Semantic Similarity Model for Risk Assessment in Forming Cloud Computing SLAs

Omar Hussain, Hai Dong, and Jaipal Singh

Digital Ecosystems and Business Intelligence Institute
Curtin University of Technology, GPO Box U1987
Perth, Australia
{O.Hussain,Hai.Dong,J.Singh}@cbs.curtin.edu.au

Abstract. Cloud computing has enabled users to access various resources and applications as a service and in return pay the provider only for the time for which they are used. Service Level Agreements (SLA) are formed between the user and provider to ensure that the required services and applications are delivered as expected. With the increase of public cloud providers, challenges such as availability, reliability, security, privacy and transactional risk demand detailed assessment during the formation of SLAs. This paper focuses on one subcategory of transactional risk while forming SLAs: namely, performance risk. We argue that performance risk assessment should be done by the user before entering into an SLA with a service provider. We propose to measure performance risk according to the specific context and assessment criteria with the aid of a semantic similarity model for the SLA requirement being negotiated in a cloud computing environment. We show through simulations that the performance risk analysis is more accurate using semantic similarity matching compared with analysis without semantic similarity matching.

Keywords: Performance Risk, Service Level Agreement, Cloud Computing, Context, Assessment Criteria, Semantic Similarity Model.

1 Introduction

Cloud Computing means different things to different people. To some, cloud computing is similar to thin-client Web-based applications, while others consider it as a computing utility that charges metered rates for every service. Some regard it as a means of efficiently processing scalability through highly distributed or parallel computing. However people look at cloud computing, it is an enabler for a new paradigm in computing.

In this paper, we define cloud computing as a model that commoditises resources, software and information as services, and delivers them in a manner similar to traditional utilities such as electricity and water. In such a model, users access services based on their requirements at a particular point in time regardless of where the services are hosted or how they are delivered.

Previous work in cloud computing focussed on technological frameworks for implementation and deployment of user services on the cloud, such as software as a

R. Meersman et al. (Eds.): OTM 2010, Part II, LNCS 6427, pp. 843–860, 2010.

service (SaaS), infrastructure as a service (IaaS), platform as a service (PaaS), etc. While these cloud computing resources are controlled and provided by the service provider, the cloud consumer needs to ensure that the quality, availability, reliability and performance of these resources meets their business functionality requirements. The consumers need to obtain guarantees from providers on service delivery to ensure that their business functions smoothly, and a means for recovery or compensation if these guarantees are not met. These are provided through Service Level Agreements (SLAs) negotiated between the providers and consumers.

An SLA is an extremely important document, as it (1) identifies and defines customer needs and expectations, (2) provides a mechanism to weight, verify, evaluate and enforce the agreed criteria, and (3) provides an appropriate level of granularity to trade-off between expressiveness and complexity. A typical SLA should contain a definition of required services, the methods for monitoring and measuring service performance, methods for compensation or indemnity for services provided, processes to manage unplanned incidents, customer duties and responsibilities to support the service delivery, security policies and procedures, disaster recovery, and termination of SLA.

Thus, a well-defined SLA will provide a framework for understanding, reduce areas of conflict, encourage dialogue in the event of disputes, and eliminate unrealistic expectations between provider and consumer [1]. As cloud computing provides different cloud offerings (IaaS, PaaS, and SaaS), there is a need to define different SLA meta-specifications. Some work has been done in defining cloud SLA models, performance criteria and measurement [2], and development of a standardised Web SLA language [3]. However, this work cannot be directly applied to every type of cloud service [4].

The cloud provider must monitor any changes in the cloud environment in order to make real-time evaluation and adjustment for SLA fulfilment. Fast and effective decision models and optimisation algorithms are needed for this. Providers may also need to reject resource requests when SLAs cannot be met. These operations need to be carried out in a nearly automatic fashion due to the promise of "self-service" in cloud computing [5]. Once a business association has been established, this real-time information is also used by consumers to measure the quality of cloud service provided and calculate the probability of an SLA violation occurring [6, 7].

While this is important, we believe it is far more beneficial for a consumer to assess and manage this risk before an SLA is formalized. This is achieved through transactional risk assessment before entering into an SLA with a service provider. This assists the consumer to make an informed decision when selecting the most appropriate service provider with which to form an SLA from a given set of possible cloud service providers.

Therefore, we propose a transactional risk framework to aid consumers in preselecting an IaaS cloud service provider. We will highlight the importance of risk in decision-making when forming an SLA and existing approaches for providing risk-based decision making in Section 2. In Section 3, we propose a new transactional risk assessment framework to model risk criteria and identify the similarity between different risk criteria faced by other agents. The risk assessment models for making a decision in selecting a cloud service provider are defined in Section 4. The results are shown in Section 5. Finally, we conclude the paper in Section 6.

2 Related Work

2.1 Assessing Transactional Risk

SLA negotiations are carried out between the consumer and service provider so the consumer can decide which provider can provide services that maximises the successful achievement of the consumer's desired outcomes and minimises any losses. To do this, the consumer will make a decision by analysing criteria that it considers to be important, such as reliability, availability, security, privacy, trust and risk in the business interaction.

These concepts assess and address all the factors which have the potential to affect the interaction negatively. The assessment criteria will be used by the consumers to form a business contract (SLA) with the service provider. The notion of risk will inform the consumer of the consequences of failure of its collaboration with the provider. An interaction between a provider and consumer is dynamic, and the risk is likewise dynamic. Therefore, the analysis of each of these SLA criteria concepts is important at different time periods during the collaboration for making an informed interaction-based decision.

Various approaches have been proposed in the literature that analyse each of these concepts in forming an SLA. However, these approaches consider the notion of risk as a subset of trust, security and privacy which can be mitigated by analysis of these concepts. In reality, this is not the case. Risk expresses the occurrence of those events that will lead to experiencing a negative outcome along with the level and magnitude of possible loss that an interacting buyer can experience. Both these representations play an important part in decision making and are not determined by the analysis of trust, security or privacy in the collaboration. Thus, any decision taken in a cloud service interaction cannot be considered as being fully informed without the analysis of risk.

ISO/IEC Guide 73 defines risk as the combination of the probability of an event and its consequences (whether positive or negative) [8]. The process of how risk is analysed is termed as risk analysis, which is a combination of various sub-steps like Risk Identification, Risk Assessment, Risk Evaluation and Risk Management. Risk Management is the process of treating risk in a methodical way to obtain benefits and sustainable values from each activity [9]. But for this process to occur, a risk assessment must be carried out. Risk assessment determines the probability of a risk event occurring along with its associated threats or consequences. There are three primary methods for assessing risk [6]: qualitative, for classifying risk without determining numerical values of all assets at risk and threat frequencies; quantitative, which calculates the magnitude and probability of risk occurring; and semi-quantitative (or hybrid), which is less numerically intensive than the quantitative method and classifies (prioritises) risks according to consequences and foreseen probabilities.

