

STRATEGISING CONSUMER LOGISTIC REQUIREMENTS IN ECOMMERCE TRANSACTIONS: EVALUATION OF CURRENT IMPLEMENTATIONS

Farid Meziane and Mohd Khairudin Kasiran
School of Computing, Science and Engineering,
University of Salford, Salford M5 4WT, UK

ABSTRACT

Lack of trust has been identified as one of the reasons why many visits to eCommerce websites do not turn into proper transactions. To support trust, an information framework model based on research on eCommerce trust has been developed. The model identifies the kind of information a consumer expects to find on an eCommerce website and that is shown to increase his trust toward the online merchant. An information extraction system has been developed to help gather the required information from the websites. In this paper, we first validate the information model through a questionnaire using a consumer sample. This is then followed by an evaluation of the current implementation of eCommerce websites with regards to the developed trust information model.

1. INTRODUCTION AND MOTIVATION

One of the areas that has been revolutionised the most by the development of the internet is business and commerce. Business transactions are no longer bound by geographical boundaries, time differences or distance barriers. This new business setting is known as eCommerce. ECommerce is the success story of the last few years. However, there are many hindrance factors which cause it to fail to reach its full potential. Han and Noh [8] found that several critical failure factors of eCommerce need to be addressed seriously by the eCommerce industry to ensure its usage will continue to grow. Their findings are mainly on the dissatisfaction of customer on the unstable eCommerce system, a low level of personal data security, inconvenience system and disappointing purchases. Other problems with eCommerce include delays in deliveries, quality of the goods and fraud. Indeed, consumers' loss to Internet fraud has increase from US\$3.2 millions in 1999 [2] to more than US\$ 14.5 millions in 2002 [16] and this is increasing every year. This has affected consumers' trust towards online business. The question that many customers are asking is "who to trust in the cyber space?" and most importantly, how to quantify trust? Many variables should be considered when attempting to quantify or just trying to understand the trust relationship between the vendor and the customer.

In this paper, we present a model for eCommerce trust, its implementation and an evaluation of current eCommerce websites based on our model. The remaining of the paper is organised as follows. In section 2 we review some of the models used to evaluate trust in eCommerce and in section 3 we present our proposed trust model. In section 4 we summarise the extraction system developed to support the gathering of the various variables required by the trust model and in section 5 we present three methods for the evaluation of the trust model. We conclude and evaluate our system in section 6.

2. ECOMMERCE TRUST

The concept of trust is becoming the driving force behind the design, evaluation and use of eCommerce websites and is getting a lot of interest from many researchers. Many approaches have been used to understand and evaluate trust. Indeed, while browsing an eCommerce website customers are faced by many uncertainties. Araujo and Araujo [1] classified these uncertainties and risks as belonging to one of these two categories: Technology related (security, privacy, and integrity) or business related (misuse of personal information and incorrect fulfilment of transactions)

The MoTEC Model [5] was developed taking into account the psychological, marketing and management concerns of trust. A list of factors likely to influence the development and maintenance of trust in the domain of Business to Customer (B2C) e-Commerce have been identified and then integrated as components of the MoTEC Model. There are four components in the MoTEC models and these are:

- *Pre-interactional Filters:* This look at the reputation of the company, the strength of its brand and the customer's interaction history with the organisation.
- *Interface Properties:* This look at the content, usability and structure of the organisation's website.
- *Informational Content:* This includes information about products and services, company's history, values and commitments, privacy and statement.
- *Relationship Management:* This include the communication and interaction facilities with the organisation

Lately, reputation based systems have been used as a medium to evaluate trust. Xiong and Liu [19] developed an adaptive trust model for quantifying and comparing the trustworthiness of peers in Peer-to-Peer eCommerce communities based on a transaction-based feedback system. In their model, trustworthiness was defined by an evaluation of the peer in terms of its reputation in providing services to other peers in the past. The trust model is then defined based on five factors that include the amount of satisfaction a peer obtained, the number of transactions, the credibility of peers that submitted feedback, a transaction context factor and the community context factor. Manchala [13] has identified models based on Boolean relationships, fuzzy logic, transaction processes and transaction automaton.

3. THE PROPOSED TRUST MODEL

3.1 The Trust Information Model

The model we have developed is based on the information that is shown to increase customer trust when present on a merchant website as supported by the literature in Ecommerce trust and a questionnaire conducted on a sample of users, mainly students from the United Kingdom and Malaysia. However, the presence of the information alone is not sufficient. The validity and veracity of the information is very important if we want to provide a valid and consistent instrument to measure the trust of a merchant's website.

