Evaluating trust in electronic commerce : a study based on the information provided on merchants' websites

Meziane, F and Kasiran, MK

http://dx.doi.org/10.1057/palgrave.jors.2602430

Evaluating Trust in Electronic Commerce: A Study Based on the Information Provided on Merchants Websites.

Farid Meziane and Mohd Khairudin Kasiran
Informatics Research Institute,
University of Salford, Salford M5 4WT, UK

ABSTRACT

Lack of trust has been identified as a major problem hampering the growth of Electronic Commerce (EC). It is reported by many studies that a large number of online shoppers abandon their transactions because they do not trust the website when they are asked to provide personal information. To support trust, we developed an information framework model based on research on EC trust. The model is based on the information a consumer expects to find on an EC website and that is shown from the literature to increase his/her trust towards online merchants. An information extraction system is then developed to help the user find this information. In this paper, we present the development of the information extraction system and its validation. This is then followed by a study looking at the use of the identified variables on a sample of EC websites.

Keywords: Marketing Sale (MK); Credit and Risk Scoring (CR); Information Technology (IY); Data Mining (DA)

1 INTRODUCTION AND MOTIVATION

1.1 Introduction

New technologies have deeply modified traditional forms of social relations and communications, in particular norms, social rules, hierarchies, familiarity, reputation, delegation and trust [Castelfranchi and Pedone, 2003]. This is certainly true for Business-to-Consumers (B2C) Electronic Commerce (EC). For centuries consumers used to do

their shopping in shops and market places. They can communicate with the sale staff, see the shops interiors and looks and try, touch and smell the goods. Consumers may not conduct any risk evaluation at all, because shopping is a habit they do not perceive as risky [Riegelsberg and Sasse, 2001]. However, the new technologies and different communication media have created different shopping experiences.

B2C EC refers to consumers ordering products or services and paying for them through the Internet [Lim, 2003]. B2C EC has seen a phenomenal growth since the development of the internet and there is a growing interest from many organisations to use EC to improve their competitiveness and reach a wider customer base. Indeed, in EC, business transactions are no longer bound by geographical boundaries, time differences or distance barriers. Cazier et al. [2006], stated that within the well known 4 P's (Product, Price, Promotion and Place) marketing model, place has become irrelevant in EC and should be substituted by "Perception". Similarly, more consumers are adopting EC as it eliminates intermediaries, minimises the cost of the product and provides consumers with world wide market access [Gritzalis and Gritzalis, 2001]. Between 40 and 44% of internet users indicated online shopping as their primary activity [Pitkow and Kehoe, 1997;Center for the Digital Future, 2004].

However, there are many hindrance factors which cause EC to not reach its full potential and consumers lack of trust has often been identified as one of the main reasons for the disappointing development of B2C EC [Luo, 2002; Merrilees and Fry, 2003; Corbitt et al., 2003;Cazier et al, 2006]. This leads consumers to perceive the Web as a world of chaos, offering both opportunities and threats [Cheskin, 1999]. There are several critical failure factors that need to be addressed by the industry to ensure EC usage will continue to grow [Han and Noh, 1999]. Studies and reports by consumer associations, government organisations and academics identified some of these factors to include consumers dissatisfaction on the unstable EC systems, a low level of personal data security, disappointments with purchases such as non delivery of goods, hidden charges and difficulties in getting a refund and fraud [Han and Noh, 1999; Luo, 2002; Merrilees and Fry, 2003; Patton and Jøsang, 2004; Cazier et al, 2006]. These concerns are well justified as consumers' loss to Internet fraud has increased from US$3.2 millions in 1999 [Ba,

2001] to more than US$ 14.5 millions in 2002 [National Fraud Information Centre, 2002] and this continues to increase. It is generally perceived that it is relatively easy to set up a company in the digital world that appears legitimate but is actually a fraud [Ngai and Wat, 2002]. The question that many consumers 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 consumer. It is in this environment of risk and uncertainty that EC merchants must develop strategies for establishing trustworthiness, and that systems should be developed to assist consumers in assessing the level of trust they should place in an EC transaction [Patton and Jøsang, 2004].