Risk will have different representations according to the area in which it is being determined. For example, if the risk being determined relates to the security aspects while forming the business association, then its analysis represents security risks. Previous work in risk assessment decision making generally considers the probability of an agent cheating [10] and the costs associated with an interaction [11]. Some other works consider risk to be a component of trust [12-15] but do not quantify the negative consequences in their model or consider it in decision making.

When risk is being determined during the decision-making stage of forming an SLA contract, its analysis represents the transactional risk. Measuring the loss or its impact by analysing the level and degree of transactional risk in the interaction is very important when making an informed interaction-based decision. The sub-categories of transactional risk to be assessed when forming an SLA are ‘performance risk’ and ‘financial risk’.

In this paper, our focus is on performance risk. Performance risk represents the probability to which the risk assessing agent (service consumer) will not achieve the expectations of its business interaction. This is mainly due to the incapability or non-cooperation of the risk assessed agent (service provider) in committing to the expectations of the business interaction as decided initially. These agents may be an individual user, small and medium enterprises (SME) or businesses that want to achieve certain aims or desired outcomes. An agent can also be a software or web service. In our previous work, we proposed an approach by which the risk assessing agent determines the performance risk of a risk assessed agent in a business interaction [16] as explained below.

In any business interaction, the level of failure is not just two extremes, High or Low, but different levels of possible failures. We used a Failure Scale to capture those varying levels, with six different severities of failures as shown in Table 1. Each FailureLevel (FL) value on the scale quantifies and represents a different magnitude or severity of failure in the interaction.

The association of a consumer with a service provider on which the SLA is being formed may be either limited to the current period of time or may extend to a point of time in the future. To determine the performance risk of a service provider, the consumer should determine its ability to commit to the expectations at that point in time. This is achieved by determining the FL of the service provider to commit to the expectations of the SLA at that point in time. If the time period extends to a point of time in future, then the service consumer has to predict the FL of the service provider in committing to the expectations at that future period of time.

To consider the time-specific nature of transactional risk while doing so, we adopt the methodology proposed by Chang et al. [17] and determine the time space of the interaction, then divide it into different non-overlapping, mutually exclusive time slots, and identify the time spot of the interaction. Time spot represents the point in time where the service consumer initiates its association with the service provider as illustrated in Figure 1. The time space is divided into two broad phases, namely:

Table 1. The Failure Scale

Semantics of Failure Level	Probability of Failure	FailureLevel
Total Failure	91-100 %	0
Extremely High	71-90 %	1
Largely High	51-70 %	2
High	26-50 %	3
Significantly Low	11-25 %	4
Extremely Low	0-10 %	5

a) pre-interaction start time phase, representing the period before the consumer starts its association with the provider, and b) post-interaction start time phase, representing the period after the initiation of the association. Our method enables the consumer to utilize the impression or capability of the service provider in the pre-interaction time period and utilize it to predict its FL in the post-interaction time period.

To consider the dynamic and variable property of time related with transactional risk assessment, the service consumer should ascertain the FL of a service provider in each pre-interaction start time slot [16]. Some important characteristics of transactional risk that need to be considered are its:

- Context specific nature, which represents the purpose for which the business association is being carried out. Performance risk cannot be quantified successfully without taking into consideration the context in which the interaction is being formed.
- Assessment criteria specific nature, which represents the specific outcomes which the risk assessing agent wants to achieve in its interaction. Based on the context, the assessing agent will measure only the desired assessment criteria instead of all possible criteria.

There are two ways by which the consumer determines the FL of a service provider in committing to the expectations:

- By utilizing its own past interaction history and/or
- By utilizing the recommendations from other users.

We proposed that the risk assessing agent gives first preference to its own past interaction history (if it is in the expectations of its future association) to determine the FL of the risk assessed agent in a pre-interaction time slot. If it does not have any past interaction history in the specific expectations of its business interaction, then it solicits recommendations from other users and assimilates it to determine its FL value in the pre-interaction start time slot. Once the FL of each assessment criterion in a time slot has been determined, they should be combined according to their significance to ascertain the combined FL of the risk assessed agent in that pre-interaction start time slot. The determined FL of each assessment criterion in a time slot will be a value in the range of 0-5. But scenarios may arise where for a given assessment criterion in a pre-interaction start time slot, the risk assessing agent may not have either its own past interaction history or obtains recommendations from other users. In such scenarios, due to the incomplete information present, we consider that the assessing agent will err on the side of caution (assuming the worst scenario) and considers an FL value of zero (0) for that assessment criterion in that time slot.

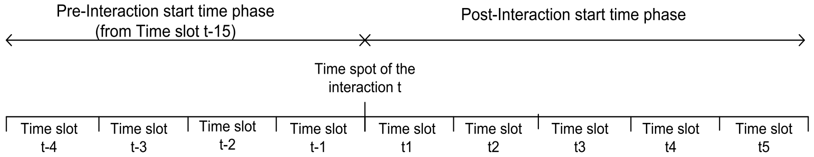

Fig. 1. Division of Time Space of the Interaction

2.2 Problem Definition

The absence of past interaction history for assessment criteria has led to very conservative outcomes when analysing performance risk. However, such outcomes might not be the best result as the risk assessed agent could still meet the assessed criteria even though the risk assessing agent has no prior knowledge of such. It is important to note that an FL value assigned for an assessment criterion will be propagated further as:

a) The FL of risk assessed agent in a pre-interaction start time slot is dependent on its FL for each assessment criterion in that time slot.
b) The performance risk of the risk assessed agent in the post-interaction start time phase is dependent on the FL values in the pre-interaction start time phase.
c) The sub-category of Financial Risk is dependent on the performance risk determined in the post-interaction start time phase, and
d) The level of transactional risk is dependent on the performance risk and financial risk determined in the post-interaction start time phase.

So it is important to make an informed decision about the FL of an assessment criterion in a pre-interaction start time slot. In this paper, we address this problem by proposing an approach where, in the absence of both direct past interaction history and recommendations from other agents, the FL of an assessment criterion can be determined by utilizing its level of similarity with the assessment criteria present. We achieve this by utilizing an ontology-based semantic similarity model.

