A Trust Vector Approach to Transaction Context-Aware Trust Evaluation in E-commerce and E-service Environments

Haibin Zhang, Yan Wang
Department of Computing
Macquarie University
NSW 2109, Australia
{*haibin.zhang, yan.wang*}*@mq.edu.au*

Xiuzhen Zhang
School of Computer Science and IT
RMIT University
Melbourne, Victoria 3001, Australia
xiuzhen.zhang@rmit.edu.au

***Abstract*—At some e-commerce websites (such as eBay), a trust value of a seller is computed based on the ratings of past transactions given by buyers, which can only reflect the general or global trust level of a seller without any transaction context information taken into account. As a result, a buyer may be easily deceived by a malicious seller in a forthcoming transaction. For example, with the notorious *value imbalance problem*, a malicious seller can build up a high trust level by selling cheap products and then start to deceive buyers in selling expensive products. In this paper, we first model all contextual transaction factors that reflect the nature of transactions, and thus influence the evaluation of transaction trust. In addition, instead of providing a single trust value, we propose a trust vector approach that takes into account the contextual factors in transactions. Our model systematically categorize these factors into service aspect and transaction aspect. In particular, the computation of the elements in this trust vector is associated with both the context of past transactions and the context of a forthcoming transaction, so as to comprehensively indicate the trust level of a seller for the forthcoming transaction. The computed trust vector can be taken as the reputation profile of the seller. Empirical studies illustrate that it is important and necessary to introduce contextual transaction factors in evaluating the trust level of sellers objectively.**

I. Introduction

In e-market environments, when a buyer wants to buy a product, trustworthiness is a key factor in selecting a seller who sells this product [15]. At eBay[1], after each transaction, a buyer can provide a rating (+1, 0, or -1) to the trust management system according to transaction quality. After accumulating over a time period (e.g., six months), a single positive feedback rate is calculated to indicate the trust level of the seller. However, in a simple trust management system, buyers are vulnerable to frauds from malicious sellers. For example, if a seller completed 100 transactions honestly with good transaction quality at a price of $10 for each transaction, his positive feedback rate could be 100%. To potential buyers, this seller appears quite trustworthy. Then, the seller might attract a buyer to complete a transaction of $\$1k$, and deceive by not delivering the purchased product or delivering a fake. In the literature, this problem is called *value imbalance* [6], and several real world cases have been reported [10]. For instance, an Australian reporter tracked down a deceiver at eBay who tricked people for more than AU$\$10k$. A Californian deceiver, in the same way, managed to earn a high positive feedback rate and deceived victims for over US$\$300k$.

The reason causing huge monetary losses is due to the lack of context consideration in transaction trust evaluation. A single trust value (like the positive feedback rate provided at eBay) has concealed contextual transaction information. From buyers' point of view, they are more concerned about the trust level of a seller in a potential forthcoming transaction. But a single trust value fails to predict the likelihood of a seller in a successful forthcoming transaction. In e-commerce environments, different transactions have different *nature* and *contexts*; even the same seller needs to be regarded differently with respect to the trust in different forthcoming transactions, rather than using the same and thus static trust value. Therefore, increasingly more studies introduce the discussion in transaction context of e-commerce environments [17, 18, 22].

Meanwhile, we need to first point out that the *value imbalance problem* is only one type of the *transaction context imbalance problem*, any type of which may cause fraud in forthcoming transactions and lead to monetary losses of customers. For example, at Alibaba[2], which was founded in 1999 and supports both B2B and B2C online trading with 50 million users, following a few cases of fraud, buyers are explicitly suggested to manually check if the products offered by a supplier are in the same category as the products that the supplier usually sells [1]. This suggestion aims to prevent frauds with the *context imbalance problem*. Also, it indicates that transaction reputation evaluation should be "transaction context-aware" and consider the contexts of both past transactions and the forthcoming one. Note that the transaction context imbalance problem also exists in e-service environments where there are service providers and service customers. In the context of this paper, we do not explicitly differentiate e-commerce and e-service but taking e-commerce as an example in analysis.

In this paper, we target the transaction context imbalance problem in e-commerce and e-service environments. Our work and contributions can be summarized as follows:

(1) First, we model transaction context with several contextual attributes and first define the concept of *transaction context imbalance* problem and several types of it in e-commerce environments (*see Section III*).

(2) Secondly, when there are no or not enough ratings from the transactions with the same context as the forthcoming transaction, we propose a set of methods to calculate transaction context similarity, the results of which are used in inferring the trust level of a forthcoming transaction (*see Section IV*).

(3) Thirdly, a trust vector approach to transaction context-

[1] http://www.ebay.com/

[2] http://www.alibaba.com/

aware trust evaluation is proposed. With our proposed approach, a set of trust values are computed which can outline the *reputation profile* of a seller. Even with the same seller, the computed values in the trust vector may vary with the context changes in different forthcoming transactions. In such a way, potentially malicious transactions with the transaction context imbalance problem can be easily identified and prevented (*see Section V*).

(4) Finally, the results of conducted empirical studies illustrate that it is important and necessary to introduce contextual transaction factors in evaluating the trust level of sellers objectively (*see Section VI*).

II. Related Work

A. Trust Evaluation without Contextual Information

Recognizing the importance of trust evaluation, a large number of existing studies aim to build up trust evaluation models in different application fields. For example, in Peer-to-Peer (P2P) information sharing networks, a "global" trust value of a given peer is calculated via collecting binary trust ratings [5]. However, these existing trust evaluation models focus on computing a single *"final trust level"* (e.g. a value in the range of $[0, 1]$ [22, 19]) to reflect the "general" or "global" trust status of every seller. However, these approaches do not take any contextual information into account. With such a result, a buyers can hardly know under what kind of circumstances the seller has obtained a high trust value. So it fails to predict the likelihood of the seller for a successful forthcoming transaction.

B. Context-Aware Trust Evaluation

There are some existing studies considering the relationship between trust evaluation and context information. Griffiths [4] proposed a Multi-Dimensional Trust (MDT) model. The trustworthiness of a particular task can be modeled along several dimensions (e.g., timeliness, quality and cost), letting a trustor specify the weights of each dimension for trust evaluation based on personal preference. Thus given the same seller, the trust results computed for different buyers may be different. Similarly, in REGRET [11] and RATEweb systems [7], multi-dimensional attributes (e.g., delivery and product quality) were adopted when calculating trust. But these models overlook the fact that some transaction context attributes (e.g. transaction amount) may vary in historical transactions. Therefore, they also fail to predict the likelihood of the seller in a successful forthcoming transaction.

