RISK MANAGEMENT OF DIGITAL CERTIFICATES IN AD HOC AND P2P NETWORKS

Tong Zhou

Comcast Cable

Lein Harn

University of Missouri – Kansas City

ABSTRACT

In nonhierarchical Public Key Infrastructure (PKI), any user can be a Certificate Authority (CA) to issue digital certificates to other users. As there is no single root CA, it is difficult to check the validity of certificates issued by unknown CAs. It is very risky to trust them without in-depth analysis. How users issue certificates in the real world has not been studied. Solomon Asch's conformity experiment reveals that peoples' decisions are influenced by others. To reduce the risk of trusting malicious certificate issuers, we propose two novel methods, micro method and macro method, for users to make trust decisions based on the relationships among the CAs. They will improve the security in ad hoc networks and Peer-to-Peer (P2P) communications.

Index Terms— trust, risk, certificate, stranger, cluster

1. INTRODUCTION

Ad hoc and P2P applications have increased dramatically challenging the traditional centralized structure. In those applications, it is very often that people encounter strangers. Therefore, security is critical and needs further studies. Many security solutions employ public key cryptographic schemes, such as digital certificate, for authentication and confidentiality. A digital certificate is used to bind a person's public key and identity using a CA's digital signature to prevent impersonation attack. Both hierarchical (i.e. X.509) and nonhierarchical PKI (i.e. PGP or Pretty Good Privacy) can be used to secure communications between two users. Hierarchical PKI requires a root CA, which may not exist in all cross-domain scenarios. Nonhierarchical PKI has the flexibility to allow any user to be a CA. However, the trust between the CAs is complicated for management. PGP defines trust level, which allows a user to assign three levels of trustworthiness to a CA's certification capability. A person only accepts a stranger's certificate if it is issued by a CA that is completely trusted or two CAs that are marginally trusted by the person. In ad hoc environments, a user will encounter more unknown CAs than known ones.

978-1-4244-1643-1/08/$25.00 ©2008 IEEE

To solve the PGP limitation, we can leverage social networking. The small-world phenomenon tells us that people are connected through an average of six or less intermediate hops (six degrees of separation) [5]. Those solutions rely on an extensive search of possible paths to a stranger. In real life, trust is not always transitive. So, the longer the path length is, the higher the risk exists in trusting friends' friends. Another solution is to utilize a centralized reputation system, which collects opinions from users and determines the ratings for CAs. Usually a new user needs a period of time to establish reputation. In an emergency, the reputation system may not be reachable.

In this paper, we propose a new approach that overcomes the current limitations and complements the existing solutions. In the case that a stranger is introduced by a number of unknown CAs with which a user has no connection, we study how the CAs are related to each other. An issuer is supposed to verify the public key and the key owner's identity. In reality, sloppy and malicious CAs exist. We cannot expect that all the users follow the rule exactly. In the social world, we observe that people are influenced by others. Psychologist Solomon Asch did several social influence experiments in the early 1950s. One of the experiments reports that one third of the subjects conformed to the majority even when the majority is absolutely wrong. The result leads us to believe that in nonhierarchical PKI, a CA's decision is likely to be influenced by other CAs. When a user requests a CA to issue a certificate, the user may present the certificates issued by other CAs. We imagine that it is likely that the CA under the influences or even pressures of those CAs decides to issue the certificate. The decision may be different from the CA's independent judgment. We also suspect that the more a CA trusts the other CAs' certification capability, the higher the possibility that the CA decides to conform to them. The influences reduce the strictness of the CA's certification because its buddies have already made the consent. A more serious threat is that members of a crime group create cross-signed certificates for themselves and distribute the certificates electronically.

We use the sociogram, which is a graph representing a social network, to study the relationship among the CAs and propose two risk control methods, a micro method and a macro method, for a user to decide whether the public key of a stranger should be trusted or not. Our key contribution

is to help users reduce the risk of trusting a stranger who is introduced by other strangers in the absence of a reputation system. With the proposed risk control methods, users will be more confident in ad hoc and P2P applications.

The remainder of the paper is organized as follows. Section 2 discusses the related works. Section 3 reviews the clustering coefficient. Section 4 describes the applications. The micro and macro methods are presented in Section 5 and 6 respectively, and compared in Section 7. Section 8 provides a conclusion.

