SCIS 2002 The 2002 Symposium on Cryptography and Information Security
Shirahama, Japan, Jan. 29 - Feb.1, 2002
The Institute of Electronics, Information and Communication Engineers

審査能力について考慮した多次元トラストメトリックに関する考察
Consideration about the Multi Dimensional Trust-metrics with judge abilities

田村　仁* 松浦幹太* 今井秀樹*
Jin Tamura Kanta Matsuura Hideki Imai

あらまし　公開鍵を利用したインフラが広がるにつれ、相手や相手の鍵に対する信頼の問題はますます重要になってくる。今までにも信頼度の定量化（トラストメトリック）に関する研究はいくつもなされており、中でも[BBK94]では信頼度を６つのクラスに分類しながらもそれらに対して汎用的な定量化がなされている。しかしながらその中で「エンティティが実情とかけ離れた信頼情報を発するかどうか」に関わるクラス(non-interference)については、他の５つと並行して議論されるべきではないと考えられ、そのクラスの不正や誤りに対する感度が高いという問題がある。本論文では、まずこのような問題点を整理し、その改良方法を提案する。また提案方式について典型的な信頼度分布を仮定して数値実験による従来方式との比較を行い、その結果、提案方式によって不正に対する感度を下げることができることを示す。さらに、既存の方式では見られなかった、多次元出力方式を提案する。

キーワード　トラストメトリック、信頼度、direct trust、recommendation trust、審査能力

1　はじめに

オープンなネットワークが広がるにつれて、ユーザがどうやってお互いを認証、信頼するかということが困難になってくる。その解決の一つとして近年代表的なものでいえばPKI（Public Key Infrastructure）方式があげられる。しかしながら、全エンティティが単一の認証局からの認証を受けることは現実的に難しいので、証明書の連鎖を利用することになる。その結果、多数の証明書を入力情報として、認証局も含めエンティティ同士がどうお互いの信頼度を定量化し評価するかというトラストメトリックが重要になっている。

トラストメトリックは、PKI に限らず、より一般の場合についても重要な研究課題であり、今までもいくつもの研究がなされている。中でも[BBK94]で提案されている方式（以降ではBBK 方式と記す）は、次に示す特徴などのため、後のトラストメトリックの研究にも大きな影響を与えた（[M96], [RH97], [RS97]）。

① ６つの信頼度クラスを設けている。

② エンティティ間で信頼度を direct trust 値と recommendation trust 値（つまり、ユーザを推薦する能力）を区別して割り当てながら定量化を行っている。また、その区別が信頼度計算方式においても明確に反映されている。これは、下位のエンティティがさらに下位を審査する能力も考慮できる点などで、意義がある。

③ ①の６つのクラスにわたり、信頼度一般について考察している汎用的な方式となっている。そのため現在でも、代表的なトラストメトリックである。

しかしながら、６つのクラスの中で「エンティティが実情とかけ離れた信頼情報を発するかどうか」に関わるクラス(non-interference)については、他の５つと並行して議論されるべきではないと考えられ、そのクラスの不正や誤りに対する感度が高いという問題がある。本論文では、まずこのような問題点を整理し、改良方式を提案する。そのために我々は新たに審査能力に対する信頼度（judgment trust）を定義する。また、典型的な信頼度分布を仮定して数値実験による比較を行い、その結果、改良方式によって不正に対する感度を下げることができることを示す。さらには、既存の方式では見られなかった多次元出力方式を提案し、その意義を簡単に考察する。

ここより先、本論文の構成は以下の通りである。まず２章で BBK 方式の概要を述べ、３章ではその中の問題点と、それらによって起こりうる攻撃を示す。また４章で

*東京大学生産技術研究所　〒153-8505 東京都目黒区駒場 4-6-1.
Institute of Industrial Science, University of Tokyo, Komaba 4-6-1, Meguro-ku, Tokyo 153-8505, JAPAN.
jin@imailab.iis.u-tokyo.ac.jp, kanta@iis.u-tokyo.ac.jp, imai@iis.u-tokyo.ac.jp