Hence the variables retained are those that can be verified by other means such as email, a telephone call or the verification with a third party. The information trust model is summarised in Figure 3.1 and described in the next subsections.

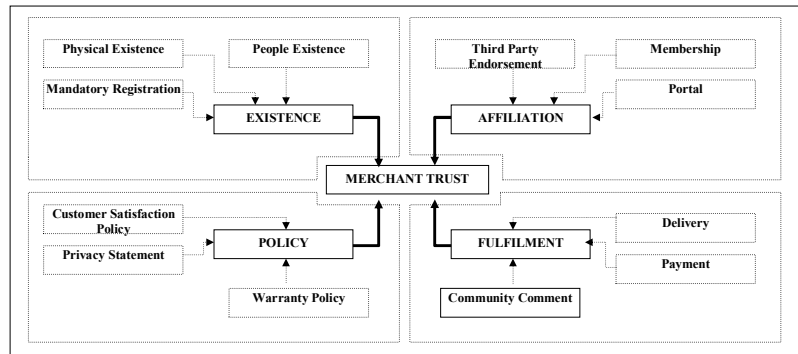


Figure 3.1: The trust information model

3.2 The Existence Component

In a trust relationship the existence of the parties involved need to be established before trust can be developed between them. Since consumers are the initiators of eCommerce transactions, merchants need to communicate that they “officially” exist behind their websites. Providing information about the company's physical existence such as address and telephone number can convey the message that the company is reachable outside the cyber world which in turn will give more control and alternatives to the user to initiate communication when needed. In addition, providing information regarding registration with certain governmental bodies would increase the trust. Information such as the company's registration number and the registering body will help the consumer to verify the merchant validity. The variables retained for the existence model are: physical existence (E1), people existence (E2) and mandatory registration (E3).

3.3 The Affiliation Component

A strong trust relationship can be established between two parties if it is being developed through direct experience. However, for new users, recommended trust can be used to establish the initial trust relationship [3, 17]. Several possible methods of affiliation that can be looked at closely in eCommerce environment are third party endorsement, membership registration and portal linkages. The influences of third party endorsement for example will become more significant to unknown merchants where the perceived risk is higher than well-known merchant like Amazon and eBay. Membership registration to certain bodies and organisations can be used to create recommended trust in areas where skill and expertise is important. Merchant trust can also be sparked through the digital entrance affiliation or portal. Well trusted portals usually gather trusted merchants in their digital supermarket. The variables retained for the affiliation component are: third party endorsement (A1), membership (A2) and portal (A2).

3.4 The Policy Component

Merchant policies are important in creating merchant trust since they set the guidelines for the methods and procedures used in running the business. In eCommerce, policies such as privacy policy, customer satisfaction policy and guarantee policy can help consumers evaluate the trustworthiness of a merchant. These policies can influence the level of risk involved in the transaction. Policy such as money back guarantee can lower consumers' risk by giving more control to the user towards the output of the transaction since they can return the product without total loss if they are not satisfied with their quality. As stated in many researches, one of the main factors that hindered consumer from being involved actively in the digital market is related to consumer privacy [3, 9]. Consumers are afraid that their personal data will be sold to other parties or being used in marketing databases. The variables retained for the policy components are: customer satisfaction policy (P1), privacy statement (P2) and warranty policy (P3).

3.5 The Fulfilment Component

It is important for online merchant to communicate their ability to fulfil their duties with regards to delivery and payment methods to consumers. Since consumers have fulfilled or partially fulfilled their duty by paying for the goods instantly when completing the online transaction by providing for example credit card details, merchants need to tell consumers how and when they will deliver the product. The information that needs to be included for example is the delivery method, the company's name and order tracking method. Tracking the merchant's reputation is considered to be an antecedent toward establishing trust environment towards the merchant [11]. Reputation conveys some information about the merchants' performance as well as behaviour in the past. A positive reputation can create basic building block of merchant trust and carry some assumption that the merchant will perform and behave in the same manner in the future. The variables retained for the fulfilment module are: delivery (F1), payment (F2) and community comments (F3)

4. THE INFORMATION EXTRACTION SYSTEM

4.1 Overall Approach and System's Architecture

The information identified in the trust model is helpful only if customers can have access to it in a reasonable amount of time. If performed manually, this may take a long time and hence discourage many customers to search for the information. To increase the usability of the model, we have developed an information extraction system that will help customers finding the required information.

