時系列トラストの検証法に関する研究

［研究代表者］河辺義信（情報科学部情報科学科）
［共同研究者］水野忠則（情報科学部情報科学科）
大久保一彦（日本電信電話(株)）
福永利徳（日本電信電話(株)）
五郎丸秀樹（日本電信電話(株)）

研究成果の概要

近年の大規模災害においては，被災者の安否情報や救援情報の交換にソーシャルメディアが積極的に活用され，早期の安否確認や人命救助などに役立てられている．しかし災害時に交換される情報は，すべてが信頼できるとは限らない．たとえば，2018 年の大阪府北部地震では，その当日中にツイッター上に「シマウマが脱走した」「京セラドーム大阪の屋根に亀裂が入っている」「京阪電車が脱線している」などのデマ情報が流され，リツイートされて混乱を引き起こした．ソーシャルネットワークで交換される情報がどれだけ信じられるのか，情報の発信元のユーザはどれだけ信じてよいのか，といった，「トラスト（信頼)」の評価が重要となっている．

本共同研究の目的は，トラストをフォーマルに定式化し，正しさを形式検証することである．時々刻々と変化するトラストを検証するため，本研究では分散アルゴリズム理論を適用して，トラストに関する安全性を検証した．加えて，トラストの反対にあたる状態は「トラストがない状態」だという観点から，信頼の欠乏に関する検討も行った．

研究分野：形式手法，プログラム理論，セキュリティ検証，プライバシ，トラスト
キーワード：トラスト，分散アルゴリズム理論，トレース集合，定理証明ソフトウェア，ファジィ理論

1. 研究開始当初の背景

近年，大規模な事故・災害・病気の蔓延などの非常事態においてソーシャルメディアが利用されている.そこで交換される情報は,すべてが信頼できるとは限らない.情報がどれだけ信じられるのか，などの「トラスト（信頼)」の評価が重要になっている．トラストを数理的に扱う試みとして，Ｍａｒｓｈらの定式化がある．ここではトラストが−１から１の間の評価値（トラスト値）で与えられ，さらに人の心的状態を「トラスト（信頼)」「ディストラスト(不信)」など 4 種類に分類している．

Ｍａｒｓｈらは，完全なる不信（トラスト値＝−１）から完全なる信頼（トラスト値＝１）までを，１次元的な指標で扱った．しかし，従来の理論では「信頼しているが,不信感もある」という矛盾や「信頼も不信もない」といった無関心を扱えなかった．これを解決するため，研究代表者は予備研究において,信頼と不信の組をトラスト値とする「２次元トラスト値」を導入している．

2. 研究の目的

上述の研究では,トラストに関する矛盾や無関心を扱うことに成功し，類似の理論（主観論理）に対する優位性も示している．しかし，以下の 2 点に課題があった．

① 予備研究の結果は，ある瞬間におけるトラストに関する結果である.しかし実際のソーシャルメディア上のメッセージや参加者のトラストは,時々刻々と変化する．そのため，大規模災害時における被災者やボランティア間の協力関係の構築を行うために,トラスト値の変化を定式化し，予測できるようにする必要がある．

② Ｍａｒｓｈらの研究では，トラストの対義語はディストラストであるが，トラストの対義語は「信頼がないこと（信頼の欠乏)」だと主張する学派がある．トラスト値の時系列的な推移を「信頼量の増減」と考えれば,直感的な理解が容易になると期待できるが,信頼の欠乏については，定式化が行われていなかった．

本研究の目的は,これらの課題を解決することである．

3. 研究の方法

２次元トラスト値を状態と考えれば，トラストの時系列的な変化は，オートマトンで扱える．Ｉ／Ｏオートマトン理論では，実行列をあらわす「トレース」を用いて，「悪い振舞いが発生しない」ことを表す安全性と呼ばれる性質を分析できる．２次元トラスト値にあてはめれば，たとえば「ディストラストの状態に至らない」は安全性である．このアイディアに基づき，Ｉ／Ｏオートマトン理論を応用して，トラストを分析する．

さらに「信頼の欠乏」を考えるため，２次元トラスト値からトラスト量への写像を定める．これについては，ファジィ理論の結果（逆転項目平均法）を適用し，トラスト値の時系列的な増減をモデル化する．