2. RELATED WORKS

Many discussions on decentralized trust problem are based on the finding of six degrees of separation and the assumption of trust transitivity, which means if A trusts B and B trusts C, then A trusts C. Recent research in trust has mainly focused on trust concept, management, representation and reasoning [6]. Theodorakopoulos and Baras view the trust evaluation as a generalized shortest path problem on a weighted directed graph and propose a trust computation scheme [1]. Josang *et al* discuss a parallel trust combination method including conflicting recommendations, and how to calculate the reputation score [2]. PeerTrust [3] is a reputation-based trust supporting framework to minimize threat in a P2P online community. It includes basic trust parameters and adaptive factors in computing trustworthiness of peers. Certified Reputation [7] is a peer-level rating scheme. The relationships between the rating agent and the rated agent may impact the rating. Cooperating partners may exaggerate each other's performance. Competing agents may underrate their opponents. No relationship may imply impartial ratings.

The studies of trust in technical and business domains are often related to social studies, which use the clustering coefficient as a major measurement for the sociogram analysis. Schank and Wagner propose an efficient approximation algorithm for the clustering coefficient [4]. Asch's social influence experiments in the early 1950s have had a profound impact on group behavior studies [8]. In one experiment he devised he showed the participants in his experiment a line followed by 3 other lines and asked them which line matches the first line. He arranged the real subject at last to answer the question and instructed all the others to give incorrect answers. To his surprise, 37 of the 50 subjects conformed to the incorrect answer at least once. In Rolfe's conditional decision-making study, he provides a mathematical expression of the probability of adoption as function of adoption among friends [9].

Attacks on trust evaluation and decision-making schemes have not been fully addressed in recent studies. With the subjective and dynamic natures of trust, it is difficult to find a universal and secure method for all contexts. In this paper, we propose two methods for reducing the possibility of trusting malicious CAs based on the clustering coefficient. To the best of our knowledge, there is no discussion on how to avoid clustered CAs. The solutions will help people control the risk when communicating with strangers.

3. CLUSTERING COEFFICIENT (CC)

The clustering coefficient introduced by Watts and Strogatz in 1998 is widely used for analyzing small-world phenomenon in social studies [2]. It measures how closely the neighbors of a node in a graph are connected. In this paper, a CA is represented as a node, which may be connected to another CA through a weighted and directional edge. The trust level is reflected as the weight of the directional connection. An arrow from node A to node B indicates that A trusts B. In the context of public key trust, we calculate the clustering coefficient of a node S as follows.

$$CC_T(S) = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} T_{ij} / P(n,2) \quad (1)$$

where $P(n,2)$ is the maximum possible directional trust. T_{ij} denotes the trust level, which refers to the degree to which node i trusts the certification capability of node j. n is the total number of CAs which issue certificates to S. For simplicity, we assume $T_{ij}=T_{ji}$. Therefore $CC_T(S)$ can be calculated by (2)

$$CC_T(S) = \sum_{\substack{i,j=1 \\ j>i}}^{n} T_{ij} / C(n,2) \quad (2)$$

where $C(n, 2)$ is the maximum possible non-directional trust. Figure 1 provides an example to explain the clustering coefficient. Assume User A does not know X, Y, Z and W, which are CAs issuing certificates C_{XB}, C_{YB}, C_{ZB} and C_{WB} to User B respectively. The trust levels are $T_{XY} = T_{YX} = 90\%$, $T_{XZ} = T_{ZX} = 90\%$, $T_{XW} = T_{WX} = 100\%$, $T_{YZ} = T_{ZY} = 80\%$, $T_{YW} = T_{WY} = 100\%$, and $T_{ZW} = T_{WZ} = 100\%$. The $CC_T(B)$ is 93%.

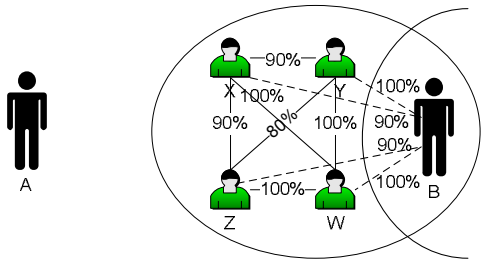

Figure 1 Clustering coefficient

4. APPLICATION SCENARIOS

In our life, we meet strangers inevitably in many occasions (i.e. chat room on the Internet). Different people handle strangers in different manners. Parents and school teachers