In the literature, context similarity calculation is regarded as an important means to deal with contextual trust evaluation problem. Uddin *et al*. proposed a CAT (Context-Aware Trust) model to compare the similarity of contexts by using key values that could describe a certain context to some extent [14]. Caballero *et al*. [3] defined a formula using task key values to calculate the similarity between two tasks in order to evaluate the trust level of different tasks. In e-commerce environments, Zhang *et al*. [23] introduced the concept of similarity between past transactions and a forthcoming transaction when evaluating transaction trust, then proposed some formulae to calculate transaction context similarity to infer the trust value of a forthcoming transaction.

However, these models focus on combining all the contextual attributes (factors) and computing a single value, which can not distinguish the contribution of each contextual attribute specifically. In our paper, first of all, we discuss and model the contextual transaction attributes in e-commerce environments. In addition, compared with single value based approaches, our proposed trust vector could reflect the trust status of a seller more precisely in any forthcoming transaction.

C. Trust Vector Approaches

In the literature, there also exist some approaches using trust vectors, but most of them have no focus on context-aware trust evaluation. In [9], Ray *et al*. proposed a trust vector that consists of experience, knowledge and recommendation. The focus is how to address these three independent aspects of trust, as listed in a trust vector. In [16], Wang *et al*. used a trust vector to describe the trust level and trust trend of sellers. In the literature, a preliminary trust vector approach for evaluating the transaction context-aware trust was first proposed by Wang and Lim [17]. In their work, the trust evaluation model took both past transactions and the forthcoming one into account, but it did not consider transaction context similarity.

III. Transaction Context

In this section, we first analyse and model transaction context for trust evaluation. Finally, we propose the concept of *transaction context imbalance* problem and describe several types of context imbalance that exist in e-commerce environments.

A. What is Transaction Context

The Webster's dictionary defines "context" as the "conditions or circumstances which affect some thing." But different application fields have many definitions of context. For instance, recommender systems, which recommend products to potential buyers, operate in the *Buyer* $\times$ *Item* space. The discussion of context in it focuses more on the *Buyer* space (e.g., age, preference and purchase history) that could affect the decision of making a recommendation [2]. In contrast, trust management systems operate in the *Seller* $\times$ *Transaction* space. As sellers's information is relatively simple and static, in our work, the discussion of context focuses on the transaction space w.r.t transaction trust evaluation.

Definition 1: *Transaction context* refers to *the factors that can determine, imply or affect the trustworthiness of a forthcoming transaction*.

B. Transaction Context Modeling

Based on Definition 1, our work considers the factors with influence on the trustworthiness of a forthcoming transaction. We categorize these factors into two classes: transaction attributes and service quality attributes.

(1) *Transaction Attributes*

• **Transaction Item**: *Transaction item* refers to the product traded in a transaction. The quality of the transaction item partially determines the nature of the transaction and thus greatly influence transaction trust [13]. In addition, the properties of a transaction item, for example, the price and product category also determine the nature of the transaction.

• **Transaction Amount**: *Transaction amount* refers to the sum of prices of all products in a transaction. A transaction of about US\$10 is obviously different from the one of about US$10K$ in nature. The larger the transaction amount is, the more likely a fraud may happen since the benefits of cheating are greater. For the sake of simplicity, in this paper, there is only one item considered in a transaction. A transaction with multiple transaction items is taken as several transactions with one item in each transaction. Hence, the transaction amount equals to the price of the item in a transaction.

• **Transaction Time**: *Transaction time* is the time when a transaction happens. Transaction trust evaluation is time-sensitive, because transaction quality may change with time [12].

(2) *Service Quality Attributes*

In the process of transactions, the quality of services affects the transaction trustworthiness as well, and buyers prefer to choose the sellers who provide high quality services. With regard to the quality of services, buyers can provide corresponding ratings in some e-commerce websites. For example, at eBay, buyers' ratings also evaluate (a) shipping time (i.e. whether the seller delivers goods on time), (b) communication (i.e. whether the seller has prompt and friendly communication with buyers), and (c) shipping charges (i.e. whether the seller charges a reasonable price for shipment). All these aspects are taken into account in our model.

C. Transaction Context Imbalance

As stated before, malicious sellers and fraudulent transactions could take advantage of transaction trust result without any context considered. Consequently, it may lead to some *transaction context imbalance*, which is first proposed in this paper and can include the following types.

• **Transaction Amount Imbalance**: A seller accumulates a high trust level by offering cheap and attractive products, then may deceive buyers with expensive products [18, 22]. In the literature, this issue is also termed as *value imbalance* [6].

• **Transaction Item Imbalance**: A seller has accumulated a high trust level by selling certain products, and then he/she can utilise this high trust value to sell different products for more profit. However, the seller should have different trust levels with respect to different products. For instance, a seller, who sold expensive *Handbags* in the past and now starts to sell a certain type of *Notebook Computer*, should not have a high trust level due to the lack of sufficient experience and reputation in selling new products with a completely different nature.

The focus of our paper is to identify and prevent potentially malicious transactions with respect to different types of transaction context imbalance.

IV. Similarity Comparison between Transaction Attributes

In order to obtain the trustworthiness of a seller in a specific transaction context, a buyer needs to take other buyers' ratings on this seller with the same transaction context to the forthcoming transaction into account. But when there are no or not enough ratings for the same transaction context, it will a good practice to derive the trust level of the seller from all the ratings in any related transaction context. In such a situation, the context similarity between the forthcoming transaction and a past transaction should be compared to weigh the rating for the past transaction [8]. According to our modeled transaction context, we present and propose methods below to compute transaction attributes similarity.

A. Similarity Comparison of Transaction Items

Considering the attributes of a transaction item, for example, a buyer plans to buy a *'Cannon EOS T3i SLR (single-lens reflex) Digital Camera'* from a seller. He would be also concerned about the trustworthiness of this seller in selling various *'Cannon SLR digital cameras'* or the trustworthiness of this seller in selling various *'SLR Digital Camera'*. In subsection IV-A1, we will propose a hierarchical structure of product categories. With this structure, any two products can be compared for similarity, with the method to be given in subsection IV-A2.

