective or objective tests, quality metrics can be classified according to the following three categories, cf. amongst others [21], [23], [24]:

- **Full Reference (FR) metrics**: Both outcome and reference are available and allow for detailed subjective and objective comparisons of voice, images, videos, download times on application level, as well as packet traces on network level, etc. Concretely, this means extraction, evaluation and comparison of QoE- and QoS-related parameters on any level in an off-line manner, which is most interesting for deriving QoE to QoS relationships. FR metrics deliver the highest accuracy, but require high computational effort.
- **No Reference (NR) metrics**: Quality information has to be extracted from the outcome, as no reference is available. This is a typical online situation with sole focus on the resulting quality as perceived by the end user, e.g. evaluated through questions, or the user's representative, e.g. an algorithm. In a networking context, NR metrics are usually lacking the possibility of discerning between quality problems stemming from the reference, e.g. quality degradations due to encoding, and additional disturbances by the network. Thus, NR metrics are not applicable for deriving QoE to QoS relationships aiming at capturing the impact of the network. NR metrics estimate the actual QoE with a low accuracy only. Common variants of NR algorithms even analyze only on network level.
- **Reduced Reference (RR) metrics**: Instead of comparing directly the reference with the outcome, parameters on application and/or network level are extracted at the sending and receiving side, which help predicting the QoE. As an example, on application level the RR Hybrid Image Quality Metric (HIQM) [23] computes various criterions of the reference image and sends them to the receiver. The extracted parameters are taken into account for estimating the quality of the received image without needing the reference image at the receiver. As a further example, on network level throughput variations and losses may be derived and compared to estimate the quality on receiver side as done in [25] and [26]. Such parameters often have their roots in FR research as a means of summarizing and interpreting the outcomes. However, as they represent key QoE and QoS parameters in a very condensed manner, they can be applied in an online in-service scenario by transmitting them between source and sink, and subsequently comparing them in order to find out about quality problems. Because of their background, they represent promising candidates to build QoE to QoS relationships upon [25], [26] and [27].

QoE metrics exist mainly for speech as well as video transmissions. The Mean Opinion Score (MOS) enables a subjective assessment of experienced speech quality, which is based on the subjective placement of voice samples by test persons on a scale from 1 (bad) to 5 (excellent) as defined in ITU-T P.800. In contrast, objective scoring mechanisms try to determine the experienced quality of speech based upon measurable values. One of these, the E-model (ITU-T G.107), maps the influence of different factors impeding the transmission of voice data onto the so-called R-factor, which is a measure of voice quality. Another, the PESQ value (Perceptual Evaluation of Speech Quality) [22], results from a comparison of two voice samples. It is typically used to evaluate transmission quality in a network using test samples.

Prior work on the topic of QoE, cf. [19] and [20], showed that VoIP is heavily impacted by network parameters such as jitter, packet delay and loss, whereas mainly effective throughput is determining the experienced quality for data services [18]. Here, an exponential interdependency was found between QoE and the according QoS parameter (e.g., packet loss) in the examined scenarios. This implicates that QoE is very sensitive to QoS disturbances in case the experienced service quality is high. Under negative conditions, i.e., a low QoE, further disturbances have a smaller effect. This sensitivity has to be taken into account by a reputation system to reflect properly the QoE.

Apart from these, numerous ways to assess the objective and subjective quality of video exist, such as the ITU-T J.144 standard for cable TV evaluation. Other mechanisms to judge video quality, multimedia content and IPTV are developed by the ITU study group 12 and especially the Video Quality Experts Group (VQEG). There are also a large number of publications on this topic, with selected examples being [28] or [29].

However, for other Internet services and applications like web service there are only a few studies available, which directly focus on the quantification of QoE. In that case, reputation could be a viable option.

4 Applications of Reputation Systems for QoE

The complexity of contemporary web services makes the quantitative QoE evaluation a multidimensional challenge for automated reasoning machines such as reputation systems. These dimensions one may identify as several user- and service oriented items, which contribute to QoE metrics. Such metrics are linked with the web services with a mix of multimedia content (audio, video, metadata), varying context (social web portals, news, science, advertisement, entertainment, e-commerce) and eventually the meaning (usefulness, importance) for the end user. In addition, the service logic and its design add a substantial input to the service ergonomics, which determines how efficient and comfortable web surfing is. Also the user's expectations and his cognition on the web service depend on the individual's profile (age, hobbies, attitude, etc.). Finally, the previous section presents that only a limited set of network measures such as QoS parameters are linked with QoE for web services. This picture of the complex QoE metric brings a reputation system as a viable solution to be applied for QoE assessment. Assuming that the given QoE metric may be split into several simple parameters the evaluation of QoE as a whole service related measure can be handled by a reputation system, which is a fully eligible candidate to collect multidimensional input data. Thus, the reputation system may perform reasoning on observations and yield results, which reflects the user's satisfaction as well as service or network performance and reliability.