2.3 Ontology-Based Semantic Similarity Models

Traditional semantic similarity models focus on measuring semantic similarity between nodes in semantic networks. Semantic networks refer to the graphic notations comprising arcs and nodes, in which nodes represent concepts and arcs represent relations between concepts [18]. Semantic networks can be used to represent simple knowledge in specific domains and a typical example is WordNet. However, limitations of semantic networks include: 1) nodes are usually single words and cannot be defined by properties; and 2) arcs are cannot be defined by restrictions and characteristics [19]. Compared with semantic networks, ontologies are a form of knowledge representation with more complex attributes. Ontologies consist of concepts and relations between concepts [20]. The advantages of ontologies include: 1) concepts can be defined by both datatype and object properties (relations); 2) object properties (relations) can be defined by multiple restrictions and characteristics. In terms of the comparison, it is not difficult to observe that ontologies can be employed to represent knowledge with more complex structures. Meanwhile, with the emergence of ontologies, new forms of semantic similarity models were developed in order to measure concept similarity in the ontology environment, known as ontology-based semantic similarity models, e.g., Dong et al.'s model [21].

In this paper, we propose an approach by which the risk assessing agent determines the level of similarity between the assessment criterion of its expectations and the other similar assessment criteria to accurately determine the FailureLevel (FL) of the risk assessed agent. The proposed approach is explained in the next sections.

3 Ontology-Based Risk Assessment Criteria Similarity Matching Framework

If an assessing agent (cloud service consumer) is unable to find a matching assessment criteria from the past cloud service interaction history of an assessed agent (cloud service provider), it will assume the worst case scenario and subsequently assign the worst FL weight of zero (0) for that criterion for any future interaction. This will influence the assessment in a negative manner, giving a high risk assessment due to uncertainty that might not accurately model the future interaction with an assessed agent.

In order to obtain more accurate values for the FL, we have extended our previous work by developing an ontology-based semantic similarity model that will measure the similarity of the current assessment criteria to the assessing agent or other cloud service consumers' previous interactions with the assessed agent. We propose to design a knowledge base which stores generic ontologies representing relationships between context-specific risk assessment criteria (figure 2). In terms of those generic ontologies, the similarity between the risk assessment criteria in the current interaction and the criteria in previous interactions can be calculated. Due to space constraints, the design of such ontologies is not discussed in this paper. This paper will explain the framework and its use in providing a failure level for a cloud service interaction between an assessing agent (cloud service consumer) and an assessed agent (cloud service provider).

In this framework, the assessing agent will store the assessment criteria of its previous interactions with the assessed agent in a database repository. The assessing

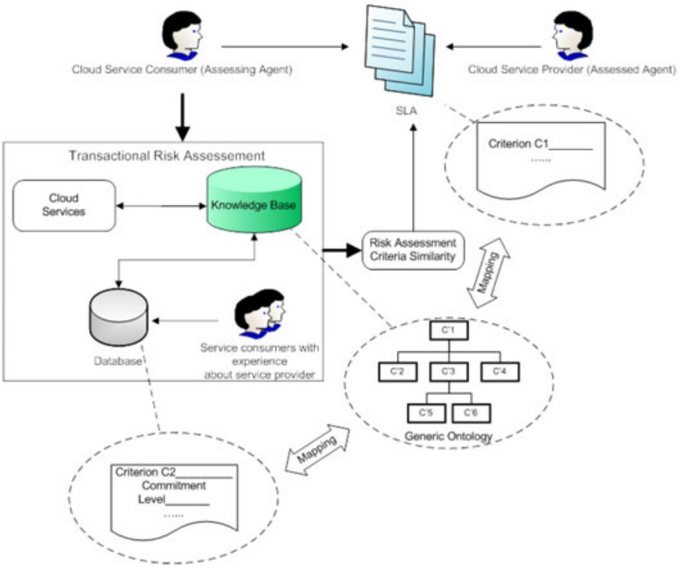
Cloud Service Consumer (Assessing Agent)
SLA
Cloud Service Provider (Assessed Agent)
Criterion C1
......
Transactional Risk Assessment
Cloud Services
Knowledge Base
Risk Assessment Criteria Similarity
Mapping
Database
Service consumers with experience about service provider
C'1
C'2
C'3
C'4
C'5
C'6
Generic Ontology
Mapping
Criterion C2
Commitment Level
......

Fig. 2. Cloud Service Transaction Risk Assessment Criteria Similarity Measure Framework

agent will further acquire the assessment criteria of previous interaction history from other service consumers that interacted with the assessed agent. There are no guarantees that these historical interactions would have used the same assessment criteria since each on-demand cloud service interaction is different for each service consumer and at different points of time. The other cloud service consumers might have used assessment criteria that are not recognised by the assessing agent's cloud service. Therefore, in our proposed framework, the assessing agent will use those generic ontologies from the knowledge base to match the assessment criteria in the current interaction with the criteria used in historical interactions with the assessed agent.

Assessing the risk for a cloud service interaction should be multi-dimensional (e.g. availability, price, latency, etc.). We maintain that each generic ontology should provide a shared representation of context-specific concepts in a cloud service risk assessment dimension. With the purpose of simplifying computation, we regard each generic ontology as a hierarchical structure of concepts linked by is-a relations. Therefore, an assessment criterion can be annotated by one or more concepts from a relevant generic ontology.

Here, we propose a semantic similarity model to measure the similarity between the current assessment criteria and the historical assessment criteria. This similarity model is designed based on the theory of Rada et al. [22]'s distance-based metric, which calculates the semantic distance between two concepts in terms of the shortest distance between the two concepts in a semantic network. In terms of the instances in Figure 2, our semantic similarity model for measuring two assessment criteria can be presented as follows:

Input: c1, c2 are two assessment criteria, O is an e-business ontology which consists of concepts (c'1…c'n) linked by *is-a* relations and its maximum depth is d.
Output: sim(c1, c2) – the similarity between c1 and c2.
Algorithm:

```
begin
for i = 1 to n
        if c'[i] ∈ c1 then
                Put c'[i] into an array C1;
        else if c'[i] ∈ c2 then
                Put c'[i] into an array C2;
        end if
end for
k = count (C1);
l = count (C2);
for i = 1 to k
        s = 2d;
        for j = 1 to l
                A[i][j] = the shortest distance between C1[i] and C2[j] in O;
                if A[i][j] < s then
                        s = A[i][j];
                end if
        end for
        d' = d' + s;
```

```
end for
for j = 1 to l
        t = 2d;
        for i = 1 to k
                if A[i][j] < t then
                        t = A[i][j];
                end if
        end for
        d’ = d’ + t;
end for
$sim(c1,c2) = 1 - \frac{d'}{(k+l)x2d'}$;
end
```

The scope of the similarity value is between the interval [0, 1], where 0 stands for nothing similar and 1 stands for completely similar. It needs to be noted that each ontology should represent an assessment dimension in a disjoint cloud service context. Therefore, this semantic similarity model cannot measure the similarity between two assessment criteria in different contexts, and we consider the similarity value should be 0 in that case. For example, the similarity value between the criterion of latency in a video service and that in an audio service should be 0, since there is no direct relationship between the two service contexts and a service provider’s performance in the audio service cannot affect his/her performance in the video service.