1.2 Research Context

Trust is a very complex concept that received attention in several areas such as psychology, sociology, political science, economics, history and socio-biology [Lewicki and Bunker, 1996; Castelfranchi and Pedone, 2003]. There are also different types of trust that include, trust in EC, software and hardware systems, information sources, infrastructures and in authorities. It is not the intention of this paper to discuss in depth the concept of trust or compare its definitions and differences as viewed by different researchers, disciplines and communities. In the context of EC, trust is defined as the "willingness of a party to be vulnerable to the actions of another party based on the expectations that the other one will perform particular actions important to the trustor, irrespective of the ability to monitor or control the other party" [Mayer et al., 1995]. In other words, trust is the willingness of an individual to behave in a manner that assumes another party will behave in accordance with expectations in a risky situation. In EC the risks are higher and consumers are very vulnerable because:

- When consumers place an order on-line, they have to reveal sensitive personal and financial information such as address and credit card number [Cazier et al., 2006]
- In EC, there is typically a delay between the time of payment and the receipt of goods. [Cazier et al., 2006]

- Customers cannot physically interact face-to-face with a human representative, so they must rely on their trust in the organisation when making purchases [Chow and Holden,1997].
- Customers do not understand the underlying technology [Riegelsberg and Sasse, 2001].

Many studies have shown that trust is a key factor in stimulating internet purchasing especially at the early stages of the development of the merchant-customer relationship [Quelch and Klein, 1996; Jarvenpaa et al., 2000; Huang et al., 2003]. The Cheskin Research Trust Study (1999) describes trust as a dynamic process that deepens or retreats as a function of experience. Since trust is based on experience over time, establishing initial trust can be a major challenge to newcomers to EC, particularly those who do not have well established off-line brands [Pichler, 2000]. Once a merchant has developed a good reputation, trust is no longer a problem and consumers focus changes to brand, navigation and technology [Patton and Jøsang, 2004]. Typical examples of EC companies with very good reputations are eBay and Amazon. It is therefore imperative to not only identify and understand the factors that promote trust online but also to provide the new and inexperienced users with tools to help them check the availability of such information on the merchant's website and make sense of this information.

In fact, more than two-thirds of users (68%) say being able to identify information on a site is very important [Center for the Digital Future, 2004]. Given the incredible diversity of information online, users are looking for source identification to support their credibility judgments on sites. The ease and efficiency with which clients have access to relevant information in a Web site influence the customers' feeling of control on it [Araujo and Araujo, 2003]. Furthermore, a survey conducted by the Yankee Group on the reasons why shoppers abandon their carts, 29% gave the difficulty to navigate the website as the reason [Thomason, 2004].

In our research, we use the literature in EC trust to identify the main variables that are shown and proved to increase the consumers trust towards EC websites. A data mining system is then developed to extract and localise these variables on websites. We have

then attempted to evaluate the trust model based on the presence of these variables on a merchant website and the users views on the relevance and importance of these variables. Finally, we evaluated a sample of EC websites with regards to the set of variables identified to find out if these variables are widely used in EC websites.

The remainder of this paper is organised as follows: In section 2 we review the literature on trust in EC and we use it as the basis to develop our trust model that is developed in section 3. In section 4 we describe the development and implementation of the information extraction system and its evaluation. We present two different models to evaluate the trust confidence based on our information extraction system in section 5. In section 6 we present an evaluation of a sample of EC sites based on our trust model and we conclude in section 7.

2 ECOMMERCE TRUST

The concept of trust is becoming the driving force behind the design, evaluation and use of EC websites and is getting a lot of interest from many researchers [Corbitt et al., 2003; Cazier et al., 2006; Koo, 2006]. Many approaches have been used to understand and evaluate trust. Indeed, while browsing an EC website customers are faced by many uncertainties. Araujo and Araujo [2003] 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). Riegelsberg and Sasse [2001] classified risks related to EC into two categories. The first category comprises those risks that stem from the internet which include (a) whether credit card data gets intercepted; (b) whether the data is transmitted correctly; (c) their own interaction with the system. The second category concerns risks that are related to the physical absence of the online retailer and include (a) whether the personal details they supply will be passed on to other parties; (b) whether the online-vendor will actually deliver the products or services. Ceaparu et al.[2002], identified the key risks that customers associate with EC as (a) Business practices: to what extent will the online retailer deliver on its promises in terms of products, services and guarantees;