1) A Hierarchical Structure of Product Category: There exist some *Products and Services Categorization Standards (PSCS)*, such as UNSPSC[3] and eCl@ss[4], which aim at grouping similar products and provide an industry-neutral hierarchical structure of product categories with up to four layers. eBay has a different schema with simply two layers in product category, and it groups products by considering some factors such as marketing and common use.

Our approach extends eCl@ss due to its reasonable classification in practice, i.e., products and services in eCl@ss are more functionally grouped, and they are subdivided for specific usage. For example, '*Digital Camera*' in eCl@ss can be further classified as '*DSLR (Digital single-lens reflex cameras)*' and '*Compact Digital Camera*', etc; but it is not subclassified in UNSPSC and eBay category.

eCl@ss standards utilize a four-layer hierarchical structure and a two-digit number for each layer. But the logical relation between some products are not reasonable. In our approach, we extend eCl@ss in two ways. First, we sort out the logical relations between products. For example, in eCl@ss, "Photo, video camera" is the supper note of "DSLR Camera", "Compact Digital Camera", "Film camera" and "Video camera". In our proposed product hierarchy, "Photo, video camera" takes "Photo camera" and "Video camera" as the child nodes. "Photo camera" further takes "Digital camera" and "Film camera" as the child nodes. "Digital camera" also has three child nodes: "DSLR", "Compact digital camera" and "Mirrorless interchangeable-lens camera (MILC)". In addition, attribute *Brand* is added to each product category to get more fine-grained analysis on the trustworthiness of a seller during "drill down" and "roll up" operations. Fig. 1 presents a small part of the extended seven-layer product category hierarchy from root to leaf.

2) Similarity Measurement within Product Category: To measure the similarity of two products p and p' within a hierarchical product category, a crucial factor is the depth d of the deepest common ancestor of the two nodes. For instance, the deepest common ancestor of *Digital camera* and *Lens* is *Photo technology* (see Figure 1), and the depth d of *Photo technology* within the product hierarchy is 2. If the deepest common ancestor is high in the hierarchy, it indicates that the two products p and p' have general classification without much similarity between them. If the deepest common ancestor is located in a lower hierarchy, it indicates that the two products have common classification with stronger similarity. Hence, the transaction item similarity S_{TI} between two products p and p'

[3]http://www.unspsc.org/

[4]http://www.eclass.de/

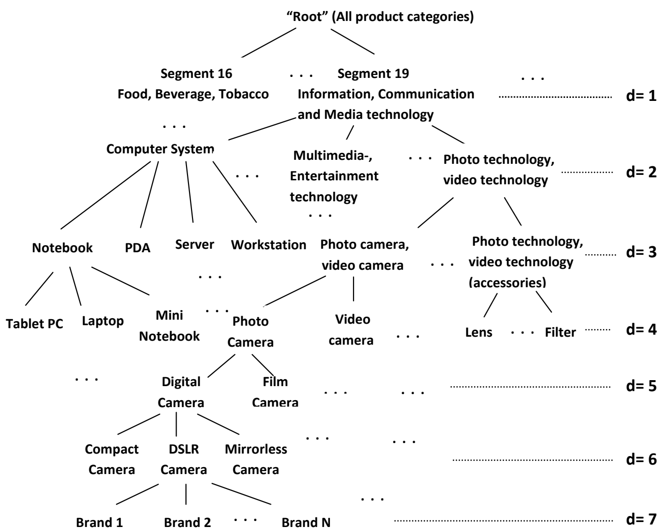

Figure 1. Part of the extended product category hierarchy of segment *"Information, communication and media"*

Table I
EXAMPLES OF TRANSACTION ITEM SIMILARITY

Product Pair (p, p')	S_{TI}	d
Cannon EOS T3i Rebel, Cannon EOS T2i Rebel	0.98	6
Digital camera , Video camera	0.83	3
Digital camera, Lens	0.66	2
Laptop, Camera (digital)	0.38	1
PDA, Food	0	0

should be a monotonically decreasing function with respect to the depth d of the deepest common ancestor of them.

In our earlier work [23], we have proposed a formula to calculate the similarity within a hierarchical structure of product category. Due to space constraints, this paper does not include detailed explanation. Instead, we only list some examples of similarity measure in Table I based on the product category hierarchy in Figure 1.

B. Similarity Comparison of Transaction Amounts

The similarity comparison of transaction amounts is important for trust evaluation, as the significant difference between transaction amount may lead to *transaction amount imbalance*. For calculating the *transaction amount similarity* S_{TA} between the amount a_p of a past transaction and the amount a_f of a forthcoming one, we first consider two typical examples below:

(1) $a_p = \$50$ while $a_f = \$550$;
(2) $a_p = \$5000$ while $a_f = \$5500$.

In existing studies, the *difference* of transaction amounts (denoted as $D_{ta} = |a_f - a_p|$) has been used in trust evaluation [17, 21], but it can not distinguish the above two examples. Although the values of D_{ta} in these two examples are both \$500, \$500 is only 10% increase in example (2) while it is 10-times increase in example (1). To this end, our model also introduces the *relative value* between a_p and a_f, denoted by R_{ta} (i.e., $R_{ta} = \frac{a_f}{a_p}$) for calculating S_{TA}. Both D_{ta} and R_{ta} are used to determine S_{TA} according to the following principles.

Principle 1: When R_{ta} is fixed, S_{TA} increases, if D_{ta} decreases.

Principle 2: When D_{ta} is fixed, S_{TA} becomes larger, if R_{ta} approaches 1.

Based on both principles, some definitions for calculating S_{TA} are proposed below.

Definition 2: If $a_f \geq a_p$, then the *transaction amount similarity* of a_f and a_p is

$$S_{TA}(a_f, a_p) = \varepsilon * f_D(D_{ta}(a_f, a_p)) + (1 - \varepsilon) * f_R(R_{ta}(a_f, a_p)) \quad (1)$$

The definition of function f_D in Eq. (1) can be based on the transaction amount category partitions in e-commerce environments that is proposed in [21].

$$f_D(D_{ta}(a_f, a_p)) = \frac{2}{e^{C(D_{ta}(a_f,a_p))*\beta} + e^{-C(D_{ta}(a_f,a_p))*\beta}} \quad (2)$$

where $C(D_{ta}(a_f, a_p))$ is an integer from 0 to 10 corresponding to the transaction amount range from 0 to over 10^5. $\varepsilon, \beta \in (0, 1]$ and we choose $\varepsilon = 0.5, \beta = 0.2$ as an example in our approach. As a result, if $a_f = 550$ and $a_p = 50$, $f_D(D_{ta}(a_f, a_p)) = 0.75$.