In order to clearly explain the proposed semantic similarity model, we provide a case study to describe the application of this model in the domain of cloud computing. With the purpose of revealing the feasibility of the proposed semantic similarity model in the cloud computing environment, we make use of a real-use scenario adopted from Amazon Web Services^{TM} (http://aws.amazon.com).

We premise that a consumer in Virginia wants to use an Amazon Elastic Compute Cloud (EC2) service (http://aws.amazon.com/ec2/). According to the actual demand of the consumer, s/he wants to use a small instance of the EC2 service (Windows). The consumer intends to obtain the risk performance of the small instance on the criterion of price. However, the consumer does not have any previous transaction with Amazon on the usage of the small instance. In contrast, s/he has the transaction history with Amazon on the usage of the large instance (Windows) and the usage of the high CPU medium instance (Windows). Here we define a price ontology in the context of the Amazon EC2 services (figure 3). Therefore, in terms of the price ontology and the proposed semantic similarity model, the similarity between the price for the small instance (S) and the price for the large instance (L) can be obtained by

$$sim(S,L)=0.667$$

Subsequently, the similarity between the price for the small instance (S) and the price for the high CPU medium instance (CM) can be calculated by

$$sim(S,CM)=0.333$$

Once our framework has measured the similarity between the current assessment criteria and the historical assessment criteria, the degree of similarities are used as weights to determine the FailureLevel (FL) of the assessment criteria.

4 Performance Risk Assessment Model

A series of further computations have to be carried out to determine the FL of the service provider in an assessment criterion. There might be different scenarios according to different factors and these are achieved as follows.

Step 1: Determine the Commitment Level of Agent 'B' in Assessment Criterion C_n.

Case 1: The risk assessing agent has previous interactions with the risk assessed agent in partly similar assessment criteria as compared to the current expectations of the SLA of its future interaction.

If the risk assessing agent 'A' has a past interaction history with the risk assessed agent 'B' in not exactly, but in partly similar assessment criteria (C_{ns}) as compared to the required assessment criteria (C_n) of the SLAs of its future interaction (termed as expectations), then we propose that the Commitment Level of agent 'B' in assessment criteria (C_1) is ascertained by:

(a) Determining the level of similarity of the assessment criteria (C_n) of the expectations and other similar assessment criteria, and
(b) Weighing the Commitment Level of agent 'B' in assessment criterion C_{ns} with the level of similarity between (a) assessment criteria and (b) the weight 'w' applied to the commitment level of the risk assessed agent to adjust and consider its status in the time slot 't-z'.

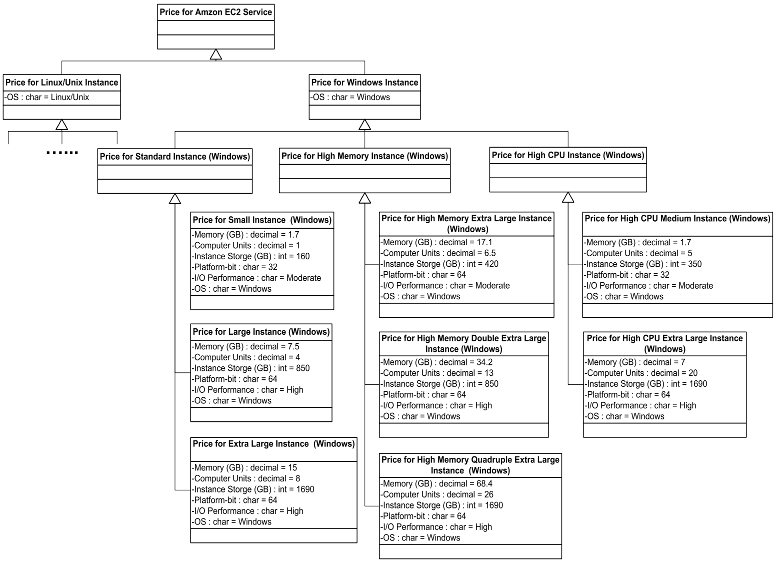
Price for Amzon EC2 Service
Price for Linux/Unix Instance
-OS : char = Linux/Unix
Price for Windows Instance
-OS : char = Windows
......
Price for Standard Instance (Windows)
Price for High Memory Instance (Windows)
Price for High CPU Instance (Windows)
Price for Small Instance (Windows)
-Memory (GB) : decimal = 1.7
-Computer Units : decimal = 1
-Instance Storge (GB) : int = 160
-Platform-bit : char = 32
-I/O Performance : char = Moderate
-OS : char = Windows
Price for Large Instance (Windows)
-Memory (GB) : decimal = 7.5
-Computer Units : decimal = 4
-Instance Storge (GB) : int = 850
-Platform-bit : char = 64
-I/O Performance : char = High
-OS : char = Windows
Price for Extra Large Instance (Windows)
-Memory (GB) : decimal = 15
-Computer Units : decimal = 8
-Instance Storge (GB) : int = 1690
-Platform-bit : char = 64
-I/O Performance : char = High
-OS : char = Windows
Price for High Memory Extra Large Instance (Windows)
-Memory (GB) : decimal = 17.1
-Computer Units : decimal = 6.5
-Instance Storge (GB) : int = 420
-Platform-bit : char = 64
-I/O Performance : char = Moderate
-OS : char = Windows
Price for High Memory Double Extra Large Instance (Windows)
-Memory (GB) : decimal = 34.2
-Computer Units : decimal = 13
-Instance Storge (GB) : int = 850
-Platform-bit : char = 64
-I/O Performance : char = High
-OS : char = Windows
Price for High Memory Quadruple Extra Large Instance (Windows)
-Memory (GB) : decimal = 68.4
-Computer Units : decimal = 26
-Instance Storge (GB) : int = 1690
-Platform-bit : char = 64
-I/O Performance : char = High
-OS : char = Windows
Price for High CPU Medium Instance (Windows)
-Memory (GB) : decimal = 1.7
-Computer Units : decimal = 5
-Instance Storge (GB) : int = 350
-Platform-bit : char = 32
-I/O Performance : char = Moderate
-OS : char = Windows
Price for High CPU Extra Large Instance (Windows)
-Memory (GB) : decimal = 7
-Computer Units : decimal = 20
-Instance Storge (GB) : int = 1690
-Platform-bit : char = 64
-I/O Performance : char = High
-OS : char = Windows

