Incorporating Prejudice into Trust Models to Reduce Network Overload (May 2005)

M. Wojcik[1], H.S. Venter[2], J.H.P. Eloff[3], M.S. Olivier[4]

hibiki[1]@tuks.co.za
{hventer[2],eloff[3],molivier[4]}@cs.up.ac.za

Information and Computer Security Architectures Research Group
(ICSA)
Department of Computer Science
University of Pretoria

Abstract— Trust and trust models have invoked a wide interest in the field of computer science. Trust models are seen as the solution to interactions between agents (computer systems) that may not have previously interacted with one another; as is often the case in the uncertain world of e-commerce. These models are seen as facilitators to the definition and development of interaction between two such agents. Trust models rely heavily on a knowledge-building process to evaluate the value, or in some instances to become aware of the risk of communicating with another agent. The observation of other agents or the sharing of knowledge between agents accomplishes this. Thus, trust models rely heavily on the flow of information between machines.

The problem this paper addresses is: How do we lessen the number of communications a single agent has to deal with in order to allow the agent to have sufficient time and resources to accurately analyse these interactions? The suggested solution involves adding a prejudice filter to current trust models. This paper investigates the value of reducing network overload by limiting communication through prejudice and suggests possible filtering factors that can be used in such a scenario. These factors are based on existing security and trust implementations in order to simplify the incorporation of prejudice into current trust models and trust model architecture.

Index Terms— category, certificate, domain, intermediary, learning, organization, prejudice, policy, recommendation, trust, trust models.

I. Introduction

The world of e-commerce is a vast dynamic domain often requiring 'virtual' businesses to communicate and establish contracts in an environment where changes in business customs can be made almost in an instant [1], thus, bringing in the need to define whom one can trust and more specifically how one trusts [2], [3].

Determining and defining whom an agent trusts and to what degree is the core of trust models. In this paper the term agent refers to a computer within an e-commerce environment using a trust model to determine trust. Trust in the world of Computer Science can be defined as an agent's belief in the dependability and capability of another agent; the value more often than not is a result of experiences, observations and/or recommendations [5].

Trust models rely on the collection and analysis of information to form trust opinions. This leads to the interesting problem of how is one to filter out unwanted flooding of communications all vying for analysis. This paper attempts to solve this problem by introducing a prejudice filter to lighten the load of information any agent needs to deal with in order to establish a trust relationship. The paper further investigates how this prejudice filter can be incorporated into trust principles already in existence.

The remainder of the paper is structured as follows. Section II is the background to the paper and serves to define trust, prejudice and trust models. Thereafter, section III investigates where and how prejudice can be implemented within trust models. Finally section IV concludes the paper.

II. Background

Concepts dealt with by this paper include trust, trust models, and prejudice. A clear understanding of these concepts is required in order to understand the aims of this paper. This section gives a broad overview of each of these concepts.

This material is based upon work supported by the National Research Foundation under Grant number 2054024. Any opinion, findings and conclusions or recommendations expressed in this material are those of the author(s) and therefore the NRF does not accept any liability thereto.

A. Trust

Trust is a subjective concept unique to each individual and each individual's worldview. It is often dynamic in nature and influenced by environment, state and situation. Nooteboom [6] defines trust as a four-place predicate stating that: "Someone has trust in something, in some respect and under some conditions." The four predicates mentioned here are: the entity trusting (someone), the entity being trusted (something), the reason and goals that define the need for trust (respect) and the conditions under which the trust is given (conditions). Thus trust involves risk.

Trust can be directed to individuals, institutions, organizations as well as socio-economic systems. Trust in systems can result in individual trust where the trust an individual has in another individual is a direct result of the trust the individual has in the organization to which the other individual belongs [6].

B. Trust Models

Trust and trust models are an area of interest in the field of Computer Science resulting in the formation of varied machine trust models. Numerous trust models have been studied in an attempt to define a common set of features in order to give a guideline as to what aspects are required in such a trust model [2], [7], [8], [9], [10], [11], [12], [13], [14]. Such an initial set of features common to most trust models as determined from these sources includes the following:

Recommendation of trust: Trust is built by an agent based on the recommendations it receives from other agents.