(b) Information protection: will private information given to the site be protected and will it only be used for the stated purpose; (c) Transaction Integrity: will the transaction be processed accurately and securely? Egger [2000] grouped the factors that are likely to influence the development and maintenance of trust in four groups. (a) Pre-interactional: reputation of the company, the strength of its brand and the customer's interaction history with the organisation; (b) Interface Properties: usability and structure of the organisation's website; (c) Informational Content: This is related to the information about products and services, company's history, values and commitments, privacy and statement and (d) Relationship Management: This looks at the communication and interaction facilities with the organisation

One of the most frequently cited concerns about online shopping is the security of monetary transactions. A recent survey showed that nearly half of the consumers expressed their fears about internet security [Center for the digital future, 2004] and this is strongly supported in the literature [Araujo and Araujo, 2003]. Ranganathan and Ganapathy [2002] reported in their study that among the factors that increases trust in conducting online transactions is the provision of alternate payment methods to online payments. In addition to security, the misuse of personal information is another serious concern for online shoppers [Anderson, 2000; Ranganathan and Ganapathy, 2002]. In a survey conducted by the Web Trends on the reasons why shoppers abandon their carts, 35% gave "the site requested too much information" as the reason [Thomason, 2004]. This is well justified as 72% of the sites surveyed by Anderson [2000] collected personal information but only 51% had a published privacy policy and only 28% of those sites notify their users about the specific personal information that is collected .

Reputation systems have also emerged as a method for fostering trust amongst strangers in EC environments. A reputation system gathers, distributes, and aggregates feedback about participants' behaviour. Resnick et al. [2000] state that these mechanisms can help people make decisions about who to trust and provide an incentive for honest behaviour. They may also have some influence on deterring dishonest parties from participating [Patton and Jøsang, 2004]. The first Web sites to introduce reputation schemes were on-line auction sites such as eBay. Xiong and Liu [2003] developed an adaptive trust model

for quantifying and comparing the trustworthiness of peers in Peer-to-Peer EC 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.

Interface web design and usability has also been found to influence user behaviour and trust towards EC websites [Basso et al., 2001, Riegelsberg and Sasse, 2001, Hu et al., 2004]. Web retailers use eye-catching graphics not only to grab a user's attention but also to convey competence or professionalism. Ease of navigation has also been found to be an important, perhaps necessary, antecedent to initial trust formation. Hu et al. [2004] have also stressed on the importance of taking into account the cultural background of the consumers when designing B2C EC websites. They argue that it is expected that people with different cultural backgrounds would respond differently to a globally generic website. There are also some factors that may influence the consumer decision making such as knowledge and experience of the use of the internet and the brand [Riegelsberg and Sasse, 2001]. .

3 THE PROPOSED TRUST MODEL

When shopping online, consumers search for information on risks and benefits and weight them against each other to reach a decision. Consumers have usually a number of questions on the shipping, service, payment and product return and policies [Ranganathan and Ganapathy, 2002]. The model we developed is based on the information present on the merchant websites that is shown from the literature reviewed above. However, the presence of the information alone is not sufficient. The veracity of the information is very important if one has to provide a valid 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 through a third party. The information trust model is summarised in Figure 1 and described in the next subsections.

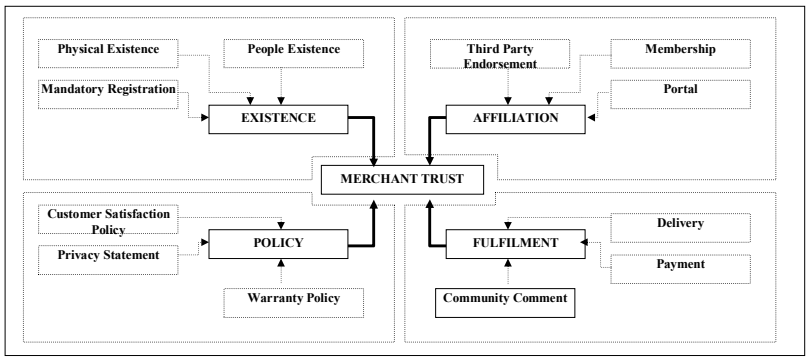

Figure1: The trust information model

3.1 The Existence Component

In EC the risk is greater due to the anonymity, distance and lack of physical interaction [Cazier et al., 2006]. In the brick-and-mortar world, customers can alleviate their concerns through face-to-face interaction with a human; the physical presence of the business offers assurance that it exists, is accessible and is trustworthy [Cazier et al., 2006]. Among the 51 web elements affecting peoples' perception of the credibility of a website identified by Fogg et al. [2001] is the inclusion of the physical address details of the organisation. Furthermore, among the recommendations made by the Nielsen user experience study [Nielsen, 2000] for communicating trustworthiness is providing the company's information that is easy to find. This is also confirmed by the study conducted by Araujo and Araujo [2003] who argued that in order to build trust an EC website should present information on the merchant background, contact details, performance history, associations, values, accomplishments, pictures, and so forth.