In addition, the definition of function f_R used in Eq. (1) is based on a threshold λ_R of *relative value* ($\lambda_R > 1$), which can be specified as a default value by the trust management authority (e.g. $\lambda_R = 20$). With λ_R, the function f_R can be defined below.

$$f_R(R_{ta}(a_f, a_p)) = \begin{cases} 0, & if \quad R_{ta}(a_f, a_p) \geq \lambda_R \\ \frac{\lambda_R - R_{ta}(a_f,a_p)}{\lambda_R - 1}, & if \quad R_{ta}(a_f, a_p) < \lambda_R \end{cases} \quad (3)$$

In Eq. (3), when $R_{ta} \geq \lambda_R$, $f_R(R_{ta}(a_f, a_p))$ is set to 0, which indicates a large *relative value* in transaction amounts between the forthcoming transaction and the past one. When $R_{ta} < \lambda_R$, $f_R(R_{ta}(a_f, a_p))$ is a projection from 0 to 1 of R_{ta}.

Definition 3: If $a_f < a_p$, then the *transaction amount similarity* of a_f and a_p is set to the maximum 1. Buyers do not need to doubt trustworthiness of the seller on transaction amount.

Following Definitions 2 & 3, when the past transaction amount a_p gets closer to the forthcoming transaction a_f, their transaction amount similarity S_{TA} has a higher value.

C. Similarity Comparison of Transaction Time

As suggested by some studies on trust evaluation, ratings of recent transactions should be assigned higher weights in trust evaluation [7, 16]. We use S_{TT} to denote the transaction time similarity between a forthcoming transaction and a past transaction, and use S_{TT} as the weight for the rating of a past transaction. The higher the transaction time similarity, the higher the weight for the rating of the past transaction.

Definition 4: During a time period $[t_1, t_n]$, where $t_k < t_{k+1}$ $(1 \leq k < n)$, t_n is the most recent transaction time, and the time for a forthcoming transaction is $t_n + 1$ $(t_n + 1 = t_{n+1})$. The transaction time similarity S_{TT} can be calculated as:

$$S_{TT}(t_k, t_{n+1}) = \gamma^{t_n - t_k}, 0 < \gamma < 1, 0 < k \leq n \quad (4)$$

In Eq. (4), the value of γ should be high (e.g., $\gamma = 0.9$) as so to avoid that the weight decreases too rapidly.

V. TRANSACTION CONTEXT-AWARE TRUST METRICS

In this section, we propose a transaction context-aware trust vector and focus on detailed calculation for each element in this vector.

A. Trust Data Representation

In order to calculate transaction context-aware trust, we assume the following data structure.

$$TR^{(t)} =< s; b; r^{(t)}; a; pr; t > \qquad (5)$$

- $TR^{(t)}$ is a transaction between seller s and buyer b at time t;
- p is the product (i.e. transaction item) purchased in transaction $TR^{(t)}$;
- a is the transaction amount;
- $r^{(t)} =< r_q^{(t)}, r_s^{(t)}, r_d^{(t)} >$ is a rating vector that buyer b gives to seller s for $TR^{(t)}$, which consists of three ratings $r_q^{(t)}$, $r_s^{(t)}$ and $r_d^{(t)}$. $r_q^{(t)}$ is the rating for the quality of product p; $r_s^{(t)}$ is the rating for the seller's service (e.g., whether s processed buyer b's order on time and whether s had prompt and friendly communication with b); and $r_d^{(t)}$ is the rating for delivery service. These ratings are also available at eBay.

B. Transaction Specific Trust Evaluation

Transaction Specific Trust is the trust evaluation bound to the transaction item in a forthcoming transaction, which includes three values. They are all calculated based on the ratings $\{r_q^{(t)}\}$.

1) Transaction Item Specific Trust (TIST): Transaction Item Specific Trust (TIST) takes into account past transactions, with the same transaction item as the forthcoming transaction. Three cases are discussed below in the calculation of *TIST*.

Case 1: If there is a sufficiently large number of ratings (named as "direct reference" ratings) from past transactions selling the same item as a forthcoming one, *TIST* can be determined from these ratings directly.

Case 2: If the transaction item in a forthcoming transaction has never been sold by a seller, the *TIST* can be inferred from the ratings on the transactions selling other items (named as "indirect reference" ratings). Transaction Context Similarity should be compared to discount these ratings [8].

Case 3: If both cases are not true, we need to use both "direct reference" ratings and "indirect reference" ratings to compute *TIST*.

Transaction Context Similarity: Before giving formulae for calculating *TIST*, we first introduce *transaction content similarity* S_{TC}, which includes transaction item similarity S_{TI} and transaction amount similarity S_{TA}.

Definition 5: With a set of m past transactions $Trans = \{TR_1, TR_2, ..., TR_m\}$ between buyers and seller s in the time period $[t_1, t_n]$ and TR_f denoting the forthcoming transaction at time t_{n+1}, we define S_{TC} as follows[5]:

$$S_{TC}(TR_i, TR_f) = \frac{S_{TI}(TR_i, TR_f) + S_{TA}(TR_i, TR_f)}{2} \qquad (6)$$

The Calculation of *TIST*: Assume a buyer is planning to buy a product p in a transaction TR_f from seller s.

In Case 1, we use θ to denote the threshold of sufficient number of "direct reference" ratings[6]. Thus, all the "direct reference" ratings are used in the calculation of *TIST*, where the transaction time similarity S_{TT} is used to weight each rating.

$$T_{TIST_{s(p)}}^{[t_1,t_n]} = \frac{\sum_{i=1}^{m_1} (r_q(TR_i)^{(t_i)} * S_{TT}(i, f))}{\sum_{i=1}^{m_1} S_{TT}(i, f)} \qquad (7)$$

In Eq. (7), m_1 is the number of "direct reference" ratings from past transaction set $Trans$, $\theta \le m_1 \le m$.