Dynamic trust: Due to the dynamic ever-changing nature of the interactions an agent has to deal with, trust requires to be continually updated.

Evaluation and ranking of trust: This involves the codification of trust whereby trust levels are given explicit values that can be translated into machine code.

Trust in policies: The policies an agent adheres to define the community an agent belongs to, defining how an agent chooses to build and evaluate trust relationships.

Delegation of trust: A form of recommendation where agents are able to delegate certain rights to other agents.

Webs of trust: Trust is propagated throughout the network in order to create webs of trust to be used by agents in order to accomplish given tasks.

Situation and trust: Trust is highly dependent on the context in which an interaction takes place.

Examples of work done in this field include trust evolution and trust update functions defined by Maarten Marx and Jan Treur [15] and a 'soft security' distributed model defined by Alfarez Abdul-Rahman and Stephen [8]. Maarten defines a trust model that defines and updates trust based on past interactions it had with an agent. While Alfarez Abdul-Rahman and Stephen [8] define a distributed trust model based on the assumption that trust is transitive, and can thus be propagated throughout the system via interaction between agents and define a Recommender Protocol to build social webs of trust.

Due to the dynamic nature of agents, trust requires the continual re-evaluation of the above-defined features. A filtering mechanism can contribute significantly to minimise the amount of workload required to do trust formation. This paper proposes such a mechanism based on prejudice discussed in section III.

Current trust models rely on flooding as an information-gathering phase [16], propagating trust through intricate communication, thus opening the doors to network overload [7]. This is further complicated by allowing agents to take proactive actions towards certain goals as a result of changes in the environment [9]. Prejudice allows trust models to filter through and simplify numerous and complex interactions, lessening the communicative load [15].

C. Prejudice

Prejudice can be defined as a negative attitude towards an entity based on stereotype, placing all entities of a certain stereotyped group into the same category [17] and is used to simplify initially a complex interactions.

Prejudice makes use of categorization. Categorization assumes that a group possesses either assumed or imagined characteristics that place them in a particular category. This allows the individual to respond to the group members based on their membership to a specific category rather than on their individual uniqueness.

Prejudice is influenced by culture. A form of cultural prejudice is that of institutional prejudice whereby assumptions are institutionalised. Forms of institutionalising assumptions include policies and practices [17].

III. The Incorporation of Prejudice into Machine Trust Models

An agent's primary goal is to minimise the risk inherent during communications with other agents. To accomplish this, agents need some method of determining the risk involved during communication with the other agents. Trust models have been proposed as a solution to this dilemma. However, as discussed in section II.B above, trust models rely on information gathering as a primary means of trust formation, that require several messages, carrying the required information, to travel through the network. This leads to network overload in an environment where the number of potentially new, previously unevaluated communications that an agent has to deal with, are vast.

The proposed solution investigates how prejudice filtering can be used to minimise the number of messages that need to travel across the network in order for an agent to successfully formulate trust. Prejudice filters need to limit the number of communications an agent needs to deal with to allow an sufficient time and resources to properly analyse the incoming communications and evaluate trust.

The goal of this paper is to extend the initial set of features described in section II.B by adding prejudice filters to the features. Nine features were explicitly chosen for investigation. These nine features were chosen due to the fact that these are principles already in use by trust models thus simplifying the incorporation of the prejudice filters. Each feature can thus be incorporated into a trust models based on any of these principles. Each feature will be best incorporated into a trust model if it extends the core principle a trust model relies on. For instance a recommendation based trust model could be extended to include recommendation based prejudice filters. The rest of this paper investigates how this may be achieved.

1. Prejudice and intermediaries

Prejudice can be incorporated into trust models by allowing a trusted intermediary to keep a list of agents that are to be trusted. This extends the concept of trusted intermediaries such as Certificate Authorities [1]. However, the intermediaries in this case, are used solely for the collection, grouping, definition and categorization of trust-related data [9]. Prejudice filters can be implemented by allowing an agent to trust 'only' other agents that are trusted by particular intermediaries. Figure 1 illustrates this concept. The intermediary that agent A trusts, trusts only Agents F, D and B, therefore, A will also only trust F, D and B. Searching for trusted agents to communicate with is done through the intermediary leaving the agent with a lighter communication load. This also has a positive affect on network overload due to the fact that all information is gathered by a single intermediary instead of having to flow to all the nodes existing within a network.