Fig. 3. Abbreviated view of a price ontology in the context of Amazon EC2 services

If there is more than one assessment criteria in the risk assessing agent's past interaction history with the risk assessed agent which partly matches the expectations of its future interaction with it, then agent 'A' should capture the similarity between each of them and accordingly weigh it to determine the commitment level of the risk assessed agent in criterion C_n. The mathematical representation of the commitment level of agent ‘B’ in assessment criteria C_n by utilizing its similarity to the other contexts is:

$$\text{CommLevel}_{\text{BCn t-z}} = (w * \frac{1}{m} (\sum_{i=1}^{m} (\text{Sim}_{\text{Cn->Cnsi}} (\text{CommLevel}^{i}_{\text{BCns}})))) \tag{1}$$

where:

‘B’ represents the risk assessed agent, ‘C’ represents the context of the SLA; ‘C_n’ represents the assessment criterion, in which the commitment level of the risk assessed agent ‘B’ is being determined; ‘C_s’ represents a context that is partly similar to context ‘C’ in which the SLA is being formed; ‘C_{ns}’ represents the assessment criterion which is partly similar to assessment criteria ‘C_n’ of the SLA; ‘CommLevel $_{\text{Cns}}$’ represents the level of commitment of the risk assessed agent for assessment criterion ‘C_{ns}’; ‘m’ represents the number of similar context and assessment criteria to context ‘C’ and assessment criteria ‘C_n’; Sim $_{\text{Cn->Cnsi}}$ represents the level of similarity between assessment criteria C_n and C_{nsi}, ‘*w*’ is the weight applied to the commitment level of the risk assessed agent to consider its status in the time slot ‘t-z’.

The variable ‘w’ is used to consider the recency of the time slot ‘t-z’ in question with respect to the time spot of the current association. It is important to take into consideration the dynamic nature of transactional risk during its assessment. This can be explained by an interaction scenario of agent ‘A’ forming an association with agent ‘B’ for one year from 25/07/2010. To elucidate, let us consider that agent ‘A’ had a previous association with agent ‘B’ in partly similar assessment criteria that completed on 15/07/2009. Assuming that:

a) agent ‘A’ does not have any past interaction experience with agent ‘B’ in the current context and assessment criteria of its business association;
b) its past interaction history matches partly with one of the assessment criteria (C_n) of its current interaction; and
c) the time period of its previous association is within the pre-interaction start time slot (PFL) of its current interaction in which it does not have any past interaction history

then agent ‘A’ can utilize its past interaction history to determine the commitment level of agent ‘B’ in C_n. But due to the dynamic nature of risk, agent ‘A’ cannot consider the impression of agent ‘B’ that it had in that previous period of time as it is quite possible that its capability to act according to the expectations may have changed during that time. So in order to consider the dynamic nature of transactional risk, it is important for agent ‘A’ to accordingly adjust the commitment level of agent ‘B’ according to the time delay factor (*w*) to consider its fresh status. We determine the weight (*w*) to be given to each time slot of the pre-interaction start time phase by:

$$w = \begin{cases} 1 & \text{if } m \leq \Delta t \\ e^{\frac{-(\Delta t - m)}{N}} & \text{if } m > \Delta t \end{cases} \tag{2}$$

where:
'*w*' is the weight or the time delaying factor to be given to the status of the risk assessed agent; 'm' represents the time slot for which the weight of adjustment is determined; 'Δt' represents the number of time slots from the time spot in which the risk assessing agent will give more importance to the freshness of the status of the risk assessed agent; 'N' is the term which characterizes the rate of decay.

The adjustment factor '*w*' adjusts the commitment level values of the risk assessed agent in the recent time slots from the time spot of the current interaction more heavily as compared to those in the far recent time slots, progressively adjusting the effect of the older values in order to take into consideration its fresh status. We consider that the risk assessing agent among the 15 time slots of the pre-interaction start time phase, does not weigh the commitment level values in the five time slots previous to the time spot of its interaction (time slot t-1 till t-5, that is characterized by Δt in Eq 2) as they are near to the time spot of its future interaction. For the importance to be given to the commitment level of the risk assessed agent in the other time slots of the pre-interaction start time phase (from t-6 till t-n), the weight according to which they have to be adjusted is a progressively declining value determined by using eq. 2.

Another advantage of adjusting the values according to their time weight avoids modelling the behaviour of agent 'B' in the future that may no longer be relevant according to the expectations of its future interaction. This is particularly important while ascertaining the FailureLevel of agent 'B' at a future period of time by utilizing its impression in the pre-interaction start time slots.

In cases when the risk assessing agent does not have its own past-interaction history in the part assessment criteria of its current business association, then it can utilize the recommendation-based method to ascertain the level of commitment of agent 'B' in the assessment criteria of its SLA. We explain the process of achieving this in the next sub-section.

Case 2: The risk assessing agent receives recommendations from other agents that are in partly similar assessment criteria as compared to the current expectations of the SLA of its future interaction.

When agent 'A' receives recommendations from other agents about agent 'B' that are in partly similar assessment criteria as compared to the current expectations of the SLA of its future interaction, then we propose that the Commitment Level of agent 'B' that assessment criteria is determined by:

a) *Classifying the recommendations according to their credibility.*

There are two broad types of groups in such classification. They are *Known* and *Unknown* recommendations. Known recommendations are the feedback from agents with whom the risk assessing agent has previous experience in soliciting and considering recommendations. Unknown recommendations are the feedback from those agents with whom the risk assessing agent does not have previous experience in considering recommendations. The known recommendations are further classified into two types, which are either *Trustworthy* or *Untrustworthy* recommendations. Trustworthy recommendations are those which the risk assessing agent considers to be correct opinions. On the other hand, untrustworthy recommendations are those which the risk assessing agent does not believe to be totally correct. We consider that that the risk assessing agent considers only recommendations that are either trustworthy or unknown when it aggregates them to determine the commitment level of the risk assessed agent. It omits taking into consideration the untrustworthy recommendations as they do not provide with the correct representation of the risk assessed agent. Further details on how the risk assessing agent considers the recommendations according to their trustworthiness are explained later.

b) Combining the recommendations to determine the Commitment Level:

From the trustworthy and unknown recommendations in the particular assessment criterion of its interest, consider the ‘Commitment Level’ value and adjust it according to the:

- level of similarity between the assessment criteria (C_n -> C_{ns}),
- credibility of the recommendation,
- time decay weight factor to be given according to the status of the risk assessed agent in that time slot.