In Case 2, when evaluating *TIST*, *transaction content similarity* S_{TC} is regarded as the weight to discount these "indirect reference" ratings, and thus *TIST* can be computed as:

$$T_{TIST_{s(p)}}^{[t_1,t_n]} = \frac{\sum_{i=1}^{m} (r_q(TR_i)^{(t_i)} * S_{TC}(i, f) * S_{TT}(i, f))}{\sum_{i=1}^{n} S_{TT}(i, f)} \qquad (8)$$

In Case 3, there are not enough "direct reference" ratings from past transaction set $Trans$. It is necessary to combine both "direct reference" and "indirect reference" ratings, and then give "direct reference" ratings higher weight ω. Moreover, the determination of weight ω should follow some principles.

Principle 3: When the number of "direct reference" ratings increases, the weight for these ratings increases as well.

Principle 4: The initial value of weight ω should be low to avoid that the trust level of a seller s drops very fast after a few misbehaviors at the beginning.

Following these two principles, our model uses a function in Eq. (9) to control the changes of ω.

Definition 6: Given parameters u and v, the weight ω can be calculated as follows:

$$\omega(m_1) = 1 - u^{m_1^{\frac{1}{v}}}, u \in (0, 1) \qquad (9)$$

where u determines the initial value of weight ω and $v \in \{1, 2, 3, ...\}$. According to Principle 4, u should be more than 0.5 (i.e., $0.5 < u < 1$). With two fixed parameters u and v, the larger m_1, the larger $\omega(m_1)$, which conforms to Principle 3. In addition, the value of v depends on the parameter θ (i.e., the threshold of sufficient number of "direct reference" ratings). For example, if θ is 10, $v = 1$ is suitable. If θ is 50, $v = 2$ is suitable.

TIST in case 3 is the weighted summation of $T_{TIST_{s(p)}}^{[t_1,t_n]}$ defined in both Eq. (7) and Eq. (8):

$$T_{TIST_{s(p)}}^{[t_1,t_n]} = \omega(m_1) * \frac{\sum_{i=1}^{m_1} (r_q(TR_i)^{(t_i)} * S_{TT}(i, f))}{\sum_{i=1}^{m_1} S_{TT}(i, f)} +$$
$$(1 - \omega(m_1)) * \frac{\sum_{j=1}^{m-m_1} (r_q(TR_j)^{(t_j)} * S_{TC}(i, f) * S_{TT}(i, f))}{\sum_{j=1}^{n-m_1} S_{TT}(i, f)} \qquad (10)$$

2) Transaction Item Similarity Based Trust (TIBT): In addition to *TIST*, the buyer may also be concerned about whether the seller obtained a high trust level in selling various products similar to p. *TIBT* is computed based on the ratings of transactions containing products similar to that in the forthcoming transaction. Its value is defined in Eq. (11):

$$T_{TIBT_{s(p)}}^{[t_1,t_n]} = \frac{\sum_{k=1}^{m_3} (r_q(TR_k)^{(t_k)} * S_{TT}(TR_k, TR_f))}{\sum_{i=1}^{m_3} S_{TT}(TR_k, TR_f)} \qquad (11)$$

where $m_3 = |\{TR_k | TR_k \in Trans, S_{TI}(TR_k, TR_f) \ge \theta_{TI}\}|$ and $m_1 \le m_3 \le m$; θ_{TI} is the threshold for transaction

[5] For the sake of simplicity, in later discussions of this paper, we denote $S_{TC}(i, f) = S_{TC}(TR_i, TR_f)$, $S_{TI}(i, f) = S_{TI}(TR_i, TR_f)$, $S_{TA}(i, f) = S_{TA}(TR_i, TR_f)$ and $S_{TT}(i, f) = S_{TT}(TR_i, TR_f)$ (defined in Eq. (4))

[6] The parameters θ can be specified by buyers or by the trust management authority.

item similarity S_{TI}, i.e., *TIBT* considers the ratings from the transactions selling products with a similarity larger than θ_{TI}.

3) Transaction Amount Similarity Based Trust (TABT): *TABT* is to outline the trust level of the forthcoming transaction in terms of transaction amount. The *TABT* is different from *TIBT*, because a seller may have lots of past transactions with the amounts similar to the forthcoming transaction but their corresponding transaction items may be in different product categories.

$$T^{[t_1,t_n]}_{TABT_{s(p)}} = \frac{\sum_{l=1}^{m_4} (r_q(TR_l)^{(t_l)} * S_{TT}(TR_l, TR_f))}{\sum_{l=1}^{m_4} S_{TT}(TR_l, TR_f)} \tag{12}$$

where $m_4 =| \{TR_l | TR_l \in Trans, S_{TA}(TR_l, TR_f) \geq \theta_{TA}\} |$ and $m_1 \leq m_4 \leq m$; θ_{TA} is the threshold for transaction amount similarity S_{TA}, i.e., *TABT* considers the ratings with the transaction amount similarity higher than θ_{TA}.

C. Service Specific Trust Evaluation

When a buyer looks for a product, in addition to its quality, services in the transaction process also have impact on transaction trustworthiness. They include sellers' services and delivery services. Thus, based on the ratings $\{r_s^{(t)}\}$ and $\{r_d^{(t)}\}$, we introduce the calculation of *Sellers' Service Trust (SST)* and *Delivery Service Trust (DST)*.

Sellers' Service Trust (SST)**:** Sellers' services include processing the order on time and having good communication with buyers, etc. The rating $r_s^{(t)}$ is used to depict a general trust level of the seller's service quality. The *sellers' service trust (SST)* value is computed based on ratings $\{r_s^{(t)}\}$.

$$T^{[t_1,t_n]}_{SST_s} = \frac{\sum_{i=1}^{n} (r_s{}^{(t_i)} * S_{TT}(i, f))}{\sum_{i=1}^{n} S_{TT}(i, f)} \tag{13}$$

Delivery Service Trust (DST)**:** Delivery services are excluded from sellers' services because they are not operated and controlled by sellers. However, delivery services indeed affect the trust level of sellers. For instance, the advertised delivery time is 3 days, but the actual one may be over one week. Hence, the rating $r_d^{(t)}$ for delivery services would be low. In particular, if a seller often obtains poor ratings $\{r_d^{(t)}\}$ on delivery, it would make them be regarded as untrustworthy. Similar to *SST*, *DST* is calculated based on ratings $\{r_d^{(t)}\}$ with S_{TT} as the weight.

$$T^{[t_1,t_n]}_{DST_s} = \frac{\sum_{i=1}^{n} (r_d{}^{(t_i)} * S_{TT}(i, f))}{\sum_{i=1}^{n} S_{TT}(i, f)} \tag{14}$$

D. Transaction Context-Aware Trust Vector

In this section, a trust vector is defined to represent the transaction context-aware trust level of a seller.