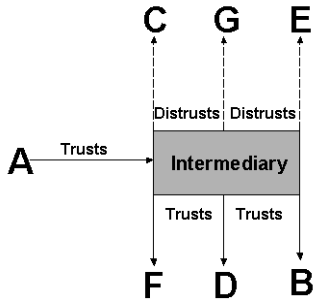

Fig1: Implementing Prejudice by means of an intermediary

2. Prejudice and categorization

Different levels of trust can be supplied by segregating agents into categories and levels. Here the agent makes assumptions about the agents it is to interact with and what privileges it will grant. It then evaluates the interaction based on those assumptions [2].

Assumptions lead to categorization. Allowing an agent to make assumptions based on categories reduces the amount of information that needs to be shared between two agents in order to participate in a transaction. This alternately reduces the number of messages and communications travelling across the network.

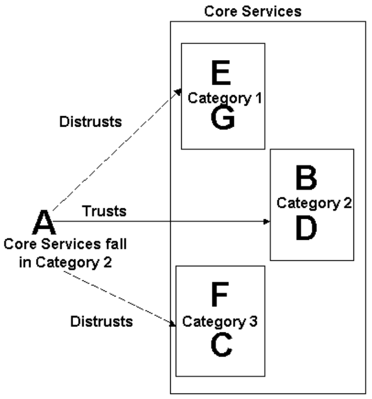

Fig2: Prejudice and Categories

Categorization is a simple way to implement prejudice and its concepts. Agents can be categorized according to their core services, products and policies. Each of these defines the fundamentals of businesses on the web. An agent can filter out and refuse to communicate with other agents that do not adhere to the same core principles [1]. For example, agents can be grouped into categories depending on their core services. An agent will then only trust agents that fall in the same category as it does as shown in figure 2.

Categorization can be further extended, allowing access rights to be divided into 'trust categories'. This allows an agent to define levels of trust by defining the category of an interaction. Thus the rights delegated to an interaction are limited by the category and assumption to which the category implies [8].

3. Prejudice and policy

Trust models tend to approach the issue of trust from a generic point and often neglect the fact that agents interact with one another according to the policies and rules defined by the environment or community to which they belong. It is important to define the conventions and prejudices that develop in such communities and use it to evaluate possible interactions with other agents [11].

Fundamental principles followed by communities refer to three basic sources of expectations. These include general rules shared by all agents; social rules that the agent's belonging to a specific group share and the institutional norms, which are defined and enforced by the institution within which the negotiating agents interact [11]. These rules define agent's expectations and include concepts such as privacy policies, what information is used for, encryption policies, transaction contexts as well as other norms used by agents during interactions.

To form a common community agents share their local policies through a process of mutual recursion. Mutual recursion relies on agents sharing one another's local policies and merging them to form a common community policy [12]. The policies are propagated throughout the network in a recursive manner ensuring all agents within the community receive the integration of these local policies.

There is an initial network communication load inherent in this approach during the establishment of a common local policy. However, once this shared community local policy has been established, less communication is required on the network simply due to the fact that any agent belonging to the community can report on and abide by the rules the community has defined. The need to analyse each individual agent from a specific community is removed and any agent wishing to communicate with multiple agents coming from a specific community, can safely assume that the policy held by all agents in the community is the same.

Prejudice filters are a good way to avoid policy misunderstandings simply by allowing the agent to disregard other agents with policies that differ vastly from its own. Possible implementations include dividing agents into world zones. This is possible due to the fact that world zones define varying cultural beliefs. These cultural beliefs tend to dominate the business world in which an agent resides. Examples of this include the varying cultural values between Asian and Western culture. Asian cultures emphasize the community while Western cultures tend to put more emphasis on individualism and public reputation [8].