Represented mathematically, the commitment level of agent ‘B’ in assessment criteria C_n by utilizing its similarity from the recommendations of other users is determined by:

$$\text{CommLevel}_{\text{BCn t-z}} =$$

$$(\alpha \ {*}(w * \frac{1}{K}(\sum_{i=1}^{K} \text{Sim}_{\text{Cn->Cnsi}} (RCV_{\text{i}} \oplus \text{CommLevel}_{\text{BCnsi}})))) +$$

$$(\beta * (w * \frac{1}{J} (\sum_{o=1}^{J} \text{Sim}_{\text{Cn->Cnso}} (\text{CommLevel}_{\text{BCnso}})))) \tag{3}$$

where:

‘B’ represents the risk assessed agent, ‘C_n’ represents the assessment criterion, in which the commitment level of the risk assessed agent ‘B’ is being determined, ‘C_{ns}’ represents the assessment criterion which is partly similar to assessment criteria ‘C_n’ of the SLA, ‘RCV_{i}’ is the credibility value of the trustworthy recommending agent ‘i’, ‘K’ is the number of trustworthy recommendations that the risk assessing agent obtains for the risk assessed agent in similar assessment criterion to ‘C_{ns}’ in time slot ‘t-z’, ‘J’ is the number of unknown recommendations that the risk assessing agent gets for the risk assessed agent in similar assessment criterion to ‘C_{ns}’ in time slot ‘t-z’, ‘ α and β ’ are the variables attached to the parts of the equation which will give more weight to the recommendation from the trustworthy known recommending agents as compared to those from the unknown recommending agents. In general $\alpha > \beta$ and $\alpha + \beta = 1$, ‘w’ is the weight applied to consider the status of the risk assessed agent in time slot ‘t-z’.

As shown in equation 3, the commitment level value of agent ‘B’ for an assessment criterion ‘C_n’ is determined in two parts. The first part of the equation calculates the commitment level value of agent ‘B’ for the assessment criterion ‘C_n’ by taking the recommendations of the trustworthy known recommending agents whereas the second part of the equation calculates the commitment level value of agent ‘B’ in the same assessment criterion ‘C_n’ by taking the recommendations of the unknown recommending agents. The recommendations from the untrustworthy known recommending agents are omitted and not considered. In order to give more importance to the recommendations from the trustworthy known recommending agents as compared to ones from the unknown recommending agents, variables are attached to the two parts of the equation. These variables are represented by α and β respectively. It depends upon the risk assessing agent how much weight it wants to assign to each type of recommendation. Furthermore, as explained in the previous sub-section, each recommendation for the risk assessed agent in a time slot is adjusted according to the weight to be given to the status of the risk assessed agent in that time slot.

The RCV of the trustworthy known recommending agent is also considered with the adjustment operator '$\oplus$' while assimilating its recommendation. This takes into consideration the accurate recommendation from the trustworthy recommending agent according to the credibility and accuracy by which it communicates its recommendations. The rules for the adjustment operator '$\oplus$' are:

$$a \oplus b = \begin{array}{ll} a + b, & \text{if } 0 \leq (a + b) \leq 1 \\ 1, & \text{if } (a + b) > 1 \\ 0, & \text{if } (a + b) < 0 \end{array}$$

Step 2: Determine the FailureLevel of agent 'B' in Assessment Criterion C_n.

Once the commitment level of a risk assessed agent for an assessment criterion has been determined then it should be mapped on the Failure Scale to determine its FailureLevel value (PFL) to complete that SLA in that time slot. The commitment level of agent 'B' for an assessment criterion shows its level of capability to meet the particular criterion according to the expectations. To determine the FailureLevel of agent 'B' for that criterion, the extent of its inability to complete the given assessment criterion has to be determined. To achieve this, the risk assessing agent should:

(a) *Map the commitment level of that assessment criterion on the Failure Scale (FS).*

Doing so, agent 'A' determines the capability of agent 'B' to meet that assessment criterion on the Failure Scale. As mentioned earlier, the levels on the Failure Scale between 0 and 5 represent varying degrees and magnitudes of failure. Hence, for ascertaining the FailureLevel of the risk assessed agent in an assessment criterion, its commitment level for that criterion should be mapped on the range of (0, 5) on the Failure Scale, as it is within these levels that its capability to complete the assessment criterion has to be ascertained on the Failure Scale. The trustworthiness or the reputation of the risk assessed agent in an assessment criterion can be represented on the Failure Scale (FS) by:

$$\text{CommLevel}_{\text{BCn t-z FS}} = \text{ROUND}\ (\text{CommLevel}_{\text{BCn t-z}} * 5) \tag{4}$$

where:

'CommLevel $_{\text{BCn t-z FS}}$' represents the commitment level of agent 'B' in time slot 't-z' and in assessment criterion 'C_n' on the Failure Scale; 'CommLevel $_{\text{BCn t-z}}$' represents the commitment level of agent 'B' in assessment criterion 'C_n' and in time slot 't-z'.

(b) Determine the probability of failure of agent ‘B’ in committing to that assessment criterion according to its expectations.

By ascertaining the difference between what agent ‘A’ expects in an assessment criterion and how far agent ‘B’ can fulfil it according to its commitment level for that criterion, agent ‘A’ should determine the probability of failure to achieve that assessment criterion in that time slot. The FailureLevel of the assessment criterion in that time slot is then achieved by mapping the probability of failure of that assessment criterion to the Failure Scale (which is between 0 and 5).

Agent ‘A’ expects agent ‘B’ to complete the assessment criterion according to its expectations. This expectation of agent ‘A’ can be quantified with a value of 5 on the Failure Scale, as it represents the lowest probability of failure of the assessment criterion and expresses the maximum commitment by agent ‘B’ to its expectations. The probability of failure to achieve an assessment criterion ‘C_n’ according to the expectations in interacting with the risk assessed agent ‘B’ in a time slot ‘t-z’, according to its trustworthiness or reputation in this can be determined by:

$$\text{Probability of Failure}_{\text{BCn t-z}} = \left(\frac{5 - \text{CommLevel}_{\text{BCn t - z FS}}}{5}\right) * 100 \tag{5}$$

The determined probability of failure to achieve assessment criterion ‘Cn’ according to the expectations, in interacting with the risk assessed agent ‘B’ and in time slot ‘t-z’ will be on a scale of 0-100 %. The risk assessing agent from this can determine the FailureLevel (PFL) of the risk assessed agent ‘B’ in assessment criterion ‘C_n’ and in time slot ‘t-z’ on the Failure Scale ($\text{PFL}_{\text{BCn t-z}}$) by:

$$\text{PFL}_{\text{BCn t-z}} = \text{LEVEL (Probability of Failure}_{\text{BCn t-z}}) \tag{6}$$

Once agent ‘A’ determines the FailureLevel of each assessment criteria of its expectations, either by utilizing its own past interaction history or recommendations in those assessment criteria (proposed in) or in the absence of those by utilizing the similarity of other assessment criteria in its own past interaction history or recommendations (proposed in this paper) then the next step is to combine then to ascertain the FailureLevel of the risk assessed agent ‘B’ in a pre-interaction start time slot ‘t-z’ ($\text{PFL}_{\text{Pt-z}}$). This is shown in the next step.