Definition 7: Given a buyer b who is planning to buy a product p from a seller s in a forthcoming transaction TR_f, the transaction context-aware trust vector $CTT^{[t_1,t_n]}_{s(p)}$ of seller s during the time period $[t_1, t_n]$ is defined as follows:

$$CTT^{[t_1,t_n]}_{s(p)} = \; < T^{[t_1,t_n]}_{TIST_{s(p)}}, T^{[t_1,t_n]}_{TIBT_{s(p)}}, T^{[t_1,t_n]}_{TABT_{s(p)}}, T^{[t_1,t_n]}_{SST_s}, T^{[t_1,t_n]}_{DST_s} > \tag{15}$$

where $T^{[t_1,t_n]}_{TIST_{s(p)}}$, $T^{[t_1,t_n]}_{TIBT_{s(p)}}$ and $T^{[t_1,t_n]}_{TABT_{s(p)}}$ are the specific trust values bound to the transaction item p in the forthcoming transaction. This trust vector can be provided to buyers to outline the reputation profile of the seller s.

VI. Empirical Studies

In this section, some empirical studies are presented to illustrate the effectiveness our proposed trust vector approach for transaction context-aware trust evaluation.

A. Study 1 - Transaction Item Specific Trust (TIST)

In this study, an example is used to study the changes of *TIST* when a seller provides new products.

Example: Four sellers s_1, s_2, s_3 and s_4 provide a latest popular model of Apple MacBooK Pro laptop (e.g., *MC700LL/A*) with an attractive price of around \$900 at time t ($t > t_{50}$). Assume that the four sellers sold different products respectively before, but since t_{51} they start to sell this popular laptop. The products that they have sold before are listed in Table II below.

Table II
THE PRODUCTS SOLD BY FOUR SELLERS

	transaction context			
	time period $[t_1, t_{50}]$		time period t_{51} and after	
Sellers	Product	Price	Product	Price
s_1	Apple iPad3	\$600	MacBook Pro Laptop	\$900
s_2	Canon A2200 Digital Camera	\$150	MacBook Pro Laptop	\$900
s_3	Luxury Watch	\$2500	MacBook Pro Laptop	\$900
s_4	Food	\$10	MacBook Pro Laptop	\$900

Furthermore, the example adopts the rating model where each rating $r_q^{(t)}$ is an integer in $\{1, 2, 3, 4, 5\}$[7]. We normalize ratings to $[0, 1]$, namely $\{0, 0.25, 0.5, 0.75, 1\}$. A good quality transaction means that the seller provides "good quality" product, thus the value of $r_q^{(t)}$ lies in the range $[0.75, 1]$. In contrast, the $r_q^{(t)}$ for a poor quality transaction with "poor quality" product lies in the range $[0, 0.25]$. We adopt other parameters $\gamma = 0.9$ in Eq. (4), $v = 2$ and $u = 0.7$ in Eq. (9), $\lambda_R = 20$ in Eq. (3), and $\theta = 20$. We also assume that the four sellers all had good quality transactions during time period $[t_1, t_{50}]$ and obtained similar high rating values. After time t_{50}, s_1 and s_3 still have good quality transactions but s_2 and s_4 start to provide poor quality transactions (e.g., they sell refurbished laptops).

Evaluation: Following a typical trust evaluation approach without context consideration in [20], the general transaction trust (*GTT*) can be calculated as:

$$T^{[t_1,t_n]}_{GTT_s} = \frac{\sum_{i=1}^{n} (r_q{}^{(t_i)} * S_{TT}(TR_i, TR_f))}{\sum_{i=1}^{n} S_{TT}(TR_i, TR_f)} \tag{16}$$

where $TR_i \in Trans$ is past transactions, TR_f is the forthcoming transaction and S_{TT} is transaction time similarity (defined in Eq. (4)), used to weight the rating of TR_i.

Analysis: As shown in Figure 2, the *TIST* values and the *GTT* values of four sellers are very close to each other in time period $[t_1, t_{50}]$ (i.e., $T^{[t_1,t_{50}]}_{GTT} \approx 0.89$), when the products they sell do not change.

At time t_{51}, when the four sellers start to provide new products, their *TIST* values decrease differently. The extent of decrease depends on the similarity between past transactions

[7]This rating model provides more accurate information than the tripple-rating model with 1 for positive, 0 for neutral and -1 for negative as eBay [7]

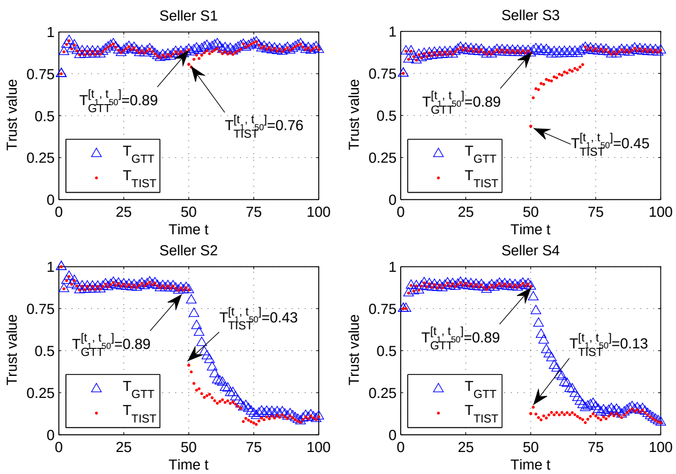

Figure 2. The changes of Transaction Item Specific Trust (TIST) for four different sellers

and the new transaction at time t_{51}. For example, s_1, who has approximately the same transaction items (iPad3 *vs* Laptop) and transaction amounts ($600 *vs* $900) as those in the past transactions, the *TIST* value decreases slightly only. Both s_2 and s_3, with decreased transaction context similarity (i.e., Digital Camera *vs* Laptop and $150 *vs* $900 for s_2, Luxury Watch *vs* Laptop and $2500 *vs* $900 for s_3), the *TIST* value drops remarkably. However, it also can be observed in Figure 2(c) that $T_{TIST_{s_3}}$ quickly reaches a high value after accumulating for a while (i.e., above 0.8 after t_{75}) when s_3 keeps selling high quality laptops. s_4, who starts to have transactions at t_{51} with completely different transaction context (Food *vs* Laptop and $10 *vs* $900 for s_4), his *TIST* values are quite low due to the potential *transaction context imbalance*.