It is important to define whom the agent you wish to

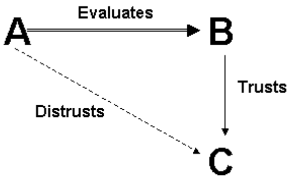

Fig3: Prejudice using Policy

communicate with, trusts. For example, as shown in figure 3, Agent A is in the process of evaluating Agent B to determine whether it trusts B as well as to analyse whom B trusts. A requests B to send a list of the agents that B trusts and analyses this list comparing it to its own. A, knows that it distrusts Agent C. A sees from the list B has supplied that B trusts C. A will give B less trust based on the fact that B trusts C. The reason A will give less trust is because A may not wish to disclose information to B that B may potentially share with C. This interaction results in A being prejudiced against all agents that trust C whether the agents that trust C would disclose information gained from Agent A to C or not [16].

4. *Prejudice and organizational certificates*

There are various mechanisms for identifying other agents one communicates with. Digital signatures vouch for people; computer addresses identify computers; and organizations represent themselves with signed certificates binding together groups of people and computer addresses [2].

These digital signatures and signed certificates can be used to implement forms of 'institution' prejudice [5]. An agent can filter out addresses and digital signatures that do not belong to an organization that it trusts. This is a form of categorization where agents use organizations to determine which agents they communicate with, communicating with only agents that belong to approved organizations [2]. This extends distrust to agents that do not belong to any organization. Using this logic it is clear that Agent A in figure 4 will trust E and G, due to the fact that it trusts organization 1, but distrust B, D, F and C.

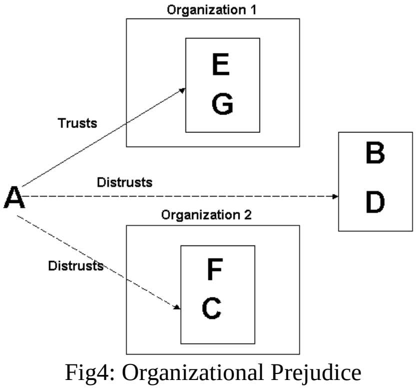

Fig4: Organizational Prejudice

Network traffic is minimised as agents only communicate with and analyse communications coming from agents that belong to trusted agents, ignoring and filtering out all other incoming traffic.

5. *Prejudice and roles*

Roles can be considered as an extension of categorization and are investigated here to see how they can be used to implement prejudice filters.

Using roles is a form of categorization that groups agents into roles and assigns to them the privileges associated with specific roles. Categorizing agents into roles makes prejudiced assumptions about the roles they play, allowing for an agent to differentiate between levels of trust given to other agents. For example, an agent wishing to act as a customer will not be given access to administrative information.

The grouping of agents into roles allows a business agent to define standard actions and privileges to a particular role instead of worrying about defining unique access to each individual agent it comes into interaction with. This simplifies the grey areas that emerge when a trust model evaluates a trust value and attempts to define trust levels. Agents that have met the criteria of a role get assigned the role for the duration of the transaction and get all the privileges and access rights associated with that role only.

Roles can be defined as shown in figure 5. Agent B requests a transaction. This instigates Agent A's analysis process. A first analysis the role into which a specific transaction would requires B to play order to succeed. A then checks if it trusts B to take on the required role. If A's level of trust in B is equal to or higher than that required by the role, A assigns the required role and all privileges to B. The communication between A and B is then limited by the constraints of the role. The limiting of reduces the number of messages that need to pass between agents lowering network traffic.

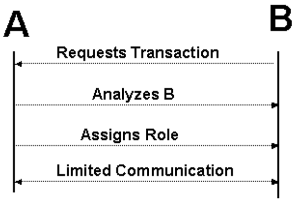

Fig5: Process of establishing role based prejudice

6. *Prejudice and domain*

One way of reducing the number of interactions an agent needs to deal, with is to allow it to only communicate with agents that fall into a limited address range, delegating communication with agents from a different range to another agents within its domain. Prejudice in a domain can be seen as an extension of organisational prejudice. However, the key difference here would be that domain prejudice can be implemented within an organization itself as a feature to limit network traffic.