often tell children to never talk to strangers for safety concerns. However, not all strangers we encounter are malicious. Many businessmen are willing to explore new opportunities from strangers. In ad hoc and P2P applications, trusting a stranger's certificates, which are issued by unknown CAs, without comprehensive analysis is very risky and may result in financial loss. A possible scenario is that User A meets a stranger S on the Internet. The two parties desire to establish a secure communication channel between them to exchange information. Authentication is a critical step for deriving a session key. A requests S to provide the certificates of S to verify that the public key belongs to S. S responds with a number of certificates issued by unknown CAs, with which A has no direct or indirect connection. Although people tend to believe that the number of malicious attackers is small in modern society, the potential negative impacts they generate may be huge. A may assume that the possibility to encounter an attacker who provides a false certification is p, $(0 < p < 1)$. The possibility that all the certifications are false is therefore p^n, where n is the number of the certificates A receives. However, based on Asch's experiment, some CAs may be influenced by others. So A may have higher possibility, p^{n-i} $(0 \le i < n)$, where i is the number of influenced CAs, to be cheated by colluding CAs. In Section 5, we use the micro method to estimate the number $(n - i)$ of independent CAs. It may be difficult but possible to discover the trust relationships among the CAs. A can query the CAs for the trust levels directly or obtain the information indirectly from other resources (i.e. the Internet). With the collected trust levels, A applies the micro or macro method discussed in section 5 and 6 respectively for making a trust decision.

5. MICRO METHOD

5.1 Fully/highly clustered group

We define a fully clustered group as a group of which all the members trust the other members' certification capability completely (trust level = 100%) and a highly clustered group as a group of which most of the members highly trust each other (trust level ≥ 90%). Figure 2 illustrates three examples. Each circle represents a CA which issues a certificate to S (not shown). In the circle, the CA's name is above the line and the degree of the CA is below the line. The degree is the total trust levels a CA trusts the other CAs. The solid line and the dotted line represent complete trust and marginal trust respectively. In Figure 2 (a), all the members trust the other members completely while in Figure 2 (b) A and D do not trust each other. In Figure 2 (c), A trusts the other members marginally (i.e. 90%). The clustering coefficients of Figure 2 (a), (b) and (c) are 100%, 93% and 97% respectively.

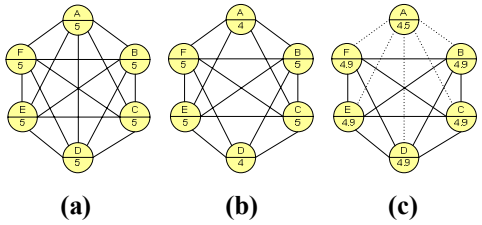

Figure 2 Different clustered groups

We use the degree of a CA to determine whether a CA belongs to a highly clustered group or not. Given an existing highly clustered group G with n nodes, if the degree of a new node X is less than T × n, X does not belong to the group. T is a configurable parameter between 0 and 1. In Figure 3, G has six CAs and T is set to 80%. X is not part of G because X's degree is only 1.9, which is less than the required threshold 4.8.

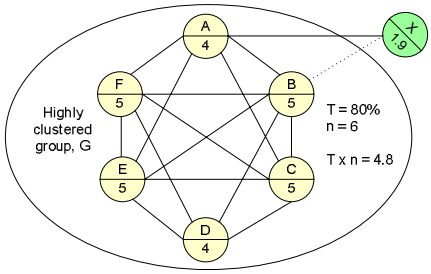

Figure 3 Group affinity determination

5.2 Influence threshold

Solomon Asch's psychology experiment reveals the phenomenon that people tend to conform to the majority, even if the majority gives an incorrect answer. This leads us to examine how people issue certificates in a self-managing environment, where each CA is supposed to verify that the public key matches the identity of the requestor responsively. It is difficult to identify which individuals are influenced by others and yield to the pressure. People are more likely to be influenced by trusted friends than unrelated strangers. We use the trust level to show how people are connected and then predict how they may influence each other. Although the trust level does not exactly match the social relationships among people, they often have overlaps.

A possible situation is that a user with certificates issued by several CAs requests a CA to issue one more certificate to him or her. The requested CA trusts those CAs completely. If the number of the certificates from the trusted CAs exceeds a certain threshold, the CA is under social pressure to conform to the others. We refer to the CA as an influenced CA. The influence threshold in the context of digital certification may vary from person to person. The

Possibility Density Function (PDF) of the influence threshold in the real world is to be studied in the future. Figure 4 shows a hypothetic example of the PDF, which can be used to determine the default influence thresholds in Table 1.