In EC, 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's validity. The variables retained for the existence model are: physical existence (E1), people existence (E2) and mandatory registration (E3). The physical existence variable (E1) is decomposed into address (E11), telephone number (E12) and fax number (E13).

3.2 The Affiliation Component

Trustmark seals when recognised, increase consumers perceptions of a site's trustworthiness [Cheskin, 1999; Gritzalis and Gritzalis, 2001]. A number of Trustmark seals have been developed to provide assurances about Web business practices and policies through the Web interface. One example is TRUSTe, which audits a site's stated privacy policies and allows sites to display the TRUSTe seal if privacy policies and disclosure meet specific standards[Patton and Jøsang, 2004]. This is particularly true for companies that do not have a reputation in the real world or that are new in the Internet arena where the use of third parties can provide assurance about their behaviour or about the quality of the products or services they recommend [Araujo and Araujo, 2003].

From the consumer's side, a strong trust relationship can be established with a vendor through direct experience. However, for new users, recommended trust can be used to establish the initial trust relationship [Noteberg et al.,1999]. Several possible methods of affiliation are used in EC and the most popular 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 merchants like Amazon and eBay. It is stated by Riegelsberg and Sasse [2001] that one of the leading advertisers on the internet is TRUSTe, an organisation that assigns seals to EC enterprises that is considered "trustworthy". 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 market. The variables retained for the affiliation component are: third party endorsement (A1), membership (A2) and portal (A3).

3.3 The Policy Component

Online privacy policy is understood as the set of statements explaining how consumers privacy is dealt with and protected by the web merchant. Public surveys indicate that privacy is the major concern for people using the Internet [Cavoukian and Crompton, 2000]. A study by the University of California has shown that 94.4% of Americans are concerned about the privacy of their personal information when buying online [2001]. Privacy related complaints that are made to the US Federal Trade Commission include complaints about unsolicited email, identity theft, harassing phone calls, and selling of data to third parties [Mithal, 2000]. Important requirements for EC security are the need to protect sensitive information that is stored on computers before and after an EC transaction, to verify the identity of the other party in the transaction, to ensure that no one can intercept the information being exchanged during the transaction, and in general to prevent disruption of services and applications [Patton and Jøsang, 2004].

In EC, policies such as privacy, customer satisfaction and guarantee can help consumers evaluate the trustworthiness of a merchant. These policies can influence the level of risk involved in the transaction. Merchant 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. The variables retained for the policy components are: customer satisfaction policy (P1), privacy statement (P2) and warranty policy (P3).

3.4 The Fulfilment Component

Online merchant needs 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 respective 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. About nine in ten online users want an explanation of when to expect delivery of goods or confirmation of reservations and a statement of the site's policies for returning unwanted items or cancelling [Patton and Jøsang, 2004]. The information that needs to be included is the delivery method, the company's name and order tracking method. Tracking the merchant's reputation is considered to be an antecedent of establishing a trusted environment towards the merchant [Jarvenpaa et al., 2000]. Reputation conveys 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 only useful if consumers can find it in a reasonable time. To increase the usability of the model, we have developed an information extraction system to help consumers localise the required information.

The entry point to the EC extraction system is the website's URL. The system attempts then to extract the variables of the trust model. Once the top level of the website is found, 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, physical address, peoples existence (names) and registration with other organisation. If any module fails to extract the required information then the links on the page are collected and navigation rules are applied to select the links to be used in the next iterations. The extracted information from all components is then stored in a database to be used to evaluate the trust factor associated with the merchant's website. The overall architecture of the existence module is shown in Figure 2.