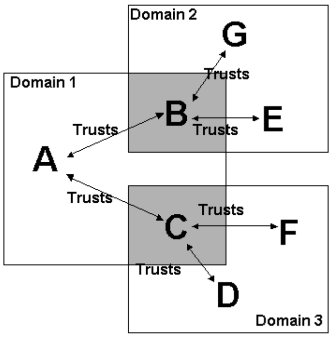

Fig6: Domain based prejudice

This allows each agent to have a limited communication circle even within its own domain, but can still communicate indirectly with agents outside it's domain by using one of the agents within its domain as a communication intermediary. The limited communication circle reduces the number of messages that flow within the network. Figure 6 illustrates this scenario. A communicates only with B and C. If G wishes to communicate with A, it does so through B, using B as an intermediary. The principle here is that G and E need to move through B to communicate with A while F and D need to move through C.

Here roles are used by an agent to indicate whether it is acting on its own accord or on behalf of another thus preventing another agent from losing trust in a particular agent simply because interactions have failed while the agent was acting as an intermediary [10].

7. *Prejudice and path*

Path length is already a form of prejudice filter used by agents in a network. This is due to the fact that extremely long paths are seen as untrustworthy. This is not necessarily the case, but this form of assumption allows an agent to cut down on interactions it has to evaluate and at the same time allow it to lower its risk and the network load [16].

Trust in another agent is highly influenced by the path of communication between two agents. The more secure the path, the safer the transaction and, therefore, the higher the possible level of trust. Thus, trust between two agents is also influenced by the trust these two agents have in agents that connect them along the path of communication. An agent has the right to refuse communication with another based on the fact that the path of communication passes through another agent that the agent in question does not trust [13]. This path can be traced and discovered using the tracert protocol to discover the path a message takes [19]. An agent thus refuses to communicate along paths that contain untrustworthy agents in a prejudiced manner.

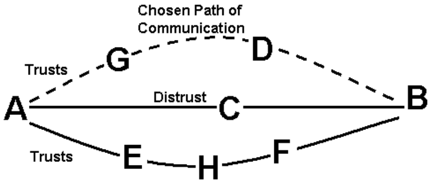

Fig7: Prejudice and path

Prejudice and path length are illustrated by figure 7. Under normal network conditions without the inclusion of trust, communication between A and B would pass through C due to the fact that, that is the path that contains the least nodes between them. However, due to the fact that either A or B distrust C, this path cannot be used. Therefore, the path with the next lowest number of nodes between the two desired points of interaction is chosen, provided that both A and B trust all the agents along that path. The path chosen in figure 7 passes through G and D since there is no distrust indicated along that path.

8. *Prejudice and recommendation*

Trust models rely heavily on 'soft security' where the key to managing and defining security is that of 'social control'. A way to manage trust often relies on a 'community of agents'. An agent trusts another agent as well as having trust values for the trust it has on the other agent's recommendations. If A trusts B and propagates that trust to C, many trust models allow C to also trust B provided C also trusts A as shown in figure 8 below. If A discovers that it no longer trusts B, it is A's responsibility to propagate the change in trust onto C. C then uses this change in information as well its own information on its experiences in interactions with both A and B to recalculate its trust values [8].

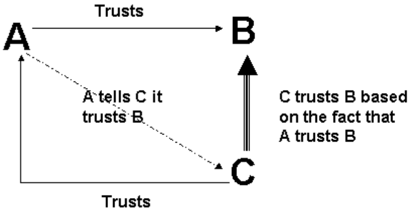

Fig8: Trust and recommendation

Intermediaries discussed in section III.1 as well as other trust models rely on this principle of recommendation to pass on trust values. Trust models that do not use intermediaries use a more social approach where an agent relies on its peers to recommend possible interactions.

Communication and levels thereof between agents can be delegated. If an agent requests a service, the agent the service was requested from, can delegate the responsibility of the service to another trusted agent. This is a form of a recommendation trust model where one agent recommends another. Prejudice can be implemented in a manner that allows an agent to only trust recommendations from a fixed set of delegating agents [16].

Network load is minimised due to the fact that an agent simply does not communicate with any agents that are outside of the community's recommendation circle.