The entry to the eCommerce extraction system is the website's URL. The system attempts then to extract the information required for each component of the trust model. Once the top level of the website is loaded, extraction rules will be applied for each component. For example, for the Existence component the system will attempt to extract the merchant's phone number, fax number and physical address. If any module fails to extract the required information then the links on the page are collected and navigation rules are applied for a better selection of the links to be used in the next steps. The

extracted information from all components is then stored in a database that will be used to evaluate the trust factor associated with the merchant's website. The overall architecture of the system when applied to the existence module is shown in Figure 4.1.

We adopted a learning approach for the system. We have first identified 50 websites and hand crafted the extraction rules for the various factors defined in the trust model. The stores are selected randomly using several ShoppingBot. Requests are then made to purchase few items such as a book, a digital camera and chocolate.

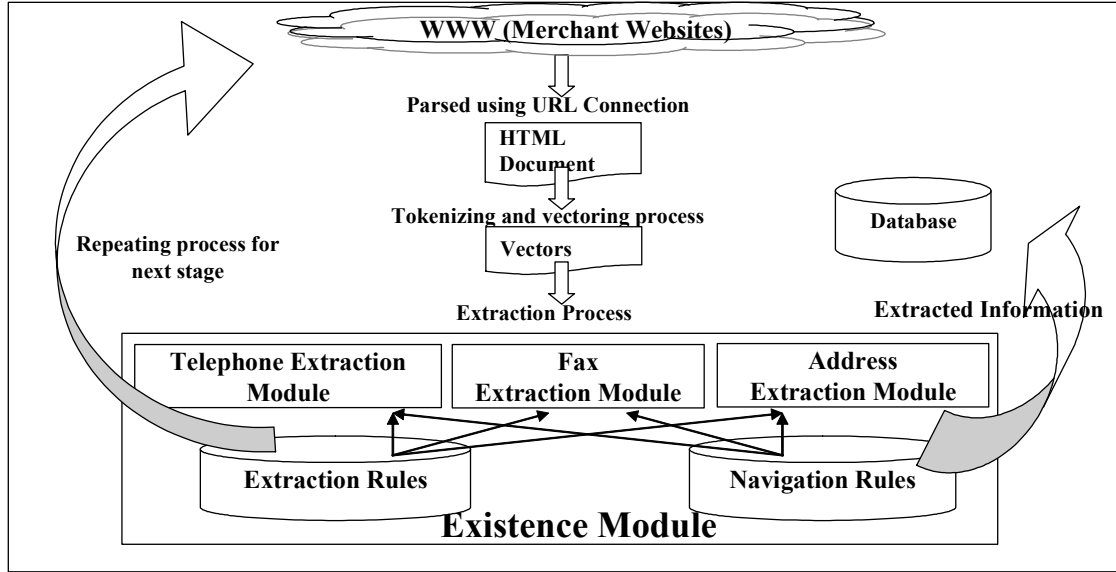


Figure 4.1: Extraction system's overall architecture

4.2 Extraction Rules

The format of an extraction rule is defined as follows:

```
extraction_rule = precede_expression; item_structure; follow_expression
```

Where `precede_expression` is the information that precedes a trust item, `item_structure`, the item's structure and `follow_expression` is the expression that follows it. For example, a telephone number can be preceded by the string "call us at:" and the structure of the phone number is a numerical value. See [14] for the detailed definition of the extraction rules for the existence component. We formulate the rules by using Horn clauses as in [4]. An example rule for a telephone number is:

```
telephone(String):- before(String, String1),
                    member(String1, TPList),
                    tstructure(String1),
                    after(String, String2),
                    member(String2, TFList)
```

Where “ T_{PList} ” and “ T_{FList} ” are the sets of strings we expect to precede and follow a telephone number respectively and “ $t_{structure}$ ” is the structure of a telephone number.