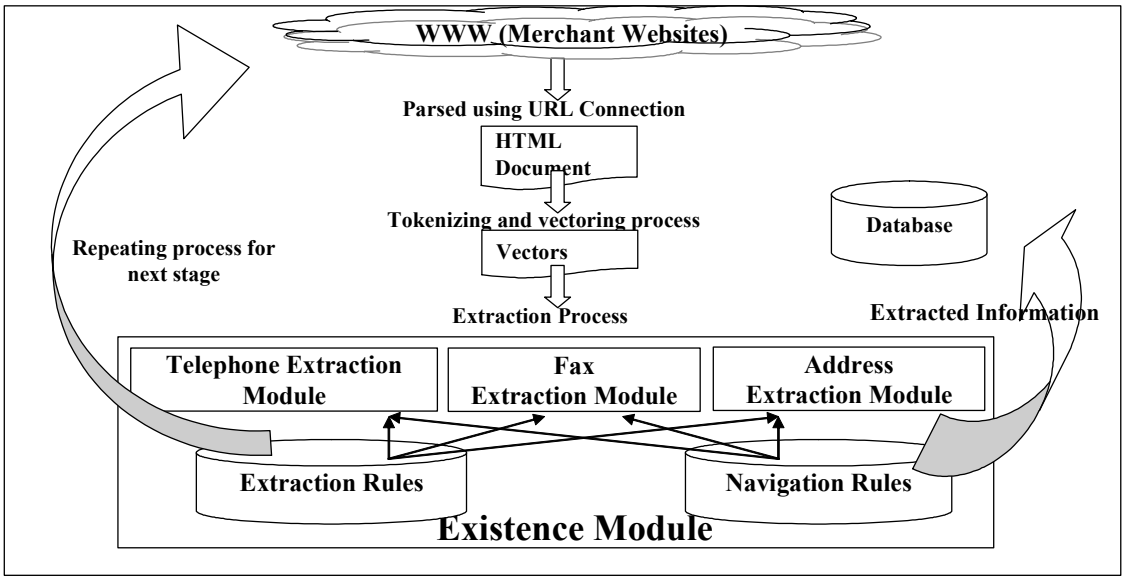

Figure 2: Extraction system's overall architecture

4.2 Extraction Rules

To develop the extraction rules, We have first randomly identified 50 websites using several ShoppingBots and requests were made to purchase few items such as books, digital cameras and chocolate. We have then hand crafted the extraction rules for the various factors defined in the trust model.

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 strings "call us at:", "Telephone" and "by phone" and the structure of the phone number is a numerical value. See [Meziane and Kasiran, 2003] for the detailed definition of the extraction rules for the existence component. An example rule for a telephone number is illustrated by:

```
telephone(String)= before(String, String1);
                   member(String1, TPList);
                   tstructure(String);
                   after(String, String2);
                   member(String2, TFList)
```

Which means that a string “String” is a telephone number if and only if there exists a string “String1” that comes before “String” and “String1” is a member of the set TPList, which contains the keywords that are know to precede a telephone number and “string” has one of the structures associated with a telephone number and there exists a string “String2” that comes after “String” and “String2” is a member of the set TFList that contains the keywords that are know to follow a telephone number.

4.3 The Navigation Process

Websites are abstracted as collections of hypertext documents that are composed of nodes and links including external links. The nodes represent documents, including multimedia 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 [Frei and Stieger, 1992]. Furthermore, we distinguish two types of links, referential links and semantic links [Frei and Schäuble, 1991]. Referential links are used for a better organisation and easy reading of a document. A semantic link is a link that can be indexed by one or more words from a predefined set of keywords and leads to a specific web page. Semantic links are used to primarily target those links that have a high probability of containing the information the system is attempting to extract. Hence, improving the overall search time of the extraction process as a single node may contain hundreds of links. For example if a link contains the string “Company Information”, then this will probably provide us with the details of the company including its phone and fax number and its physical address. In addition to the source and destination nodes already associated with a link, we now associate a list of indices for each link.

The link name and target URL are first tokenised (sentences split into single words). 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 initial 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

We evaluated the information extraction system using a different set of 100 websites selected randomly. Table 1 summarises the performance of the system in terms of the precision with regards to the extracted variables.

Table 1: Precision of the information extraction system

Components	Existence			Affiliation			Policy			Fulfilment		
Variables	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 varies from 94% for the privacy statement variable to 31% for third party endorsement variable. The main reason for the low precision 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. 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 EC is a very difficult task [Nefti and Meziane, 2004]. In our research, we have identified two approaches to evaluate 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 and the parameterised approach.