9. *Prejudice and learning*

Prejudice doesn't necessarily need to be explicitly defined. An agent can learn it. To accomplish this, an agent relies on 'first impressions'. The agent defines a category for the new agent it comes into contact with via an information-gathering process. The agent then proceeds to attempt a transaction with the other agent in question.

If this transaction fails, the agent tags certain general information given by the agent the interaction failed with, for instance, which organization the agent represents, and refuses further transactions from agents from that category.

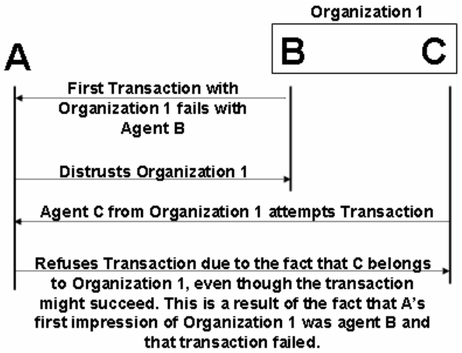
Organization 1
A
B
C
First Transaction with Organization 1 fails with Agent B
Distrusts Organization 1
Agent C from Organization 1 attempts Transaction
Refuses Transaction due to the fact that C belongs to Organization 1, even though the transaction might succeed. This is a result of the fact that A's first impression of Organization 1 was agent B and that transaction failed.

Fig9: Establishing prejudice through first impressions

Figure 9 demonstrates this by means of a sequence diagram. As shown A refuses any transactions from Organization 1 simply based on the fact that its first impression of Organization 1 came from B with whom A's transaction failed. A does not care that interactions with other agents from Organization 1 might succeed. If the transaction succeeds, a trust relationship is formulated. It is important to note that the first impression is the key one. If the failure in transaction occurs at a later stage it does not terminate the agreement or the relationship between agents but simply lowers the trust value thereof [8].

Network load is thus gradually reduced as an agent learns prejudice.

IV. Conclusion

This paper has introduced the concept of prejudice and investigated how it may be incorporated into current trust models to minimise the interaction load among agents. Means of incorporating prejudice into current infrastructure that have been investigated include intermediaries, categories, policies, certificates, roles, domains, path definitions, recommendations and learning. Each of these areas explore how various measures that are already in use, can be used as means of filtering out the overload of communication on a network.

This paper provided a conceptual discussion, which requires further investigation regarding implementation issues. The initial set of features for machine trust models defined for the purpose of this paper needs to be explored in more detail and should be made more concrete to be used as a guideline for future trust models, as well as for the evaluation of current trust models defined.

More in-depth work needs to be done on protocols for the definition and interpretation of the features discussed, which are required by trust models. These concepts are required for the evaluation of trust. It is hoped that we can move towards some kind of data representation standardization that would allow various agents to easily understand each other and find the information they are looking for when working towards trust formation.

Future work

Future work will include experimental implementation and evaluation of the concepts discussed. Example applications will be designed, allowing the resulting data to be analysed and quantified. Each feature’s impact on network performance will be looked at. Finally, a comparison of each feature’s impact, under various situational constraints, will be performed quantifying situational influence on performance.

References

[1] Siyal, M.Y. & Barkat, B. 2002. A novel trust service provider for the internet based commerce applications. Internet research: electronic networking applications and policy, 12(1):55-65.

[2] Khare, R., and Rifkin, A., Weaving a Web of Trust. In: World Wide Journal, Volume 2, Number 3, 1997, pp. 77-112.

[3] Britner, M.J., Servicescapes: the impact of physical surroundings on customers and employees, Journal of marketing, Volume 54, 1992, pp. 69-82.

[4] Yang, Y., Brown, l., Newmarch, J., and Lewis, E., eCommerce Trust via the Proposed W3 Trust Model, the PACCS01Conference Proceedings, July 2001, Australia, pp. 9-14 .

[5] Nooteboom, B., (2002) Trust: Froms, Foundations, Functions, Failures, and Figures, Edward Elgar Publishing, Ltd. Cheltenham UK, Edward Elgar Publishing, Inc. Massachusettes, USA, ISBN: 1 84064 545 8.