は改良方式を提案し、５章で数値実験による分析と多次元出力に関する考察を行う。最後に６章でまとめる。

2　BBK 方式

２．１　信頼度のクラス分け

[BBK94]では､信頼度に関して次の６つのようにクラス分けを行っている。

① key generation
確かな質の鍵を生成しているかどうかについての信頼度。

② identification
得られているデータ（例えば公開鍵など）がしっかりとエンティティとバインドされているかどうかについての信頼度。

③ keeping secrets
個々の信頼度の情報について漏れていないかどうかについての信頼度。

④ non interference
他のエンティティのやりとりに対する干渉（例えば、盗聴やなりすましなどによるもの）をしないかどうかについての信頼度。

⑤ clock synchronization
正確な時刻を刻んでいるかどうかについての信頼度。

⑥ performing algorithmic steps
きちんとプロトコルの仕様通りにすすめているかどうかについての信頼度。

２．２　Direct Trust と Recommendation Trust

上記した各信頼度クラスそれぞれについて次のような二種類の信頼度が設けられている。

- Direct Trust
  相手エンティティに対してどれほど信頼するか。
- Recommendation Trust
  相手エンティティが他のエンティティを推薦する能力に関してどれほど信頼するか。

各エンティティは、相手エンティティに対し信頼する度合いに応じて0から1までの実数値（0，1も含む）を割り当てる。次節（２．３）では、便宜上エンティティ A がエンティティ B に対して direct trust をしていることを $A \to B$、recommendation trust をしていることを $A \Rightarrow B$ と表すことにする。また本論文ではエンティティ A、エンティティ B、・・・を誤解の生じない限り単に A, B, L と書くことにする。

２．３　信頼度の計算方法（定量化）

エンティティは、自分もトラストアンカー（自分が絶対的な信頼を置いているエンティティ）も Direct Trust を割り当てていないエンティティに対する信頼度を以下の３つの規則を用いて計算していく。ここでは自分またはトラストアンカーを A として、A の B に対する信頼度を算出することを想定する。

[規則①]

recommendation trust 値 r_1 のパス $A \Rightarrow \mathrm{L} \Rightarrow C$ と、r_2 の recommendation trust パス $C \Rightarrow D$ が存在したときパス $A \Rightarrow \mathrm{L} \Rightarrow C \Rightarrow D$ の recommendation trust 値は $r_1 \cdot r_2$ とする。

[規則②]

recommendation trust 値 r のパス $A \Rightarrow \mathrm{L} \Rightarrow X$ と、値 d の direct trust パス $X \to B$ が存在したとき、パス $A \Rightarrow \mathrm{L} \Rightarrow X \to B$ における direct trust 値 $d(A,B)$ は

$$d(A,B) = 1-(1-d)^r \quad \text{とする}$$

[規則③]

A から B までの間、最後が direct trust パス $X_i \to B (1 \le i \le m)$ で終わるパスが各々 n_i 本存在し、各々の direct trust 値を $d_{i,1}, \mathrm{L}, d_{i,n_i}$ とすると、それらを合成した時の A から B への direct trust 値 $d_{com}(A,B)$ は

$$d_{com}(A,B) = 1-\prod_{i=1}^{m} \sqrt[n_i]{\prod_{j=1}^{n_i}(1-d_{i,j})}$$

とする。

3　問題点及び攻撃モデル

３．１　問題点

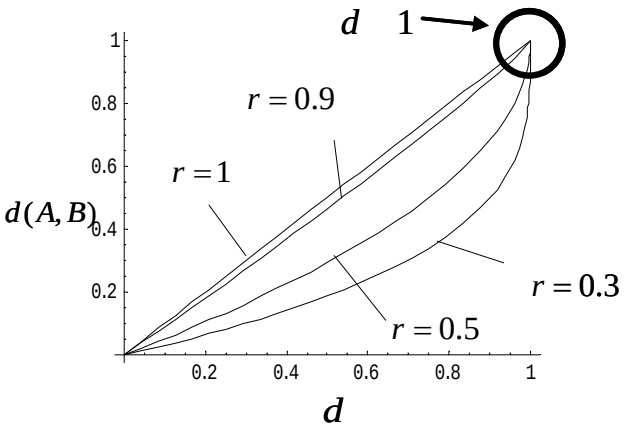
BBK 方式では、６つの信頼度クラスに対し汎用的な定量化を行っている。しかしながら non interference のようなクラスに関しては、例えばある特定のエンティティが信頼度を意図的に操作しようとすることなどについても考慮されなくてはいけないので、他の５つのクラスと全く同列に議論されるべきではない。これが具体的にどう問題になるかを、以下で指摘する。