The reputation of web services and their users are usually considered as long-term scoring system [31]. For QoE the monitoring and evaluation is usually performed on short time basis. This requirement is a challenge for a reputation system engine in order to at least collect input QoE data with an appropriate resolution, which quarantines an accuracy of the output reputation. Some proposal in handling the real-time event by dedicated reputation systems can be found in [16], [37].

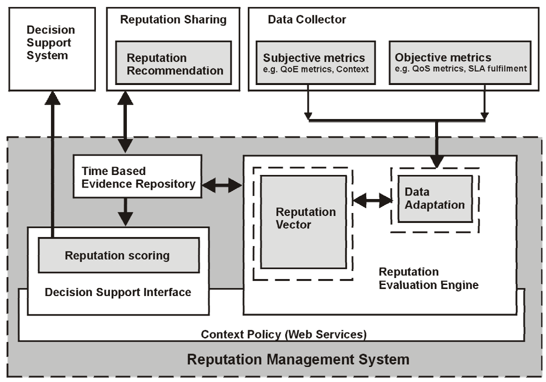

Fig. 2. Generalized reputation system model for context- and time-aware applications

In a variety of reputation system proposals [35] we may identify a common function set that takes part in a reputation forming and delivering knowledge for decision support systems. In order to perform QoE assessment, we propose the following framework architecture of a reputation system (Fig. 2). A *data collection* layer is responsible for feeding a *reputation evaluation engine* with measurements such as subjective scoring of context aware web transactions between parties as well as objective metrics related to network QoS parameters or Service Level Agreement fulfillment. The reputation evaluation engine *adopts* and *normalizes* input data in order to extract and emphasize the characteristic features of the opinions, which are under evaluation with the reputation system. *Reputation vector* is an internal metric, which reflects a history and context of the scoring and being stored in an *evidence repository*.

Reputation metrics are exchanged between interested parties with reputation sharing subsystem, especially employed in a distributed realization of reputation systems. The last common item is a decision support interface, which plays an important role in producing the output related to the reputation metrics and the service context for the given timeframe.

Within the three classes of reputation systems presented in Section 2, everyone has capabilities, which are suitable for QoE evaluation. They can handle multidimensional QoE nature in a distributed web service environment.

Probabilistic methods possess an innate mechanism for calculation of statistical correlations between data and metrics of QoE parameters. This feature is useful when a certain web portal delivers several web services with a significantly different context. The statistical calculation of interactions and collected data may result in a precise reputation generalization but limits input data sensibility in time.

Fuzzy based reputation systems are effective in an evaluation of scoring for social services considering time frames of interactions. One of the drawbacks of this solution might be a limited granularity of input parameters. The advantage of this system is an ability to perform an accurate ranking, even if the tampering and attacks on the system are possible. Moreover, the fuzzy systems, due to mechanisms of aggregation, enable taking into account several different properties or roles of a party being evaluated. Such systems could be useful in advanced models of reputation for web services where the three aspects of quality mentioned in Section 2 would be taken into account: the contents, presentation and transmission quality.

In deterministic approach the sensibility and time resolution may be adapted and parameterized for particular features of input data. This makes the reputation systems of this group a good candidate for QoE assessment of web services. However, the heuristic reputation modeling within these methods may lead to biased results and long-term outcomes could be misleading in reasoning. The possible overcome to this issue is a hybrid approach where probabilistic and deterministic systems are combined allowing for self cross-checking the reputation evaluation. How to design the hybrid reputation system applicable for QoE could be a subject for further research.

5 Conclusions

This paper concerns applicability of reputations systems for assessing QoE for web service in Future Internet. The presented framework is a generic architecture proposal for reputation systems, which provide mechanisms to manage subjective opinions in a web society and yield general scoring of particular users' behavior as well as service and network reliability. QoE parameters express the level of satisfaction of the users, which may vary greatly in time and depend on a service context or its type. This multidimensional nature of QoE metrics can be handled by reputation systems, which produces time and context related scoring on the users, service and network operator.

The application of the reputation systems for QoE assessment faces the challenges of adaptation QoE metric features into the data collection module with a need of definition how the input measurements are correlated with a user behavior and service context. This part is not clearly covered in literature and drives a new research areas related to the QoE user behavior modeling.