After time t_{51}, when the four sellers have completed a series of transactions, the *TIST* values and the *GTT* values tend to be close to each other again due to the increased number of "direct reference" ratings. The *TIST* values can be determined from them directly (Case 1 in Eq. (7)), which are the same as the *GTT* values (calculated in Eq. (16)).

Summary: Obviously, it is unreasonable to use *GTT* to indicate the trust level of sellers, because it can not reflect the changes of trust level when a seller starts to provide new products. In contrast, our proposed *TIST* takes into account the transaction context similarity. Only for those sellers who provide transaction items and transaction amounts similar to a forthcoming transaction, the trust value will be as high as before (such as s_1). Therefore, *TIST* can be used to identify some potentially malicious transactions with the *transaction amount imbalance problem*; typically, such as s_4, who obtains a quite low *TIST* value with our model.

B. Study 2 - Comparison with Existing Context-Aware Trust Evaluation Approaches

In this study, we compare our approach with existing context-aware trust evaluation approaches.

*1) **Comparison with REGRET**:* In REGRET [11], multiple dimensions of context are considered in trust evaluation. A buyer can specify the weight in each dimension based on personal preference, and then an aggregated value can be computed for trustworthy seller selection.

Assume two other sellers s_5 and s_6 obtained the following ratings in time period $[t_1, t_{10}]$ as listed in Table III. We compare the trust value of two sellers computed based on REGRET model and our approach as shown in Table IV, and we set the same parameters as the example introduced in Section VI-A.

Table III
RATING VECTORS OF SELLERS s_5 AND s_6

transaction time	s_5 $r_q^{(t)}$	$r_s^{(t)}$	$r_d^{(t)}$	transaction item
t_1	0.75	1	0.75	iPad3
t_2	0.75	0.75	1	MacBook Pro Laptop
t_3	0.5	1	1	iPad3
t_4	1	0.75	0.75	iPad3
t_5	0.75	0.75	0.75	MacBook Pro Laptop
t_6	0.5	1	0.75	iPad3
t_7	0.75	0.75	1	iPad3
t_8	1	1	0.75	MacBook Pro Laptop
t_9	0.75	0.75	1	iPad3
t_{10}	0.5	0.75	1	iPad3

transaction time	s_6 $r_q^{(t)}$	$r_s^{(t)}$	$r_d^{(t)}$	transaction item
t_1	1	0.75	1	Canon A2200 Digital Camera
t_2	1	1	1	Canon A2200 Digital Camera
t_3	0.75	0.75	0.75	Canon A2200 Digital Camera
t_4	0.5	1	1	MacBook Pro Laptop
t_5	0.75	1	0.75	Canon A2200 Digital Camera
t_6	1	1	0.75	Canon A2200 Digital Camera
t_7	0.5	0.75	1	MacBook Pro Laptop
t_8	1	1	0.75	Canon A2200 Digital Camera
t_9	0.25	0.75	1	MacBook Pro Laptop
t_{10}	1	1	1	Canon A2200 Digital Camera

In REGRET, as both sellers sell this laptop with an attractive price $900, they have the same trust level on price. In addition, based on the ratings $\{r_q^{(t)}\}$ in time period $[t_1, t_{10}]$, we can know that the trustworthiness of a new transaction selling this laptop is 0.71 for s_5 and the trustworthiness of service quality is 0.89 for s_5 based on ratings $\{r_d^{(t)}\}$. Similarly, we can know that the trustworthiness of a new transaction selling this laptop is 0.74 for s_6 and the trustworthiness of service quality is 0.90 for s_6 (see Table IV). Thus, with any weights specified by a buyer that are used for both sellers simultaneously, the aggregated trust value of s_6 is always greater than s_5, i.e., s_6 is more trustworthy, denoted as $T_{s_5}^{[t_1,t_{10}]} < T_{s_6}^{[t_1,t_{10}]}$.

Table IV
THE TRUST VALUE OF TWO SELLERS COMPUTED BASED ON REGRET AND OUR APPROACH

Seller	REGRET: the trustworthiness of transaction	REGRET: the trustworthiness of service	Our approach: T_{TIST}	T_{TIBT}	T_{TABT}	T_{SST}	T_{DST}
s_5	0.71	0.89	0.75	0.72	0.72	0.86	0.88
s_6	0.74	0.90	0.40	0.76	0.40	0.90	0.90

In Table III, the ratings $\{r_q^{(t)}\}$ of s_5 for selling MacBook Pro laptop are higher than s_6. Compared to the single value result $T_{s_5}^{[t_1,t_{10}]} < T_{s_6}^{[t_1,t_{10}]}$ from REGRET system, our approach is more reasonable when selecting a seller for a specific product, since the computed trust vector is particularly bound to each forthcoming transaction. As we can see in Table IV, s_6 has lower T_{TIST} value (0.40 *vs* 0.75) and T_{TABT} value (0.40 *vs* 0.72) than s_5.

*2) **Comparison with the trust evaluation approach in [17]**:* In [17], Wang *et al*. proposed a preliminary trust vector approach for evaluating the transaction context-aware trust, which includes 1) *product specific trust (ST)*, calculated from the ratings of all past transactions selling the same product as the forthcoming transaction; 2) *product category specific trust*

(SCT), based on the ratings of past transactions having the same product category as that in the forthcoming transaction; 3) *transaction amount specific trust (TAST)*, based on the ratings of past transactions having the same transaction amount category as that in the forthcoming transaction; and 4) *global weighted trust (GWT)*. The *GWT* value is the weighted average of ratings from all past transactions in a recent time period, in which transaction amount difference between the forthcoming transaction and each past transaction is used to calculate the weight for the rating of the past transaction.

However, their model did not consider transaction item similarity (i.e., product hierarchical structure). As a result, the selection of the most trustworthy seller may be unreasonable when a seller just starts to sell a new product in the forthcoming transaction. As introduced in Section VI-A, for two sellers s_1 and s_3, we assume that s_3 sold Luxury Watch for $1000. When s_1 and s_3 start to sell a model of MacBook Pro Laptop at time t_{51}, the calculated trust vectors of s_1 and s_3 using the approach in [17] and our approach (with $\theta_{TI} = 0.8$ and $\theta_{TA} = 0.8$) are both listed in Table V, respectively.