例えば、必ずしも不正ではないとしても、あるエンティ

ティが実情とはかけ離れた信頼度を割り当てることは十分に考えられる。そのため、特定のエンティティの行動によってでてくる信頼度が大きく動かされること（つまり感度が高すぎること）はあまり望ましくない。実際に BBK 方式では、比較的高い数値の割り当てに関する感度がかなり高い。その理由を分析すると、次の２つの問題点があることがわかる。

[問題点①]

規則②における $d(A,B)$ の値が、d が１に近い領域で、d の変化量に対する変化量が大き過ぎる（グラフ３－１参照）。その結果、規則③における $d_{com}(A,B)$ の値も、X_i が B に対して割り当てる direct trust 値に対して過敏になる。特に、比較的高い trust 値の割り当てに対してこのことが顕著である。極端な例をあげるなら、X_i $(1 \le i \le n)$ のうちどれか一つのエンティティが B に対して direct trust 値として１を割り当てるなら、他の direct trust 値、recommendation trust 値やパスの構造によらず $d_{com}(A,B)=1$ となってしまう。



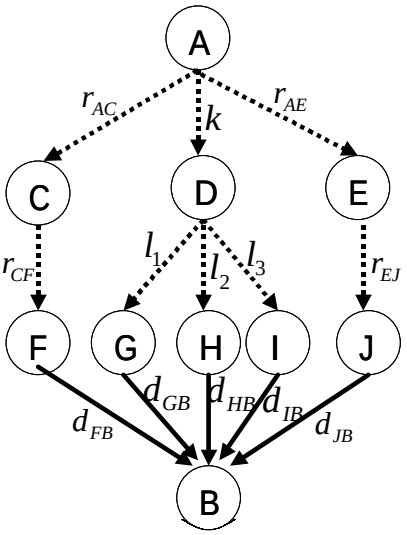
グラフ３－１. 規則②で合成する前の direct trust 値（横軸）と合成後の direct trust 値（縦軸）の関係.　r は recommendation trust 値.

[問題点②]

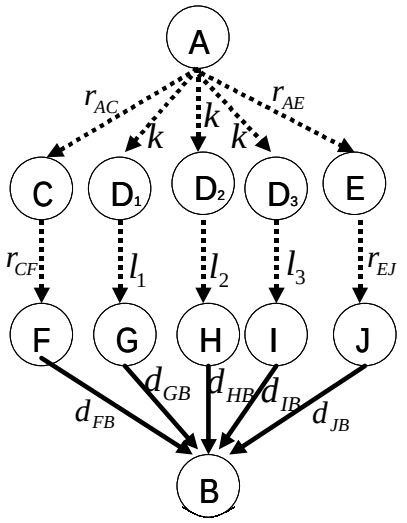
あるエンティティが複数のエンティティに対して recommendation trust 値を割り当てる場合、複数のエンティティが一つの recommendation trust 値を割り当てるのと同じ影響が出てしまうという問題点。これによって特定のエンティティの割り当てる信頼度への依存度が高くなりうる。例えば、次に示す図 A と図 B では $d_{com}(A,B)$ の値が等しい。

３．２　攻撃モデル

まず、A は、$d_A(0 \le d_A \le 1)$ 以上の信頼度（direct trust 値）を示す時にはその相手エンティティの公開鍵を信頼するという個人的なポリシーを有するエンティティであ



図A



図B

ると仮定する。攻撃としては、B が、エンティティ A の recommendation trust パスにおいて配下にある、あるエンティティ D と結託することにより $d_{com}(A,B)$ の値を $d'(A,B)$ に高めるものを考える。この時、もともと

は $d_{com}(A,B) < d_A$ であったにもかかわらず $d'(A,B) \geq d_A$ となれば攻撃成功となる。ここでは二通りの攻撃方法を示す。

[攻撃①]

単純な攻撃である。結託者 D が B に対し本来存在しない direct trust を発し、direct trust 値 d を割り当てる。攻撃の趣旨から、d には比較的高い数値を割り当てる（つまり、$0 < d \leq 1$）。

[攻撃②]

[RS97]でも多少ふれられている、[攻撃①]よりもかなり強力な攻撃方法である（5．2図c参照）。まず結託者 D は、自分のすぐ配下に複数の偽のエンティティ $D_1, D_2, \cdots, D_n$ を設け、更に $D_1, D_2, \cdots, D_n$ に対し recommendation trust 値 $r_i (1 \leq i \leq n)$ を、$D_1, D_2, \cdots, D_n$ から B に対して direct trust 値 $d_i (1 \leq i \leq n)$ を割り当てる。こちらにおいても、攻撃の趣旨から、$r_i (1 \leq i \leq n)$、$d_i (1 \leq i \leq n)$ には比較的高い数値を割り当てる。