Step 3: Determine the FailureLevel of agent ‘B’ in time slot ‘t-z’.

The FailureLevel of agent ‘B’ in time slot ‘t-z’ is determined by weighing its FailureLevel to complete each assessment criterion of the expectations in that time slot, with the significance of the assessment criteria as shown in Equation 7.

$$\text{PFL}_{\text{Bt-z}} = \text{ROUND}\left(\sum_{n=1}^{y} S_{\text{Cn}} * \text{PFL}_{\text{BCn t-z}}\right) \tag{7}$$

where:

‘S_{Cn}’ is the significance of the assessment criterion ‘C_n’; ‘$\text{PFL}_{\text{PCn t-z}}$’ represents the FailureLevel of the risk assessed agent ‘P’ in assessment criterion ‘C_n’ in time slot ‘t-z’; and ‘y’ is the number of assessment criteria in the expectations.

5 Discussion

We simulated the case study discussed in Section 3 in order to determine the performance risk in the pre-interaction timeslot using our original model [16] and our proposed semantic similarity matching framework. We considered that when the consumer is forming the SLA with Amazon for using the small instance of EC2 service (Windows); (a) there are 5 assessment criteria, and (b) the time space is formed such that there are 15 timeslots (ts) in the pre-interaction time phase and 10 timeslots in the post-interaction time phase. We determine the performance risk (PFL) of Amazon in committing to those criteria (C1 to C5) in timeslots t-15 to t-1. Due to space limitations, we show only the determined PFL value of Amazon in timeslots t-1 to t-5 in table 2. The shaded rows of each timeslot show the PFL using the proposed method while the unshaded rows use the original model. As can be seen in t-4, the PFL determined using the original method is 0 on the failure scale as the consumer did not have any past interaction history or receive any recommendations about Amazon in C1-C5. But by using semantic similarity matching, we used those assessment criteria that are similar to C1-C5 for determining the PFL as 2 in the failure scale. Figure 4 shows the improvement in the PFL value of Amazon in timeslots t-15 to t-1.

Table 2. Calculation of PFL in Timeslots t-5 to t-1

TS	AssCrit C1	AssCrit C2	AssCrit C3	AssCrit C4	AssCrit C5	PFL
t-5	CommLevel: 1 Source: OWN	CommLevel: 0 Source: OWN	CommLevel: 1 Source: OWN	CommLevel: 1 Source: OWN	CommLevel: 0 Source: OWN	3
	-	-	-	-	-	3
t-4	CommLevel: 0 Source: NONE	CommLevel: 0 Source: NONE	CommLevel: 0 Source: NONE	CommLevel: 0 Source: NONE	CommLevel: 0 Source: NONE	0
	CommLevel: 0 Sim $_{\text{Cn->Cnsi}}$: 0.9 Source: OWN	CommLevel: 1 Sim $_{\text{Cn->Cnsi}}$: 0.4 Source: OWN	CommLevel: 1 Sim $_{\text{Cn->Cnsi}}$: 0.8 Source: OWN	CommLevel: 1 Sim $_{\text{Cn->Cnsi}}$: 0.6 Source: OWN	CommLevel: 1 Sim$_{\text{Cn->Cnsi}}$: 0.75 Source: OWN	2
t-3	CommLevel: 0 Source: NONE	CommLevel: 0 Source: NONE	CommLevel: 0 Source: NONE	CommLevel: 0 Source: NONE	CommLevel: 0 Source: NONE	0
	CommLevel: 1 Sim $_{\text{Cn->Cnsi}}$: 0.8 Source: OWN	CommLevel: 1 Sim $_{\text{Cn->Cnsi}}$: 0.4 Source: REC-K RCV: 1	CommLevel: 1 Sim $_{\text{Cn->Cnsi}}$: 0.5 Source: REC-U	CommLevel: 1 Sim $_{\text{Cn->Cnsi}}$: 0.3 Source: REC-U	CommLevel: 1 Sim $_{\text{Cn->Cnsi}}$: 0.6 Source: OWN	2
t-2	CommLevel: 1 Source: OWN	CommLevel: 1 Source: REC-K RCV: 0.87	CommLevel: 1 Source: REC-U	CommLevel: 1 Source: REC-U	CommLevel: 1 Source: REC-K RCV: 0.74	4
	-	-	-	-	-	4
t-1	CommLevel: 0 Source: OWN	CommLevel: 0 Source: NONE	CommLevel: 0 Source: NONE	CommLevel: 0 Source: NONE	CommLevel: 0 Source: NONE	0
	-	CommLevel: 1 Sim $_{\text{Cn->Cnsi}}$: 0.4 Source: OWN	CommLevel: 1 Sim $_{\text{Cn->Cnsi}}$: 0.4 Source: REC-K RCV: 0.6 CommLevel: 1 Sim $_{\text{Cn->Cnsi}}$: 0.4 Source: REC-U	CommLevel: 1 Sim $_{\text{Cn->Cnsi}}$: 0.9 Source: OWN	CommLevel: 1 Sim $_{\text{Cn->Cnsi}}$: 0.8 Source: OWN	3