Following the model in [17], with any specified weights, considering the aggregated trust values of two sellers, we always have $T_{s_3}^{[t_{41},t_{50}]} > T_{s_1}^{[t_1,t_{50}]}$ at t_{51}. That is because both s_1 and s_3 have the same T_{ST}, T_{SCT} and T_{TAST} values, but T_{GWT} of s_3 is higher than s_1 (0.86 *vs* 0.67). In contrast, in our model, transaction item similarity is introduced. Thus, seller s_1, who has sold more transaction items similar to the one in the forthcoming transaction, obtains a higher transaction specific trust value. Compared with s_3, s_1 has higher T_{TIST} value (0.76 *vs* 0.45) and T_{TIBT} (0.89 *vs* 0) value.

Moreover, qualitatively, in [17], transaction context is limited to transaction attributes only, but more transaction contextual information should also be taken into account like our proposed approach, such as service quality attributes.

Table V
THE CALCULATED TRUST VECTOR OF TWO SELLERS AT TIME t_{51} PROPOSED IN [17] AND IN OUR APPROACH

	The approach in [17]				Our proposed approach		
Sellers	T_{ST}	T_{SCT}	T_{TAST}	T_{GWT}	T_{TIST}	T_{TIBT}	T_{TABT}
s_1	0	0	0.89	0.67	0.76	0.89	0.89
s_3	0	0	0.89	0.86	0.45	0	0.89

VII. CONCLUSION AND FUTURE WORK

In this paper, we propose a trust vector approach to transaction context-aware trust evaluation. The trust vector includes a set of valuable information to outline the sellers' reputation, which is available to customers and can help identify and prevent some potentially malicious transactions with *transaction context imbalance*. The proposed approach is based on real e-commerce websites (e.g., eBay) and thus can be easily incorporated.

Regarding the future work, transaction context-aware trust computation is dynamic and complex because hierarchical structures are taken in account (e.g. in product category). In order to clearly outline a seller's reputation profile, if a user adjust argument/layer in product category, accordingly, different ratings from different transaction contexts need to be taken into account for computation. If this trust vector is applied in e-commerce environments with millions of transactions, all these factors incur high computational complexity. Therefore, our research will focus on designing new data structures and efficient algorithms to facilitate buyers' context-aware trust enquiries on each element of the trust vector.

REFERENCES

[1] Protect yourself from fraudsters pretending to be gold suppliers, http://resources.alibaba.com/article/232530/protect_yourself_from_fraudsters_pretending_to_be_gold_suppliers.htm.

[2] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6):734–749, 2005.

[3] A. Caballero, J. Botia, and A. Gomez-Skarmeta. On the behaviour of the trsim model for trust and reputation. In *German Conference on Multi-Agent System Technologies*, pages 13–24, 2007.

[4] N. Griffiths. Task delegation using experience-based multi-dimensional trust. In *AAMAS*, pages 489–496, 2005.

[5] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina. The eigentrust algorithm for reputation management in p2p networks. In *WWW*, 2003.

[6] R. Kerr and R. Cohen. Modelling trust using transactional, numerical units. In *ACM PST*, 2006.

[7] Z. Malik and A. Bouguettaya. Rateweb: Reputation assessment for trust establishment among web services. *The International Journal on Very Large Data Bases*, 18(4):885–911, 2009.

[8] L. Mui. Computational models of trust and reputation: Agents, evolutionary games, and social networks. *PhD Thesis, University of MIT*, 2003.

[9] I. Ray and S. Chakraborty. A vector model of trust for developing trustworthy systems. In *European Symposium on Research Computer Security*, pages 260–275, 2004.

[10] B. Rietjens. Trust and reputation on ebay: Towards a legal framework for feedback intermediaries. *Information and Communications Technology Law*, 15(1):55–78, 2006.

[11] J. Sabater and C. Sierra. Regret: Reputation in gregarious societies. In *ACM AGENTS*, pages 194–195, 2001.

[12] S. Spitz and Y. Tuchelmann. A trust model considering the aspects of time. In *ICCEE*, pages 550–554, 2009.

[13] T. Tran and R. Cohen. Improving user satisfaction in agent-based electronic marketplaces by reputation modelling and adjustable product quality. In *AAMAS*, pages 828–835, 2004.

[14] M. Uddin, M. Zulkernine, and S. Ahamed. Cat: A context-aware trust model for open and dynamic systems. In *ACM Symposium on Applied Computing*, pages 2024–2029, 2008.

[15] L.-H. Vu, M. Hauswirth, and K. Aberer. Qos-based service selection and ranking with trust and reputation management. In *IEEE International Conference on Cooperative information system*, pages 466–483, 2005.

[16] Y. Wang and L. Li. Two-dimensional trust rating aggregations in service-oriented applications. *IEEE Transactions on Service Computing*, 4(4):257–271, 2011.

[17] Y. Wang and E.-P. Lim. The evaluation of situational transaction trust in e-service environments. In *ICEBE*, pages 265–272, 2008.

[18] Y. Wang and K.-J. Lin. Reputation-oriented trustworthy computing in e-commerce environments. *IEEE Internet Computing*, 12(4):55–59, 2008.

[19] Y. Wang and V. Varadharajan. Trust2: Developing trust in peer-to-peer environments. In *SCC*, pages 24–31, 2005.

[20] Y. Wang and V. Varadharajan. Two-phase peer evaluation in p2p e-commerce environments. In *IEEE International Conference on e-Technology, e-Commerce and e-Service*, pages 654–657, 2005.

[21] Y. Wang, D. S. Wong, K. J. Lin, and V. Varadharajan. Evaluating transaction trust and risk levels in peer-to-peer e-commerce environments. *Information Systems and e-Business Management*, 6(1):25–48, 2008.

[22] L. Xiong and L. Liu. A reputation-based trust model for peer-to-peer ecommerce communities. In *CEC*, pages 275–284, 2003.

[23] H. Zhang, Y. Wang, and X. Zhang. Transaction similarity-based contextual trust evaluation in e-commerce and e-service environments. In *ICWS*, pages 500–507, 2011.