The usage of reputation may be a beneficial for service providers in terms of SLAs fulfillment or retaining QoE on the satisfaction level for users sharing the same network or service resources. In the scope of application advantages the reputation systems for QoE evaluations are able to support automated decision-makers and adapt web services or networks for keeping QoE on a satisfactory level. The benefits of such adoption are as follows:

- Aligned with business objectives (Tele Management Forum [38]).
- Reputation system may be treated as an input for Decision Support Systems (business intelligence systems).

- The outcome of the analysis may be used to trigger remedy actions for retaining QoE on the satisfying level in the network service for users sharing the same resources:

a) load balancing of network traffic driven by reputation,

b) limiting the number of concurrent web session for a user when QoE degradation is detected,

c) influence to the content adaptation mechanisms for real-time sessions (dynamic audio or video codecs changes) [29].

References

1. Amazon Auctions. http://auctions.amazon.com
2. eBay. http://www.ebay.com
3. Carrara, E., Hogben, G., (Eds.): ENISA Position Paper, Reputation-based Systems: a security analysis, European Network and Information Security Agency (ENISA), October 2007 URL:http://www.enisa.europa.eu/doc/pdf/deliverables/enisa_pp_reputation_based_system.pdf
4. Jøsang, A., Ismail, R.: The beta reputation system, Proceedings of the 15th Bled Electronic Commerce Conference, Bled, Slovenia, June, 2002.
5. Jøsang, A., Haller, J.: Dirichlet Reputation Systems, In The Proceedings of the International Conference on Availability, Reliability and Security (ARES 2007), Vienna, Austria, April 2007.
6. Jøsang, A., Bhuiyan, T., Xu, Y., Cox, C.: Combining Trust and Reputation Management for Web-Based Services, Proceedings of TrustBus2008. Turin, September 2008.
7. Jøsang, A., Hayward, R., Pope, S.: Trust Network Analysis with Subjective Logic, In Proceedings of the 29th Australasian Computer Science Conference (ACSC2006), CRPIT Volume 48, Hobart, Australia, January 2006.
8. Jøsang, A.: Logic for Uncertain Probabilities, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 9(3):279–311, June 2001.
9. Sabater, J., Sierra, C.: Reputation and social network analysis in multi-agent systems, Proc. 1st Int. Joint Conf. on Autonomous Agents and Multi-Agent System (AAMAS 2002), Bologna, Italy, 15-19 July 2002, pp. 475-482.
10. Carter, J., Bitting, E., Ghorbani, A.A.: Reputation Formalization for an Information-Sharing Multi-Agent System, Computational Intelligence, Vol. 18, No. 4, 2002, pp. 515-534.
11. Ramchurn, S.D., Sierra, C., Godo, L., Jennings, N.R.: A computational trust model for multi-agent interactions based on confidence and reputation, Proc. 6th Int. Workshop of Deception, Fraud and Trust in Agent Societies, Melbourne, Australia, 14-15 July 2003, pp. 69-75.
12. 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.
13. Shi, X.-b., Liu, F., Du, L., Chen, X.-h.: A cheating detection mechanism based on fuzzy reputation management of P2P MMOGs, 8th ACIS Int. Conf. on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, pp. 75-80.
14. 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.