今回は紙面上の都合により、次節では[攻撃①]に関する結果のみを示し、[攻撃②]については４章で実験結果とともに考察することにする。

３．３　単純な攻撃の可能性

ここでは[攻撃①]が成り立つ trust 値割り当てが常に可能であることを示す。

A の D に対する recommendation trust 値を $r(A,D)$ とし，また D から B に対する direct trust d が存在しない時（つまり攻撃を受ける前）と、存在した時（攻撃を受けた後）の A の B に対する direct trust 値をそれぞれ $d_{com}(A,B)$、$d'_{com}(A,B)$ とおく。すると、$d'_{com}(A,B)$ は、パスの構造に関係なく次式で表せる。

$$
\begin{aligned}
&d'_{com}(A,B) \\
&= 1-(1-d_{com}(A,B))[1-\{1-(1-d)^{r(A,D)}\}] \\
&= 1-(1-d_{com}(A,B))(1-d)^{r(A,D)}
\end{aligned}
$$

攻撃成立の必要十分条件は、上式が d_A よりも大きいことであり、それは次式と同値である。

$$\frac{1}{r(A,D)} \log\left(\frac{1-d_A}{1-d_{com}(A,B)}\right) < \log(1-d) \quad \text{（式1）}$$

ここで、「攻撃を行う」という仮定により $d_{com}(A,B) < d_A$ であり、さらに $0 < r(A,D) \leq 1, 0 < 1-d_A < 1-d_{com}(A,B)$ で、それぞれが定数だから、（式１）となるような $d(<1)$ が明らかに存在する。つまり、どのような構造に対してもこの攻撃は成功することになる。

4　方式の提案

４．１　審査能力

A からみたエンティティ X の審査能力に対する信頼度（judgment trust）$j(A,X)$ を次のように定義する。

$$j(A,X) = \sqrt[n]{r(A,X)}$$

ただし、$j(A,X)$ は、A と recommendation trust パスでつながっているエンティティに限って定義されることとし、また、n とは A から X までの間に含まれている recommendation trust パスの個数である。つまり、$j(A,X)$ とは、A から X までのそれぞれの recommendation trust 値の相乗平均にあたる。こうすると、A からのパスが増えても recommendation trust 値が急激に小さくなりすぎることはない。このような judgment trust を用いて、以下のように BBK 方式の改良を行った。

４．１　[問題点①]に関する改良

[規則２]を次のように改める。

[規則２’]

recommendation trust 値 r のパス $A \Rightarrow \cdots \Rightarrow X$ と、値 d の direct trust パス $X \rightarrow B$ が存在したとき、パス $A \Rightarrow \cdots \Rightarrow X \rightarrow B$ における direct trust 値 $d(A,B)$ は

$$d(A,B) = 1-(1-j(A,X)\cdot d)^r \quad \text{とする。}$$

我々が[規則２’]を適用した主な理由は、次の２点である。

[理由①]
エンティティは自分のパスの配下にあるエンティティに対する信頼度のみを操作可能であるので、パス全ての相乗平均をとることによって特定のエンティティの操作に対する感度を下げる。これによって、出力の direct trust 値が１になるのは、少なくとも一つの A から B までのパスにおける recommendation trust 値全てと、direct trust 値にともに１が割り当てられている時のみとなる。
[理由②]
この規則により、$d(A,B)$ の上限が $j(A,X)\cdot d$ となるが、ここにエンティティ X の審査能力としての意味合いを込めた。

４．２　[問題点②]に関する改良

次のような規則を加える。
[規則４]
パス構造が決定したとき、複数のエンティティに対して recommendation trust 値を割り当てているエンティティ（ただし A は除く）の recommendation trust 値 $r(A,E)$ を次のように計算する。（ここではそのエンティティを E とし、もともとの recommendation trust 値としては、エンティティ $E_i\,(1\le i\le n)$ に対してそれぞれ $r_{original}(E,E_i)$ を割り当てているとする）

$$
\begin{aligned}
&r(E,E_i)\\
&= j(A,E)+\mathrm{sgn}(r_{original}(E,E_i)-j(A,E))\cdot\left|r_{original}(E,E_i)-j(A,E)\right|^n\\
&= j(A,E)+\frac{r_{original}(E,E_i)-j(A,E)}{\left|r_{original}(E,E_i)-j(A,E)\right|}\cdot\left|r_{original}(E,E_i)-j(A,E)\right|^n\\
&= j(A,E)+\left\{r_{original}(E,E_i)-j(A,E)\right\}\cdot\left|r_{original}(E,E_i)-j(A,E)\right|^{n-1}
\end{aligned}
$$