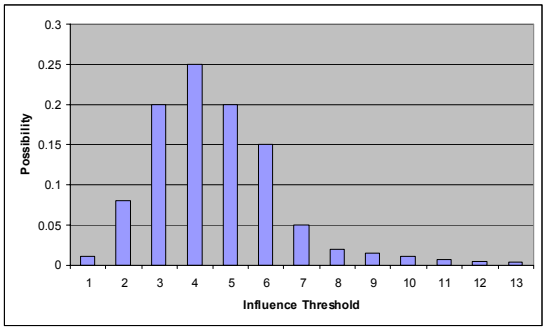

Figure 4 Distribution of influence threshold

5.3 Micro method procedure

The sociogram for trust analysis can be viewed as a collection of nodes that are connected in different ways, including isolated nodes, single high degree nodes, highly clustered nodes and loosely clustered nodes. The proposed micro method can be used to detect the number of independent certificates (or CAs). The certificates from influenced CAs should not be considered when making a trust decision. The detailed procedure is as follows:

1) Identify highly clustered groups based on the collected trust levels. Figure 5 depicts a graph with three such groups, G1, G2 and G3.

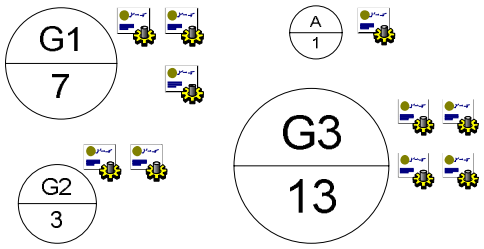

Figure 5 A graph of trust

In each circle, the group name is above the line and the number of members in the group is below the line. G1, G2 and G3 have 7, 3 and 13 members respectively. If a person belongs to two or more groups based on the degree discussed in Section 5.1, we only assign the person to one of the groups.

2) Determine the default influence threshold(s) (Table 1) based on the PDF assumption (Figure 4). A larger group may have a higher threshold as a member may want to wait for more consent from his or her buddies before making a decision.

3) Determine the adjustments and the number of equivalent independent certificates (Table 1). If the group's identity (i.e. high school or IEEE) is discovered, we can use the perceived reputation or behavior of the group to adjust the number. For example, a person may assume that high-tech professionals are more independent and responsible than teenagers in issuing certificates. The threshold can be adjusted accordingly. In addition, people in different countries or cultures have different tendencies to conform to others.

Table 1 Group Conversion

Group Name	Original No. of Certs	Influence Threshold (Default)	Total Adjustment	Equivalent No. of Certs
G1	7	3	0	3
G2	3	2	0	2
G3	13	5	-1	4

4) Remove single high degree nodes. A single high degree node is a node that connects with many (i.e. ≥ 5) other nodes However, the other nodes are low degree nodes (Figure 6). The single high degree node is very likely to be influenced by the other nodes. Therefore, it cannot be considered as an independent CA and its certificate should be ignored.

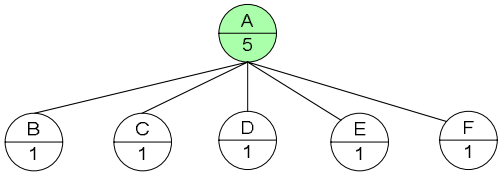

Figure 6 Single high degree node

5) Count isolated nodes. An isolated node is a node that has no trust relationship with other nodes (i.e. A in Figure 5). We consider its certificate as independent.

6) Count loosely clustered nodes. Loosely clustered nodes have few connections with others. All or a portion of them can be considered as uninfluenced nodes.

7) Calculate the total number of independent certificates and compare it with the number of required independent certificates. If the former is larger than the latter, then start the process to verify all the digital signatures on the certificates using the public keys of the CAs. If all the verifications are successful and the corresponding public keys in these certificates are the same, this public key is accepted as the stranger's public key. For simplicity, we omit other factors, such as certificate expiration and

revocation in this paper. Otherwise, the public key is considered as invalid.

6. MACRO METHOD

The macro method utilizes a predetermined clustering coefficient threshold to measure whether the CAs which issue the certificates to the stranger are too "closed" to each other at high level. The larger the clustering coefficient is, the more the issuers trust each other and therefore the more influence and pressure exist among the CAs. If the calculated clustering coefficient is lower than the predetermined threshold, the digital signatures on the certificates can be verified using the public keys of the CAs, and the corresponding public keys in these certificates are the same, this public key is accepted as the stranger's public key. Otherwise, the public key is considered as invalid.