15. Liu, J., Issarny, V.: Enhanced Reputation Mechanism for Mobile Ad Hoc Networks, Springer-Verlag Berlin Heidelberg, 2004
16. Ciszkowski, T., Kotulski, Z.: Distributed Reputation Management in Collaborative Environment of Anonymous MANETs, Proceedings of the IEEE International Conference on Computer as a Tool, EUROCON 2007, September 9-12, 2007, Warsaw, Poland, pp. 1028-1033. ISBN 1-4244-0813-X. Proceedings IEEE, IEEEXplore
17. Eliasson, Ch., Fiedler, M., Ciszkowski, T., Kotulski, Z., Mazurczyk, W.: Parameterisation of a reputation system for VoIP in P2P networks for improved communication quality and security, EuroFGI, IA.7.6 Workshop on Socio-Economic Aspects of Future Generation Internet, Karlskrona, Sweden, 27-29 May, 2008
18. Khirman, S., Henriksen, P.: Relationship between Quality-of-Service and Quality-of-Experience for Public Internet Service, In 3th Passive and Active Measurement Workshop (PAM2002), Fort Collins, Colorado, USA, March 2002.
19. Hoßfeld, T., Tran-Gia, P., Fiedler, M.: Quantification of Quality of Experience for Edge-Based Applications, 20th International Teletraffic Congress (ITC20), Ottawa, Canada, June 2007.
20. Hoßfeld, T., Hock, D., Tran-Gia, P., Tutschku, K., Fiedler, M.: Testing the IQX Hypothesis for Exponential Interdependency between QoS and QoE of Voice Codecs iLBC and G.711, 18th ITC Specialist Seminar on Quality of Experience, Karlskrona, Sweden, May 2008.
21. Engelke, U., Zepernick, H.-J.: Perceptual-based Quality Metrics for Image and Video Services: A Survey, In Proceedings of the 3rd Conference on Next Generation Internet Networks (NGI'07), Trondheim, Norway, 2007.
22. International Telecommunication Union, ITU-T Recommendation P.862, Perceptual Evaluation of Speech Quality (PESQ), an Objective Method for End-to-End Speech Quality Assessment of Narrowband Telephone Networks and Speech Codecs, 2001.
23. Kusuma, T. M., Zepernick, H.-J.: A Reduced-Reference Perceptual Quality Metric for in-Service Image Quality Assessment, in Proceedings of IEEE Symposium on Trends in Communications (SympoTIC'03), 2003.
24. Pastrana-Vidal, R. R., Gicquel, J.-C.: Automatic Quality Assessment of Video Fluidity Impairments Using a No-Reference Metric, in Proceedings of 4th International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM'06), 2006.
25. Fiedler, M., Tutschku, K., Carlsson, P., Nilsson, A. A.: Identification of Performance Degradation in IP Networks Using Throughput Statistics, in Proceedings of the 18th International Teletraffic Congress (ITC'18), Berlin, Germany, 2003.
26. Fiedler, M., Chevul, S., Radtke, O., Tutschku, K., Binzenhöfer, A.: The Network Utility Function: A Practicable Concept for Assessing Network Impact on Distributed Services, in Proceedings of the 19th International Teletraffic Congress (ITC'19), Beijing, China, 2005.
27. Hoßfeld, T., Binzenhöfer, A., Fiedler, M., Tutschku, K.: Measurement and Analysis of Skype VoIP Traffic in 3G UMTS Systems, in Proceedings of the 4th International Workshop on Internet Performance, Simulation, Monitoring and Measurement (IPS-MoMe'06), Salzburg, Austria, 2006.
28. Wang, Z., Lu, L., Bovik, A.C.: Video quality assessment based on structural distortion measurement, Signal Processing: Image Communication, 19 (2), 2004, pp. 121-132.
29. Ries, M., Nemethova, O., Rupp, M.: Motion Based Reference-Free Quality Estimation for H.264/AVC Video Streaming, 2nd International Symposium on Wireless Pervasive Computing, 2007.
30. Resnick, P., Zeckhauser, R., Friedman, E., Kuwabara, K.: Reputation systems, Communications of the ACM, 43(12):45–48, 2000.

31. Mui, L., Mohtashemi, M., Halberstadt, A.: A computational model of trust and reputation, In Proceedings of the 35th HICSS, 2002.
32. O'Reilly, T.: What Is Web 2.0, O'Reilly Network http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html
33. Buyya, Rajkumar; Chee Shin Yeo, Srikumar Venugopal, Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities, Department of Computer Science and Software Engineering, The University of Melbourne, Australia. pp. 9. http://www.gridbus.org/~raj/papers/hpcc2008_keynote_cloudcomputing.pdf, 2008
34. Liau, C.Y., et al.: Efficient distributed reputation scheme for peer-to-peer systems, Proceedings of the 2nd International Human.Society@Internet Conference (HSI), vol. LNCS 2713, Springer, 2003, pp. 54– 63.
35. Jøsang, A., Ismail, R., Boyd, C.: A survey of trust and reputation systems for online service provision, Volume 43, Issue 2, Pages 301-686 (March 2007) Emerging Issues in Collaborative Commerce
36. Mehta, B., Hofmann, T., Nejdl, W.: Robust collaborative filtering, ACM Conference On Recommender Systems Archive, Proceedings of the 2007 ACM conference on Recommender systems table of contents, USA, Pages: 49 - 56, 2007
37. Ciszkowski, T., Eliasson, Ch., Fiedler, M., Kotulski, Z., Lupu, R., Mazurczyk, W.: SecMon: End-to-End Quality and Security Monitoring System, Annales UMCS, Informatica, AI 8 (2008), pp.186-201.
38. Tele Management Forum URL: http://www.tmforum.org