4.3 The Navigation Process

Websites contain collections of hypertext documents that are composed of nodes and links. The nodes represent documents and the links the relationships between documents. A node contains the information and a link allows the navigation of other documents of the hypertext collection. A link $(n_1; n_2)$ therefore represents a connection between the source node n_1 and the destination node n_2 [7]. Furthermore, we distinguish two types of links [6], referential links and semantic links. Referential links are used for a better organisation and easy reading of a document. In our approach, we give a more restricted definition of semantic links. A semantic link is a link that can be index by one or more words from a predefined set of keys. This restricted view of semantic links is used to target primarily those links that have a high probability of containing the information the system is looking for. Hence, improving the overall search time of the extraction process as a single node may contain hundreds of links. In addition to the source node and destination node already associated with a link, we now associate a list of indices for each link.

The indexation process starts with the tokenisation of the link name and target URL. Each token is then compared to a predefined list of indices. If the link name is textual we index both link name and target URL however, if the link is an image we just index the target URL. Again we have used the initial list of 50 websites to build the index lists. For example a link name “About us” will certainly give us the address of phone number of the company and “privacy policy” will link to the merchant privacy policy.

4.4 Evaluation of the Information Extraction System

Once the extraction and navigation rules have been defined and implemented, we used the same process as before to select another set of 100 commercial websites for the evaluation of the system. Table 1 summarises the performance of the system in terms of the precision with regards to the extracted variables.

Table 4.1: Precision of the information extraction system

	E1	E2	E3	A1	A2	A3	P1	P2	P3	F1	F2	F3
Actual	90	35	52	65	50	20	55	80	60	30	99	23
Extracted	75	20	30	20	35	10	25	75	42	12	65	13
Precision	83%	57%	58%	31%	70%	50%	45%	94%	70%	40%	66%	57%

The precision of the extraction system vary from 94% for the privacy statement variable to 31% for third party endorsement variable. The main reason for the low extraction of some variables is due to the fact that the information is conveyed through images and our system is currently not able to extract information from images. Similarly, the various payments methods are shown using various credit/debit cards symbols. However, their

number is very limited and they can be recognised by the extraction system. Users are required to manually check when the system fails to extract a particular variable.

5. TRUST MODEL EVALUATION

Defining an instrument to evaluate trust in eCommerce is a very difficult task. In our research, we have identified three approaches to evaluate a merchant trust based on the information we identified in the trust model and its presence or absence on the merchant website. These approaches are the linear approach, the parameterised approach and the fuzzy logic approach and are described in the next subsections.

5.1 The Linear Approach

This is a very simplistic approach and is best used with users with no or little experience with online shopping. We do not take into account their preferences and we allocate equal weights to all the variables of the model. If a variable is present on the merchant website and is positive (for example if the warranty policy is present and the customer is allowed to return the goods after a reasonable period of time), then the value 1 is assigned to the variable otherwise the value 0 is assigned. The total for each component is calculated and divided by 3 (the total number of variables). The total of all the components is divided by 4 (the number of components), to give us a value between 0 and 1 that will represent a percentage of trust. This is summarised in equation (1).

$$T = \frac{1}{4} \left(\sum_{i=1}^3 \frac{E_i}{3} + \sum_{i=1}^3 \frac{A_i}{3} + \sum_{i=1}^3 \frac{P_i}{3} + \sum_{i=1}^3 \frac{F_i}{3} \right) \quad (1)$$

For example if there are two variables in the existence component, two in the affiliation component, three in the policy and one in the fulfilment component then the trust factor is $T = 0.25(2/3+2/3+3/3+1/3) = 0.25(8/3) = 0.66$ or 66% which represents the trust factor for the merchant. The system is only processing this factor it does not provide any suggestion on whether to trust the merchant or not. The final decision is left with the user.

5.2 The Parameterised Approach

This approach is used with more experienced users. It is shown by the questionnaire that not all users agree with all the variables and their importance vary from one individual to another. For example, some may find that people existence is not important, portal fairly important and privacy policy very important. When the various variables are extracted, the customer is required to classify each variable as important, fairly important or not important. The system then assigns a weight for each variable. 1 if the variable is judged as important; 0.5 if fairly important and 0 if not important. The parameterised approach uses equation (2) to calculate the trust factor T which is again given as a percentage.

$$T = \frac{1}{4} \left(\frac{\sum_{i=1}^3 E_i w_i}{\sum_{i=1}^3 w_i} + \frac{\sum_{i=1}^3 A_i w_i}{\sum_{i=1}^3 w_i} + \frac{\sum_{i=1}^3 P_i w_i}{\sum_{i=1}^3 w_i} + \frac{\sum_{i=1}^3 F_i w_i}{\sum_{i=1}^3 w_i} \right) \quad (2)$$