Theorem 1 For a given node *S* that has *n* neighbors, assume *n* can be divided into two fully clustered groups, where there are k_1 neighbors in one group and k_2 neighbors in the other group ($k_1 + k_2 = n$). When k_1 is equal to k_2, the clustering coefficient of S has the minimum value (3). For simplicity, we assume n is an even number.

$$CC_{T-\min,2}(S) = \frac{n-2}{2(n-1)} \qquad (3)$$

Proof for simplicity, we only prove the case when n is an even number.

$CC_{T,2}(S) = (C(k_1,2) + C(k_2,2)) / C(n,2)$
$= 1 - c \times k_1 \times k_2,$
where $c = 2 / (n \times (n - 1))$. As $k_1 + k_2 = n$, $CC_{T,2}(S) = 1 + c \times (k_1 - n/2)^2 - c \times n^2/4$. When $k_1 = n/2$, $CC_{T\text{-}min,2}(S) = (n - 2) / (2 \times (n - 1))$. QED.

Corollary 1 For a given node *S* that has *n* neighbors, assume any neighbor belongs to one of *k* fully clustered groups. When each group has the same size, the clustering coefficient has the minimum value (4).

$$CC_{T-\min,k,l}(S) = \frac{n-k}{k(n-1)} \qquad (4)$$

The clustering coefficient threshold for making a trust decision should be robust in the presence of liars. A liar is defined as a person who deliberately provides a completely false answer in response to a trust level query. No information in the liar's response is correct. For example, a user belongs to a group claims he/she has no trust relationship with any member in the group. In the design of the macro method, it is necessary to adjust the base threshold (4) to avoid attack from liars.

Corollary 2 For a given node S that has *n* neighbors, assume any neighbor belongs to one of *k* fully clustered groups, which have the same size. If there are *l* liars, the adjusted clustering coefficient has the minimum value when the liars are evenly distributed in all the groups. In the case that there are *l/k* liar(s) per group (*l/k* is an integer), the minimum value is as follows:

$$CC_{T-adj\,\min,k}(S) = \frac{(n-l)(n-l-k)}{k \times n(n-1)} \qquad (5)$$

We can derive (5) from $k \times C(n/k - l/k, 2) / C(n, 2)$.

Theorem 2 For two given nodes S1 and S2 that have *n* neighbors each, assume all the neighbors can be evenly divided into K and L fully clustered groups respectively. If K < L, $CC_T(S1) > CC_T(S2)$.

Proof $CC_T(S1) = K \times C(n/K, 2) / C(n,2)$
$= (n/K - 1) / (n - 1)$
$> (n/L - 1) / (n - 1)$, which is $C_T(S2)$ QED.

We conclude that when the neighbors of a node are evenly distributed in the fully clustered groups, the clustering coefficient has the minimum value (Theorem 1), and the more the fully clustered groups exist, and the smaller the clustering coefficient is (Theorem 2). Figure 7 shows the minimum and maximum values for a group with 24 members divided into 2 to 12 fully clustered groups (i.e. for a group with two fully clustered groups, the $CC_{T\text{-}min,2}$ is 47%).