ここでも４．１と同様の理由で $j(A,X)$ を用いた。

以上の２点についての改良を行ったものを以下では提案方式と呼ぶことにする。

５　　数値実験及び考察

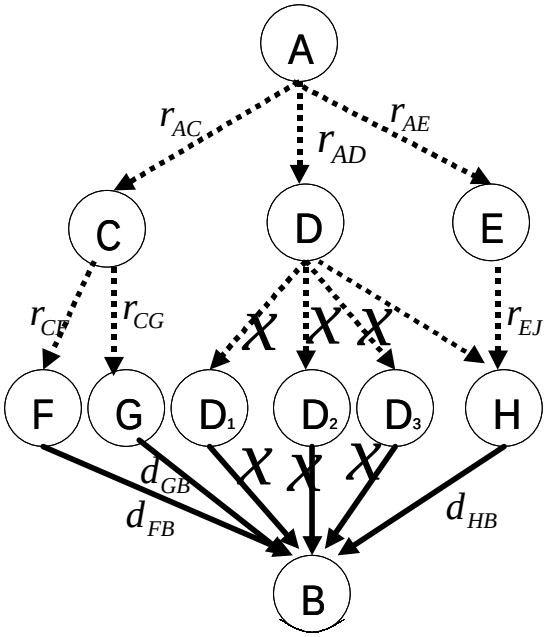
BBK 方式と提案方式を、数値実験で比較した。

５．１　実験の趣旨

提案方式が、BBK 方式に比べ
[趣旨①]
BBK 方式が目指したトラストメトリック本来の性質を保っているか。
[趣旨②]
「特定のエンティティの操作に対して $d_{com}(A,B)$ がどの程度変動するか」という感度が下げられているか。
に着目してその振るまいを調べた。この感度を下げることは、必ずしも不正に対してのみならず、特定のエンティティが実情とかけ離れた高い信頼情報を発する場合においても耐久性をもたせることとなる。

５．２　実験方法

本実験では、３．２[攻撃②]を実装し、BBK 方式と提案方式それぞれに対して攻撃を試みて、その振る舞いを調べた。攻撃をより単純化するべく、D が生成する偽のエンティティの人数は最大で３人までとし、どちらにも比較的高い値が割り当てられるであろう信頼度については一様に x を割り当て、この値を０から１まで変動させてその振る舞いを調べた。以下に本実験で用いたパス構造を図示する。



図Ｃ

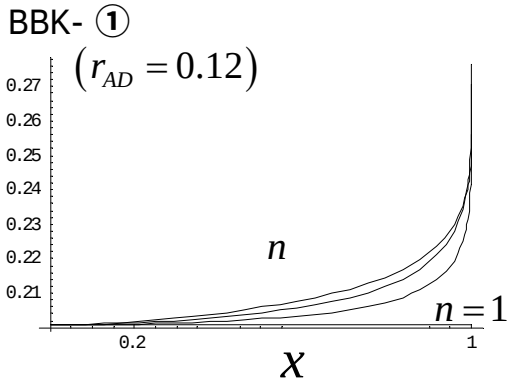
また、$\{r_{AC},r_{AD},r_{AE},r_{CF},r_{CG},r_{EJ},d_{FB},d_{GB},d_{HB}\}$ については
ケース１．ランダムな値の中での攻撃の影響
（1.0 以下のランダムな値）
ケース２．全体的に高い数値の中での攻撃の影響
（0.6 以上および 0.5 以上のランダムな値）
ケース３．全体的に低い数値の中での攻撃の影響
（0.4 以下および 0.5 以下のランダムな値）
について調べるため、それぞれについて数種類ずつサンプルを試した。