Where n is the number of non zero components (i.e. if the customers chooses a whole component as not important, then that component is not counted)

5.3 The Fuzzy Logic Approach

This approach is an extension of the parameterised approach. We think that assuming that a variable can only be important, fairly important or not important is not sufficient. There are other parameters that need to be taken into account. The most obvious one is to define a degree of importance in general rather than the three ones identified in the parameterised approach. As mentioned earlier, there is also a risk that some of the information provided in a merchant's website may not be true. Hence, fuzzy logic is being investigated as a possible best fit approach as it takes into account the uncertainties within E-commerce data and like human relationships, trust is often expressed by linguistics terms rather than numerically. The fuzzy logic approach model and results are published in a separate paper [15].

6. CONCLUSIONS AND EVALAUTION

In this paper we presented an eCommerce trust model based on the information that increases customer trust if found on a merchant's website. The model is validated by the literature on eCommerce trust and a questionnaire. We have identified two major problems with this model. The first is guaranteeing the veracity of the information. Indeed it is widely know that some information found on eCommerce websites is not correct. For example it has been reported that many websites uses third party endorsements illegally[12], others do have a privacy policy but do not respect it and there is no guarantee that the comments found on some websites are from genuine customers. We have simplified an early model and kept only those variables that can be verified by other means such as calling the third party, getting in touch with customers or phoning the company. The second problem is the localisation of the information by the customer. To support this process, we developed an information extraction system and its evaluation was presented in section 4.4.

With regards to the current implementations of eCommerce website, we have based our evaluation on the same 100 websites sample used to evaluate the information extraction system. The distribution of the variables presence on this sample is shown in Figure 6.1. The average of this distribution is 55 and the standard deviation is 25.6. Some variables have a very high rate of presence on websites this include E1, physical existence and F2 payments methods. As one would expect, eCommerce website will always convey a payments methods and some information to support their physical existence such as a phone or fax number or a physical address. The variables with the lowest presence are A3, portal and F3, community comments. Many companies seem to miss to include these

variables on their websites although it is shown that they increase customers trust. Well known eCommerce websites such as Amazon and eBay have very good portals and community comments are an important part of their business.

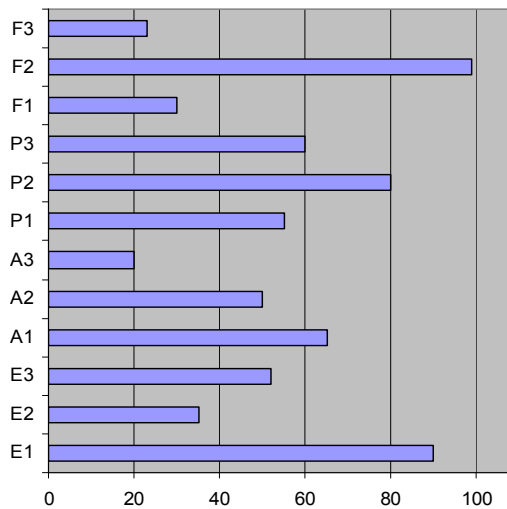


Figure 6.1: Variables distribution

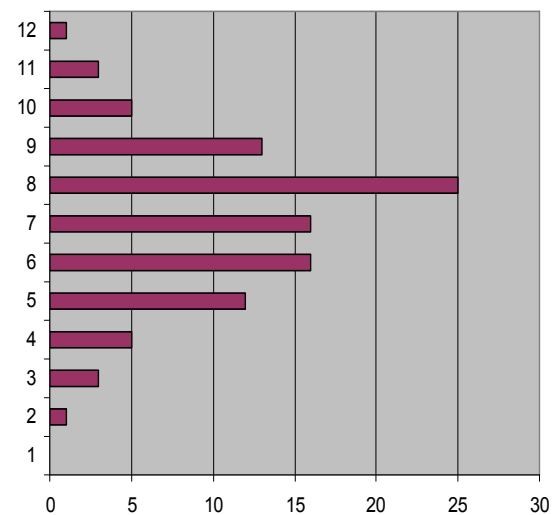


Figure 6.2: Cumulative variables distribution

In terms of the accumulation of these variables on websites, this is summarised in Figure 6.2 with an average of 8.3 and a standard deviation of 8. Most websites will have between 6 and 9 variables on their websites with an average of 8 which represents two thirds of the variables. There was only one website that has all the 12 variables and three with 11. A detailed analysis of the use of third party endorsement can be found in [12].