5.1 The Linear Approach

This approach is developed for consumers 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's 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 confidence. This is summarised in equation (1).

$$T = \frac{1}{4}\left(\sum_{i=1}^{3} \frac{E_i}{3} + \sum_{j=1}^{3} \frac{A_j}{3} + \sum_{k=1}^{3} \frac{P_k}{3} + \sum_{l=1}^{3} \frac{F_l}{3} \right) \quad (1)$$

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 confidence 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. For experienced users some variables are more important then others. 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 confidence factor T which is again given as a percentage.

$$T = \frac{1}{4}\left(\frac{\sum_{i=1}^{3} E_i w_i}{3} + \frac{\sum_{j=1}^{3} A_j w_j}{3} + \frac{\sum_{k=1}^{3} P_k w_k}{3} + \frac{\sum_{l=1}^{3} F_l w_l}{3}\right) \quad (2)$$

6 EVALUATION

With regards to the current implementations of EC websites, we have based our evaluation on the same sample of 100 websites used to evaluate the information extraction system. The distribution of the variables presence on this sample are summarised in Figure 3. The average of this distribution is 55 and the standard deviation is 25.6. Some variables have a very high rate of presence this include E1 (physical existence) and F2 (payments methods). As one would expect, EC websites will always convey a payments methods and some information to support their physical existence such as a phone or a fax number or a physical address. The variables with the lowest presence are A3 (portal) and F3 (community comments). Many companies seem to ignore these variables although it is shown that they increase consumers trust. Well known EC websites such as Amazon and eBay have very good portals and community comments are an important their reputation system is part of their business.

In terms of the accumulation of these variables on websites, this is summarised in Figure 4 which shows a normal distribution with an average of 8.3 and a standard deviation of 6. 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 [Kasiran and Meziane, 2004].

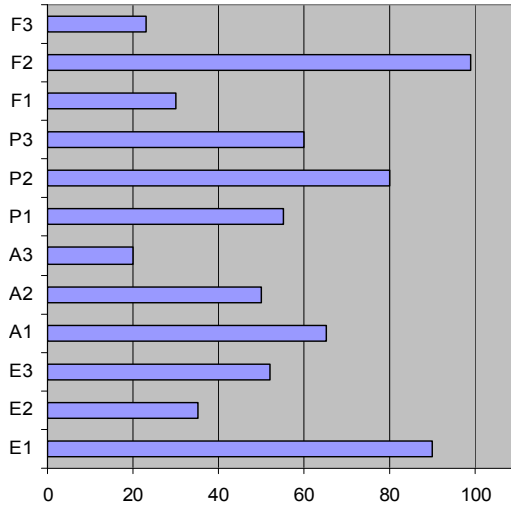

Figure 3: Variables distribution

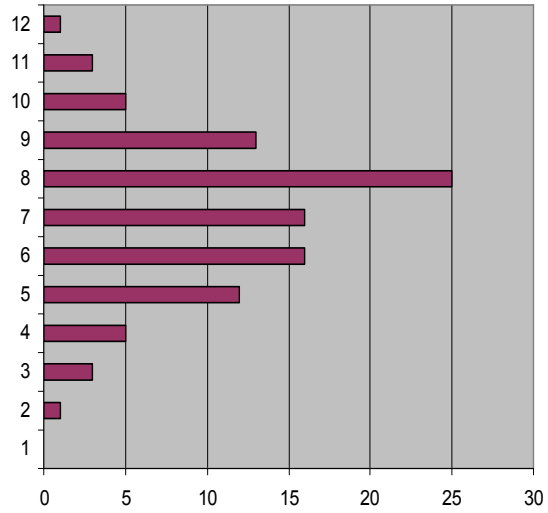

Figure 4: Cumulative variables distribution

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 find as they are deep in the website structure. This is inline with Nielsen's (2000) study that shows that only 42% of web users could successfully locate the information they wanted. The efficiency of the information extraction systems can be affected if it fails to quickly find the information. 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, our study has shown that 95% of all the variables are found in the first 4 levels of the websites. Hence, as a threshold between precision and efficiency, the extraction system stops searching after level 4. Searching after this level may affect the time efficiency of the system as the extraction time after level 4 can reach sometimes 20 minutes.