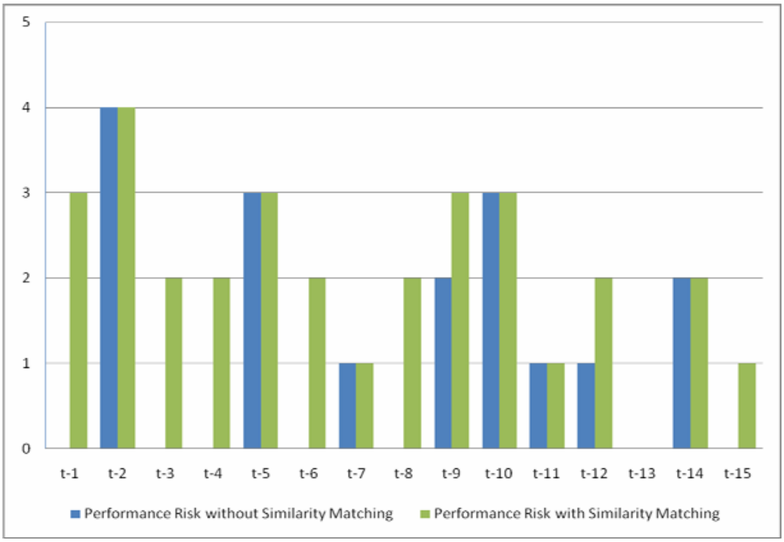

Fig. 4. Comparison of PFL using semantic similarity matching and original model over timeslots t-15 to t-1

6 Conclusion

In this paper, we proposed an improved approach for performance risk assessment that is used by a consumer to choose a cloud service provider that can meet its SLA requirements in the cloud environment. In order to determine risk, the consumer will base its decisions on the past capability of the provider. However, if there is no information that matches the capability of the provider according to the current assessment criteria, it will assume that the risk is very high for those criteria. Our proposed approach addresses this by utilizing a semantic similarity model that incorporates similar criteria from the provider's past interaction history into its performance risk assessment of the current interaction. This will help the consumer to make more informed decisions about (a) the performance risk of the provider in the post-interaction time phase, and (b) the financial risk and transactional risk in forming a SLA with the provider. We have shown that the use of semantic similarity matching improves the performance risk analysis. As part of our future work, we will extend our framework to assess the performance risk of a provider in providing multiple services to a consumer over the cloud environment, such as IaaS, SaaS and PaaS.

References

1. Kandukuri, B.R., Paturi, V.R., Rakshit, A.: Cloud Security Issues. In: Proceedings of the 2009 IEEE International Conference on Services Computing, pp. 517–520. IEEE Computer Society, Bangalore (2009)
2. Comellas, J.O.F., Presa, I.G., Fernández, J.G.: SLA-driven Elastic Cloud Hosting Provider. In: Proceedings of the 18th Euromicro Conference on Parallel, Distributed and Network-based Processing, pp. 111–118. IEEE Computer Society, Pisa (2010)

3. Nurmela, T., Kutvonen, L.: Service Level Agreement Management in Federated Virtual Organizations. In: Indulska, J., Raymond, K. (eds.) DAIS 2007. LNCS, vol. 4531, pp. 62–75. Springer, Heidelberg (2007)
4. Pearson, S., Charlesworth, A.: Accountability as a Way Forward for Privacy Protection in the Cloud. In: Cloud Computing, pp. 131–144. Springer, Heidelberg (2009)
5. Dillon, T., Wu, C., Chang, E.: Cloud Computing: Issues and Challenges. In: Proceedings on the 24th IEEE International Conference on Advanced Information Networking and Applications, pp. 27–33. IEEE Computer Society, Perth (2010)
6. Fitó, J.O., Guitart, J.: Introducing Risk Management into Cloud Computing. Barcelona Supercomputing Center and Technical University of Catalonia, Barcelona, Spain (2010)
7. AssessGrid Consortium.: D4.1 Advanced Risk Assessment. In: Carlsson, C., Weissmann, O. (eds.): Assess Grid Deliverable (2008)
8. ISO Guide 73: Risk Management Vocabulary (2009), http://www.iso.org/iso/cataloguedetail?csnumber=44651
9. ISO 31000: Risk management - Principles and guidelines (2009), http://www.iso.org/iso/cataloguedetail?csnumber=43170
10. Aberer, K., Despotovic, Z.: Managing trust in a Peer-2-Peer Information System. In: ACM (ed.): Proceedings of the Tenth International Conference on Information and Knowledge Management (CIKM 2001), Atlanta, Georgia, USA, pp. 310–317 (2001)
11. Zheng, X., Wu, Z., Chen, H., Mao, Y.: A Scalable Probabilistic Approach to Trust Evaluation. In: Stølen, K., Winsborough, W.H., Martinelli, F., Massacci, F. (eds.) iTrust 2006. LNCS, vol. 3986, pp. 423–438. Springer, Heidelberg (2006)
12. Jøsang, A., Keser, C., Dimitrakos, T.: Can We Manage Trust? In: Herrmann, P., Issarny, V., Shiu, S.C.K. (eds.) iTrust 2005. LNCS, vol. 3477, pp. 93–107. Springer, Heidelberg (2005)
13. Hassell, L.: Affect and Trust. In: Herrmann, P., Issarny, V., Shiu, S.C.K. (eds.) iTrust 2005. LNCS, vol. 3477, pp. 131–145. Springer, Heidelberg (2005)
14. Pearson, S., Mont, M.C., Crane, S.: Persistent and Dynamic Trust: Analysis and the Related Impact of Trusted Platforms. In: Herrmann, P., Issarny, V., Shiu, S.C.K. (eds.) iTrust 2005. LNCS, vol. 3477, pp. 355–363. Springer, Heidelberg (2005)
15. Wang, Y., Wong, D.S., Lin, K.-J., Varadharajan, V.: Evaluating transaction trust and risk levels in peer-to-peer e-commerce environments. Information Systems and E-Business Management 6, 25–48 (2008)
16. Hussain, O.K., Chang, E., Hussain, F.K., Dillon, T.S.: A methodology to quantify failure for risk-based decision support system in digital business ecosystems. Data & Knowledge Engineering 63, 597–621 (2007)
17. Chang, E., Dillon, T., Hussain, F.K.: Trust and Reputation for Service-Oriented Environments. John Wiley & Sons, Ltd., West Sussex (2006)
18. Sowa, J.F.: Semantic Networks. In: Shapiro, S.C. (ed.) Encyclopaedia of Artificial Intelligence. Wiley, Chichester (1992)
19. Dong, H., Hussain, F.K., Chang, E.: A hybrid concept similarity measure model for ontology environment. In: Meersman, R., Herrero, P., Dillon, T. (eds.) OTM 2009, pp. 848–857. Springer, Vilamoura (2009)
20. Gruber, T.: A translation approach to portable ontology specifications. Knowledge Acquisition 5, 199–220 (1995)
21. Dong, H., Hussain, F.K., Chang, E.: A context-aware semantic similarity model for ontology environments. Concurrency and Computation: Practice and Experience (in Press)
22. Rada, R., Mili, H., Bicknell, E., Blettner, M.: Development and application of a metric on Semantic Nets. IEEE Transactions on Systems, Man and Cybernetics 19, 17–30 (1989)