Another issue that has been noticed during the evaluation of these websites and the extraction system is the location of the variables on the websites. Some variables are very difficult to localise as they are deep in the website structure. The information extraction systems will spend a long time searching for the information which makes it inefficient and probably unusable. Similarly, if the search is performed manually, users will give up very quickly if they cannot find the information in the first few pages. However, 95% of all the variables are found in the first 4 levels of the websites. Hence, the extraction system stops searching after level 4. The extraction time after this level is around 3 minutes in average which start affecting the time efficiency of the system.

7. REFERENCES

- [1] I. Araujo and I. Araujo, Developing trust in internet commerce, *Proceedings of the 2003 IBM conference of the Centre for Advanced Studies on Collaborative research*, pp. 1-15, Toronto, Ontario, Canada, 2003
- [2] S. Ba, Establishing Online Trust through a Community Responsibility System, *Decision Support Systems*, (13):323-336, 2001.
- [3] Cheskin Research and Studio Archtype/sapient, ecommerce study trust. <http://www.studioarchtype.com/cheskin/assets/images/etrust.pdf>, 1999.

- [4] M. Craven, *et al.*, Learning to extract symbolic knowledge from the World Wide Web, *Proc. of the 15th Conference of the AAAI*, pp. 509–516, Madison, US, 1998.
- [5] F. N. Egger, Towards a Model of Trust for E-Commerce System Design, *Proc. of the CHI 2000 Workshop, Designing Interactive Systems for 1-to-1 E-commerce*, The Hague, The Netherlands, April 1-6, 2000
- [6] H. P. Frei and P. Schäuble, Designing a Hyper-media information system. In *DEXA '91*, pp. 449-454, Wien, 1991. Springer-Verlag.
- [7] H. P. Frei and D. Stieger. Making Use of Hypertext Links when Retrieving Information. In *Proc. ACM Conf. on Hypertext and Hypermedia*, pp.102-111, Milan, Italy, 1992.
- [8] K.S. Han and M.H. Noh, Critical Failure Factors that Discourage the Growth of Electronic Commerce, *Inter. Journal of Electronic Commerce*, 4(2):25-43, 1999.
- [9] D.L. Hoffman, T.P. Novak and P. Chatterjee, Commercial scenarios for the Web: opportunities and challenges, in *Readings in electronic commerce*, pp.29-53, Addison-Wesley Longman Publishing Co., Inc., 1997.
- [10] D.L. Hoffman, P.N. Thomas and M. Peralta; Building consumer trust online; *CACM* 42(4):80-85, 1999.
- [11] S. Jarvenpaa, N. Tractinsky, and M. Vitale. Consumer trust in an internet store. *Information Technology and Management Journal*, 1(1-2):45-47, 2000.
- [12] M. K. Kasiran and F. Meziane (2004), The Usage of Third Party Endorsement in Ecommerce Websites, *7th International Conference on Work with Computing Systems (WWCS2004)*, Kuala Lumpur, Malaysia, pp. 794-798
- [13] D.W. Manchala, E-Commerce Trust Metrics and Models, *IEEE Internet Computing*, March-April 2000, pp.36-44
- [14] F. Meziane and M.K. Kasiran, Extracting Unstructured Information From the WWW to Support Merchant Existence in E-Commerce, *Proceedings of NLDB03 International Conference*, pp.175-185, Cottbus, Germany, 2003.
- [15] S. Nefti, F. Meziane and M.K. Kasiran, A Fuzzy Trust Model for ECommerce, *7th IEEE International Conference on E-Commerce Technology*, July 19-22, München, Germany, 2005 to appear
- [16] National Fraud Information Centre, <http://www.fraud.org/2002intstats.htm>
- [17] A. Noteberg, E. Christiaanse, and P. Wallage, The role of trust and assurance service in electronic channels: An exploratory study, *Proceeding of the Information Industry Outlook Conference*, pp. 472-478, North Carolina, 1999.
- [18] E. Potter, WELL Topic Commercialization of the World Wide Web, in the *Internet conference on the Well*, November 16, 1994.
- [19] L. Xiong and L. Liu, A reputation-based trust model for peer-to-peer ecommerce communities, *Proceedings of the 4th ACM conference on Electronic commerce*, San Diego, CA, USA, Pages: 228 – 229, 2003.