7 CONCLUSIONS AND FUTURE WORK

In this paper we presented an EC trust model based on the literature review on EC specifically on the information that is shown to increase customer trust if found on a merchant's website. We have identified two major problems with this model. The first is regarding the veracity of the information. Indeed it is widely know that some information

found on EC websites is not correct. For example it has been reported that many websites use third party endorsements illegally [Kasiran and Meziane, 2004], 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 only kept in our model those variables that can be verified by other objective means such as calling a third party, getting in touch with customers or phoning the company. The second problem is the finding of the information by the consumer. To support this process, we developed an information extraction system. Making this information available to the user without the effort of searching for it, will remind customers about the dangers of EC and hopefully make them think before engaging in a transaction with unknown merchants.

The ideas presented in this research are being taken by some commercial organisations. Verisign (www.verisign.co.uk), the well know company that provides third party endorsement for EC sites with regards to security, is introducing what they label as the new generation of browsers. The browser they are developing is able to recognize if the website a user is looking at is endorsed by Verisign. This research might be extended as a plug-in to current web browsers running in the background and providing the collected information on users request. However, before attaining this stage, the precision of the current information extraction system needs to be improved. As more and more websites are conveying some information using images, the system needs to be extended by a module able to identify information from images.

REFERENCES

Anderson A (2000), Internet Privacy survey 2000.

Araujo I and Araujo I (2003), Developing trust in internet commerce, *Proc. of the IBM conference of the Centre for Advanced Studies on Collaborative Research*, pp. 1-15, Toronto, Canada.

Ba S (2001), Establishing Online Trust through a community responsibility system. *Decision Support Systems* (13):323-336.

Basso A, Goldberg D, Greenspan S and Weimer D (2001), First impressions: emotional and cognitive factors underlying judgments of trust e-commerce, *Proc. of the 3rd ACM conference on Electronic Commerce*, p.137-143, Tampa, Florida, USA

Castelfranchi C and Pedone R (2003), A review of trust in information technology, *the ALFEBIITE Project*, http://alfebiite.ee.ic.ac.uk/docs/papers/D1/ab-d1-cas+ped-trust.pdf

Cavoukian A and M Crompton (2000), Web Seals: A Review of Online Privacy Programs, http://www.ipc.on.ca/english/pubpres/papers/seals.pdf.

Cazier J A, Shao B B M and Louis R D St (2006), E-Business differentiation through value-based trust, *Information and Management*, to Appear.

Ceaparu I, Demner D, Hung E, Zhao H and Shneiderman B (2002), In Web we trust, Establishing strategic trust among online customers, in Rust, R T and Kannan P K (Eds), *e-Services: New directions in theory and practice*, Armonk New York, pp. 90-107.

Center for the digital future (2004), USC Annenberg School, The digital future report, http://www.digitalcenter.org/downloads/DigitalFutureReport-Year4-2004.pdf

Cheskin Research & Studio Archetype (1999), eCommerce Trust Study, http://www.cheskin.com/docs/sites/1/report-eComm Trust1999.pdf

Chow S and Holden R (1997), Towards an understanding of loyalty: the moderating role of trust, *Journal of Managerial Issues* 9(3), pp. 275-299.

Corbitt B J, Thanasankit T, and Yi H.(2003), Trust and e-commerce: A study of consumer perceptions, *Electronic Commerce Research and Applications* (2): 203-215.

Egger F N (2000), Towards a model of trust for E-Commerce system design. *Proc. of the CHI 2000 Workshop, Designing Interactive Systems for 1to1 E-commerce*, The Hague, The Netherlands, April 1-6.

Fogg, B et al. (2000), What makes Web sites credible? A report on a large quantitative study, *in: Proc. of CHI 2001*, pp. 61–68, ACM Press.

Frei H P and Schäuble P (1991), Designing a hyper-media information system, in *DEXA'91*, pp. 449-454, Wien, Springer-Verlag.

Frei H P and Stieger D (1992), Making use of hypertext links when retrieving information. *In Proc. of the ACM Conference on Hypertext and Hypermedia*, pp.102-111, Milan, Italy.

Gritzalis S and Gritzalis D (2001), A digital seal solution for deploying trust on commercial transactions, *Information Management and Computer Security*, 9(2)71-79.

Han K S and Noh M H (1999), Critical failure factors that discourage the growth of electronic commerce. *International Journal of Electronic Commerce*, 4(2):25-43.