５．３　実験結果及び考察

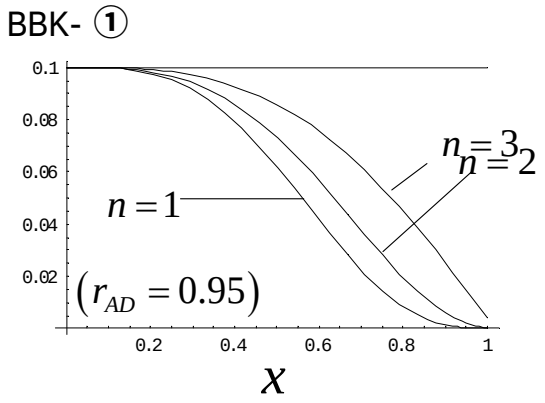
提案方式では、ケース１や２においては BBK 方式と振る舞いが似ていた。むしろ少し高めの x に対してはほとんど 1.0 を出力してしまうような BBK 方式に比べ適切であ

った。しかし、ケース３のように全体的に低いようなケースでは、かなり出力値が低過ぎ、望ましくない。これは[規則２’]により、小さな $j(A,X)$ の値にそれぞれのパスにおける $d(A,B)$ の上限として押さえられた影響が強くでているためと考えられる。ただし、実際的なケースでは、ケース１か２のように信頼度が割り当てられる方が自然であることが考えられる。

また、紙面の都合で詳細は割愛するが、同様の条件の下で BBK 方式と４．１で導入した改良のみを加えた方式(これを改良①方式と呼ぶことにする）との差分（必ず出力は BBK 方式の方が等しいかまたは大きくなる。証明は略す）をとる数値実験も行った。すると、場合によっては次のような興味深い結果が得られた。すなわち、r_{AD} ひいては $j(A,D)$ が小さいときほど差分はかなり広がった（グラフ５－１)。逆に、r_{AD} ひいては $j(A,D)$ が大きいときは x が大きくなるにつれ差分はほとんどなくなる（グラフ５－２)。以上のことをふまえ、最後に次節のような提案をする。



グラフ５－１．縦軸は BBK94 方式と提案方式のトラストメトリック出力値の差



グラフ５－２．縦軸は BBK94 方式と提案方式のトラストメトリック出力値の差

５．４　多次元メトリックスの提案

数値実験により、どの出力方式もケースバイケースで長所、短所があることがわかった。そこで、出力を既存方式のような１つのスカラーではなく、たとえば、（BBK 方式の出力、提案方式の出力、BBK 方式の出力と 改良①方式の出力の差）を提示することにより、よりユーザが状況判断しやすくなるのではないかと考える。たとえば、この場合において出力が（1.0、小、大）の時は、全体的に信頼度が低く、どこかのエンティティからは 1.0 の direct trust 値が割り当てられている。更に A からはあまり審査能力の信頼度が高く評価されていないエンティティのパスの配下の誰かから複数のエンティティに高い recommendation trust 値を割り当てられていることが予想できる。また、（1.0、1.0、1.0）ならば、必ず少なくとも一つの recommendation trust 値、direct trust 値全てに 1.0 が割り当てられているようなパスが存在する、というように、BBK 方式だけではわからないところまで判断が可能になるということが期待できる。

６　　まとめ

本論文では[BBK94]ではほとんど考慮されていなかった攻撃、すなわち不正または判断ミスにより実情とかけ離れた信頼情報を発する行為に対して、トラストメトリックの感度を考察した。そして、この感度を定性的に下げる方式の提案を行った。その際に、審査能力についての信頼度を定義した。実際の数値実験による従来方式との比較では、提案方式にある程度の有効性は見られたが、単独で用いる場合の限界点も明らかとなった。以上を受け、最後に、出力を多次元にすることで各方式の長所短所を補い合いユーザの判断を助ける新しい多次元出力方式を提案した。今後は、さらに詳しい検討が必要だと考えている。

参考文献

[BBK94] T. Beth, M. Borcherding, and B. Klein, ``Valuation of trust in open networks,”Computer Security, LNCS875, pp.3-18, Springer-Verlag, 1994.

[M96] U. Maurer, ``Modelling a Public-Key Infrastructure'' , Computer Security, LNCS1146, Springer-Verlag, pp.325-350, 1996

[RH97] A. Abdul-Rahman and S. Hailes, ``A Distributed Trust Model'', Proc. of New Security Paradigms Workshop '97, pp.48-60, Sept. 1997.

[RS97]M.K.Reiter and S.G.Stubblebine,”Toward acceptable metrics of authentication”, Proceedings of the 1999 IEEE Symposium on Securty and Privacy, pp. 10-20,May 1997