4. 研究成果

(1) トラストに関する「安全性」の検証手法

トラストの安全性を，論理式で表し検証する．述語

$$
\begin{aligned}
&stepTrustSafe(s)\\
&\Longleftrightarrow\\
&\quad (s.tr,\, s.dis) \notin D\\
&\qquad \Longrightarrow\ \forall a \in sig(X)\, \forall s' \in states(X)\\
&\qquad\qquad [\, s \xrightarrow{a}_X s' \ \Longrightarrow\ (s'.tr,\, s'.dis) \notin D \,]
\end{aligned}
$$

を用いると（これは「状態 s がディストラストでないならば，その次の状態もディストラストでない」を表す），「信頼を失うことはない」という安全性は

$$
\begin{aligned}
&\forall s \in start(X)[\,(s.tr, s'.tr) \notin D\,]\\
\wedge\ \ &\forall s \in state(X)[\,stepTrustSafe(s)\,]
\end{aligned}
$$

を証明することで保証できる．ただし，これは必ずしも効率的な手法ではない．そこで本研究では，効率的な証明法を考案した．具体的には，仕様（オートマトン）をふたつ用いる．一方は安全性が自明な仕様Ｓ，他方は通信システムの仕様Ｉである．ＳとＩのトレース包含が示されれば，Ｓの安全性からＩの安全性を導ける．本研究ではＩとしてＳＮＳシステムの設計図を記述し，Ｉ／Ｏオートマトン理論の「フォワードシミュレーション法」を適用した．この証明は定理証明ソフトウェアを用いて自動で行った．図１に，定理自動証明の様子を示す．本実験では，小型サーバ（Dell PowerEdge T110．CPU は Celeron G1620@2.7GHz．メモリは 32GB）を用い，仮想計算機上で，数十秒で証明を完了できた．以上により，安全性検証が可能であることが，実証的に示された．

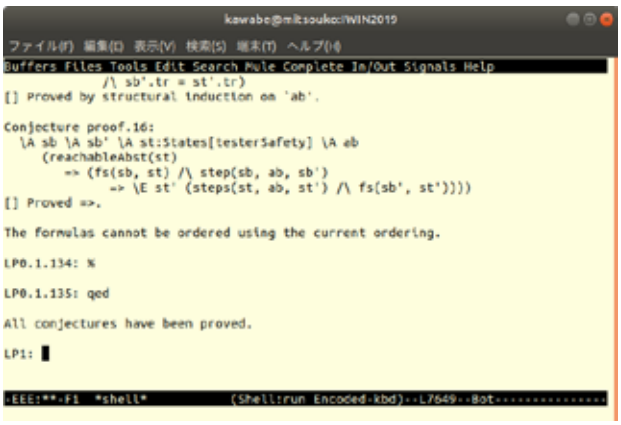

図１：計算機による自動定理証明

(2) 「信頼の欠乏」のモデリング

信頼の欠乏と類似の概念として「信頼不足」という概念が知られており，これはＭａｒｓｈらによるトラスト研究の初期の論文にも議論がある．本研究では，信頼不足を拡張して信頼の欠乏を定義することを試みた．

信頼不足や信頼の欠乏を考えるとき，暗黙的にトラストを量で測っていることになる．つまり，対象を信用している状態では十分なトラスト量があり，信頼不足の状態ではトラストの量が十分ではない．また信頼の欠乏の状態では，トラストの量がないと考えられる．

このアイディアに基づき，本研究では，２次元的トラスト値からトラスト量への写像をｑ（ｔ，ｄ）＝ｔ−ｄで与えた．この変換式は単純だが，ファジィ理論の「逆転項目平均法」という手法（真の度合いと偽の度合いから，「正味の」真の度合いを求める手法）の応用であり，変換の妥当性はファジィ理論に基づき，保証される．

さらに，信頼性理論における「システムを安定運用する際の，不具合受け入れの許容量」と同様の考え方に基づき，信頼の欠乏を数理的に定義することを試みた．

5. 本研究に関する発表

(1) Toshinori Fukunaga, Hideki Goromaru, Tadanori Mizuno, Kazuhiko Ohkubo, Yoshinobu Kawabe, “How to Theorem-Prove Trace-Based Safety Properties,” Int. J. of Informatics Society, accepted (2020).

(2) 河辺 義信，水野 忠則，五郎丸 秀樹，“信頼の欠乏の数理的定義に向けて”，日本知能情報ファジィ学会・ソフトサイエンス研究部会，第 30 回ソフトサイエンスワークショップ，2020 年