Hu J, Shima K, Oehlmann R, Zhao J, Takemura Y and Matsumoto K (2004), An empirical study of audience impressions of B2C web pages in Japan, China and the UK, *Electronic Commerce Research and Applications*, (3):176-189.

Kasiran M K and Meziane F (2004), The usage of third party endorsement in Ecommerce websites, *7th Intern. Conference on Work with Computing Systems*, Kuala Lumpur, Malaysia, pp. 794-798.

Koo D M (2006), The fundamental reasons of e-consumers' loyalty to an online store, *Electronic Commerce Research and Applications*, (5):117-130.

Lewicki R J and Bunker B B (1996), Developing and maintaining trust in work relationships, in Kramer, R M and Tyler T R (Eds), *Trust in Organizations: Frontiers of Theory and Research*, Sage, Thousand Oaks, CA, pp. 114-39.

Lim, N.(2003), Consumers' perceived risk: sources versus consequences, *Electronic Commerce Research and Applications* (2):216–228.

Luo X (2002), Trust production and privacy concerns on the internet: a framework based on relationship marketing and social exchange theory, *industrial marketing management*, 31(2):111-118.

Mayer R C, Davis J H and Schoorman F D (1995), An integrative model of organisational trust, *Academy of Management Review*, 30(3):709-734.

Merrilees B and Fry M (2003), E-Trust: the influence of perceived interactivity on e-tailing users, *Marketing Intelligence and Planning*, 21(2):123-128.

Meziane F and Kasiran M K (2003), Extracting unstructured information from the WWW to support merchant existence in E-Commerce, *Proc. of the international conference on application of natural language to information systems*, pp.175-185, Cottbus, Germany.

Mithal M (2000), Illustrating B2C Complaints in the Online Environment, the Joint Conference of the OECD, HCOPIL, ICC: Building Trust in the Online Environment: Business to Consumer Dispute Resolution, The Hague, http://www1.oecd.org/dsti/sti/it/secur/act/online_trust/presentations.htm.

National Fraud Information Centre, 2002, http://www.fraud.org/2002intstats.htm.

Nefti S, Meziane F and Kasiran M K (2005), A fuzzy trust model for ECommerce, *7th IEEE International Conference on E-Commerce Technology*, July 19-22, München, Germany, pp. 401-404, IEEE Press.

Ngai E W T and Wat F K T (2002), A literature review and classification of electronic commerce research, *Information and Management*, Number 39 pp. 415-429.

Nielsen J R, Molich C S and Farrell S (2000), E-commerce User Experience, Technical report, Nielsen Norman Group.

Noteberg A, Christiaanse E and Wallage P (1999), The role of trust and assurance service in electronic channels: An exploratory study. *Proc. of the Information Industry Outlook Conference*, pp. 472-478, North Carolina.

Patton M A and Jøsang A (2004), Technologies for Trust in Electronic Commerce, *Electronic Commerce Research*, (4):9–21.

Pichler R (2000), Trust and Reliance - Enforcement and Compliance: Enhancing Consumer Confidence in the Electronic Marketplace. *Stanford Law School*, www.oecd.org/dsti/sti/it/secur/act/online_trust/Consumer_Confidence.pdf.

Pitkow J and Kehoe C (1997), 7th WWW User Survey. *Georgia Tech Research Corp*., www.gvu.gatech. edu/user_surveys/.

Ranganathan C and Ganapathy S (2002), Key dimensions of business-to-consumer web sites, *Information and Management*, (39):457-465.

Resnick P, Zeckhauser R, Friedman E and Kuwabara K (2000), Reputation Systems, *CACM,* 43(12), pp: 45–48.

Riegelsberger J and Sasse M A (2001). Trustbuilders and trustbusters: The role of trust cues in interfaces to e-commerce applications, Proc. of the 1st IFIP Conference on e-commerce, e-business, e-government, pp. 17-30. Kluwer.

Thomason L (2004), Keep Customers With Their Carts, Netmechanic 7(6), http://www.netmechanic.com/news/vol7/ecommerce_no6.htm

University of California-Los Angeles Centre for Communications policy (2001), the UCLA Internet report 2001: Surveying the digital future.

Xiong L and Liu L (2003), A reputation-based trust model for peer-to-peer ecommerce communities, *Proc. of the 4th ACM conference on Electronic commerce*, San Diego, CA, USA, pp. 228 – 229.