[6] Guha, R., Kumar R., Raghaven, P., and Tomkins, A., Propagation of Trust and Distrust. International World Wide Web Conference, Proceedings of the 13th international conference on World Wide Web New York, NY, USA SESSION: Reputation networks table of contents, pp. 403 – 412.

[7] Abdul-Rahman, A., and Hailes, S., A Distributed Trust Model. New Security Paradigms Workshop, Proceedings of the 1997 workshop on New security paradigms, Langdale, Cumbria, United Kingdom, 1998, pp. 48 – 60.

[8] Papadopou, P., Andreou, A., Kanellis, P., and Martakos, D., Trust and relationship building in electronic commerce. In: Internet Research: Electronic Networking Applications and Policy, Volume 11, Number 4, 2001, pp. 322-332.

[9] Lamsel, P., Understanding Trust and Security. Available: http://wiki.uni.Lu/secan-lab/ UnderstandingTrustAndSecurity.pdf. Accessed 25 April 2005.

[10] Ramchurn, S. D., Sierra, C., Godo, L. and Jennings, N. R. A computational trust model for multi-agent interactions based on confidence and reputation. In Proceedings of 6th International Workshop of Deception, Fraud and Trust in Agent Societies, , Melbourne, Australia, 2003, pp. 69-75.

[11] Carbone, M., Nielsen, M., and Sassone, V., A Formal Model for Trust in Dynamic Networks. In: Software Engineering and Formal Methods, 2003. Proceedings. First International Conference on 25-26 Sept. 2003, pp. 54- 61.

[12] Datta A., Hauswirth M., and Aberer K., Beyond "web of trust": Enabling P2P E-commerce.In: E-Commerce, 2003. CEC 2003. IEEE International Conference on Publication Date: 24-27 June 2003, pp. 303- 312.

[13] Patton, M.A., and Josang, A., Technologies for Trust in Electronic Commerce. In: Electronic Commerce Research, Volume 4, 2004, pp9-21.

[14] Marx, M., and Treur, J., Trust Dynamics Formalised in Temporal Logic. In: L. Chen, Y. Zhuo(eds.), Proc. of the Third International Conference on Cognitive Science, ICCS 2001, pp. 359-363. Available: http://www.cs.vu.nl/~treur/cve/publ.html. Accesed on 26 April 2005.

[15] Langheinrich, M., When Trust Does Not Compute - The Role of Trust in Ubiquitous Computing. Workshop on Privacy at Ubicomp 2003, Seattle, Washington, October 2003. Available: http://www.inf.ethz.ch/personal/langhein/articles/archive.html #2003. Accessed on 26 April 2005.

[16] Bagley, C., Verma, G.K., Mallick, K., and Young, L., (1979), Personality, Self-esteem and Prejudice, Saxon House, Teakfield Ltd, Westmead, Farnborough, Hants., England, ISBN: 0 566 00265 5.

[17] Held, G., Focus on PATHPING. In: International Journal of Network Management, Volume 11, Issue 4, 2001, pp 259 – 261.

[18] English, C., Nixin, P., Terzis, S., McGettrick, A., and Lowe, H., Dynamic Trust Models for Ubiquitous Computing Environments. Workshop on Security in Ubiquitous Computing, UBICOMP 2002, October 2002. Available: http://www.teco.edu~philip/ubicomp2002ws/organize/paddy. pdf. Accessed: 26 April 2005.

Marika Wojcik was born in Carletonville, South Africa in 1982. Obtained a B.Sc. IT Information and Knowledge Systems degree at the University of Pretoria in 2003, and a B.Sc. (Hons.) Computer Science degree from the University of Pretoria in 2004. Is currently studying her M.Sc. Computer Science majoring in Computer Science at the University of Pretoria.

Dr. H.S. Venter’s research interests are in computer and Internet security, including network security, Intrusion detection, information privacy, and digital forensics. He has published in a number of accredited international subject journals and attended a number of acclaimed international and national, Computer and Information Security conferences, to present his research papers. He is a member of the organizing committee for ISSA and SAICSIT.