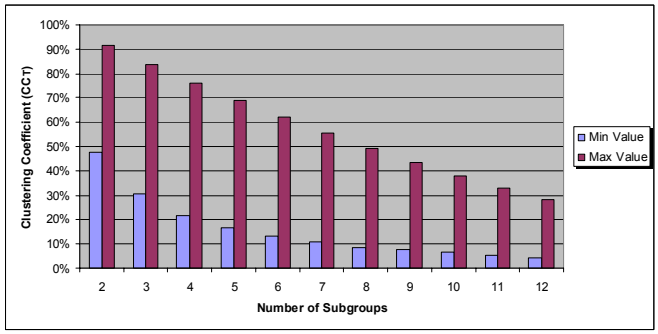

Figure 7 Maximum and minimum CC_T values of groups

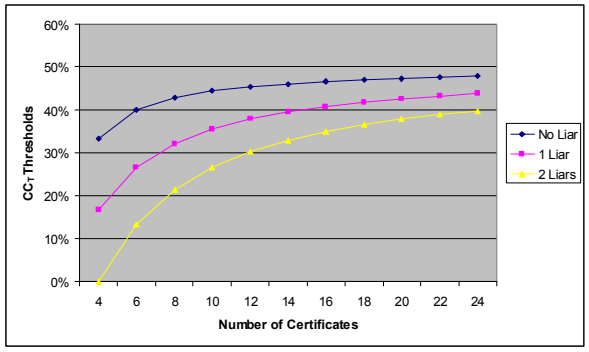

Figure 8 CC_T base and adjusted thresholds

For an example of 24 CAs, if they are divided into 2, 3, 4, 6, 8 or 12 fully clustered groups, the base clustering coefficient thresholds are 47%, 30%, 21%, 13%, 8% and 4% respectively. We recommend certificates from more than two fully clustered groups be required. So the threshold $CC_{T\text{-}min,2}(S)$ is 47%. Considering the trust levels may not be collected accurately and completely due to liars and lack of response, the base threshold is adjusted to make the macro method more secure. Figure 8 depicts the base clustering coefficient thresholds and the adjusted clustering coefficient thresholds for 2 to 24 (even number) CAs with 1 liar and 2 liars in each case. For the case of 24 CAs, the adjusted thresholds with the assumption of one liar and two liars are 43% and 39% respectively.

7. COMPARISON OF THE TWO METHODS

The micro method supports detailed analysis at node and group levels. Its advantage is to allow the decision-maker to fine tune the settings for better accuracy. The macro method uses a predetermined clustering coefficient threshold as a barometer to measure how CAs are trusted. The adjusted threshold has considered the existence of groups and liars. Once an appropriate threshold is selected, the user simply calculates the $CC_T(S)$ and compares it with the threshold. Comparing to the micro method, the macro method is easy to be implemented but lacks deeper analysis; it can still provide risk mediation.

8. SUMMARY AND FUTURE WORKS

Dealing with strangers is challenging but unavoidable. We have proposed two reasonable and flexible methods, micro method and macro method, to control the risks. The social influence from the Asch conformity experiments is considered in the micro method, which allows a user to analyze how each node is connected to the others, and discover clustered nodes. The macro method introduces the adjusted clustering coefficient threshold, which is resistant to clustered group attack and liar attack. Numerical analysis and examples are provided. Both methods will be able to enhance the security in ad hoc and P2P networks, and build consumers' confidence in electronic commerce. In the future, we aim to apply the methods in social networking applications in wireless networks and Customer-to-Customer (C2C) business models. We will interview the participants and collect real data for more accurate representation of the PDF of the influence threshold.

ACKNOWLEDGEMENT

We appreciate the feedback from Dr. Kimberly Hoagland, research fellow of Merck & Co., Inc., and Min Deng, scientist of Xenometrics LLC.

REFERENCES

[1] G. Theodorakopoulos and J.S. Baras, "On Trust Models and Trust Evaluation Metrics for Ad Hoc Networks", IEEE Journal on Selected Areas in Communications, Vol. 24, No. 2, pp. 318-328, Feb. 2006.

[2] A. Josang, R. Hayward and S. Pope, "Trust Network Analysis with Subjective Logic", Twenty-Ninth Australian Computer Science Conference, Hobart, Tasmania, Australia, Jan. 2006.

[3] L. Xiong and L. Liu, "PeerTrust: Supporting Reputation-Based Trust for Peer-to-Peer Electronic Communities", IEEE Transactions on Knowledge and Data Engineering, Vol. 16, No. 7, Jul. 2004.

[4] T. Schank and D. Wagner, "Approximating Clustering Coefficient and Transitivity", Journal of Graph Algorithm and Applications, Vol. 9, No. 2, pp. 265-275, 2005.

[5] D.J. Watts, "Six Degrees: The Science of a Connected Age", W. W. Norton & Company, 1st Edition, Feb. 2003.

[6] Ji Ma and Mehmet A. Orgun, "Trust Management and Trust Theory Revision", IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans", Vol. 36, No. 3, May 2006.

[7] Trung Dong Huynh, Nicholas R. Jennings and Nigel R. Shadbolt, "Certified Reputation: How an Agent Can Trust a Stranger", Autonomous Agents and Multiagent Systems, Hakodate, Hokkaido, Japan, May 2006.

[8] John M. Levine, "Solomon Asch's Legacy for Group Research", Personality and Social Psychology Review, Vo1. 3, No. 4, pp. 358-364, 1999.

[9] Meredith Rolfe, "Social Networks and Threshold Models of Collective Behavior", December 10, 2004. (http://home.uchicago.edu/~mrrolfe/research.html)