

令和 5 年 6 月 15 日現在

機関番号：33903

研究種目：基盤研究(C) (一般)

研究期間：2020～2022

課題番号：20K11822

研究課題名(和文) ソーシャルネットワークのためのトラスト表現と検証法の研究

研究課題名(英文) On Trust Representation and Verification Methods for Messages in SNSs

研究代表者

河辺 義信 (Kawabe, Yoshinobu)

愛知工業大学・情報科学部・教授

研究者番号：80396184

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：本研究課題では「好きだが嫌い」や「好きでも嫌いでもない」といった曖昧な評価を許す感性評価技術を応用した、新たなトラストの表現・分析手法の開発に取り組んだ。また、トラストの移り変わりに関する性質(いつか信用できるようになる、信用を裏切ることはない、など)を効率的に分析する手法の検討を行うとともに、SNS上のメッセージ(水害時の救助要請メッセージ)への分析例を通じて、本技術の適用可能性を確認した。さらに感性評価技術(FCR法)における統合値の算出方法(曖昧さを含むトラスト値から、正味の信用度を求める方法)について、2軸(信用、不信)に限らない3軸以上の場合を扱う拡張を検討し、基礎的な結果を得た。

研究成果の学術的意義や社会的意義

ひとたび災害が起こるとSNSには膨大なメッセージが流れるが、救助要請とは無関係なものも多く、要救助者のメッセージは埋もれてしまう。人命に関わるなかで僅かな「本物の救助要請」を選び出す作業は容易ではないが、本研究では、信用度と不信度の組を計算することにより、救助要請としてのトラストを自動的に与えられるようになった。また災害時に限らない一般の場合でも、ネットワーク上のメッセージやユーザの信用度/不信度を、本研究の手法で適切に評価できる。従来手法ではかなわなかった、対象へのこうした柔軟なトラスト評価を、本研究は可能とした。さらに本研究では、トラストの遷移(移り変わり)についての分析も可能となった。

研究成果の概要(英文)：This study developed a new trust representation and analysis method by applying a fuzzy logic that allows ambiguous trust evaluations such as "I like it, but I don't like it." We also developed techniques for efficiently analyzing properties related to trust transitions (e.g., the information will become trustworthy someday). We confirmed the applicability of our approach by conducting a case study on analyzing messages on SNS during a disaster. Specifically, we evaluated whether some postings were genuine requests to call a rescue during a flood. In addition, we examined an extension of calculating the integration value in the fuzzy logic (i.e., calculating the net trustworthiness from the trust value, including ambiguity) to handle cases with three or more items, not limited to two (trust and distrust).

研究分野：情報セキュリティ，ファジィ論理

キーワード：トラスト ソーシャルネットワーク ファジィ論理 分散アルゴリズム理論 検証 災害時分析

1. 研究開始当初の背景

今日においてはツイッターやフェイスブック等のソーシャルメディアを使った情報交換が日常的に行われるようになってきている。一方で、いわゆるフェイクニュースの問題が取り沙汰されるように、こうしたソーシャルメディア上の情報がいつも信用できるとは限らない。実際、たとえば2018年の大阪府北部地震の際には、「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」などのフェイクニュースがツイッターに流され、広範囲に拡散して問題となった。

ソーシャルメディアから信用できる情報を選び取るためには、メッセージや参加者の「トラスト(信用度)」を定式化し計算・推論できるようにする必要がある。これに対するアプローチとしては、1990年代にMarshら[1]がトラストを数理的に扱う「トラスト計算」の研究を興しており、それを発展させる形で、トラスト研究の流れが情報セキュリティ分野を中心に世界中に広がった。海外においてはIFIP(情報処理国際連合)のTrust Management 国際会議(IFIPTM)やIEEE TrustCOMといった専門の国際会議が開かれており、また国内においても情報処理学会の「情報セキュリティ心理学とトラスト(SPT)研究会」や「トラスト勉強会」などの研究集会在定期的に開かれ、トラスト計算の研究が盛んに行われている。

Marshらのトラスト計算では、メッセージや参加者のトラストは数値として与えられ、それに基づいて信用や不信などのトラストの諸概念が分類・定義されてきた。たとえば、信用とはトラストの数値が極めて高い場合であり、「対象をどれだけ信じて協力できるか」を表す指標として用いられる。一方、不信はトラストの数値が負の場合にあたり、「対象が自分の足をどれだけ引っ張るか」の指標として使われる。このほか、相手を信用できるほどトラストの数値は高くはないが負の値でもない状態である「ミストラスト」と呼ばれる概念も存在する。研究開始当初においては、ネットワーク内の矛盾した情報流通下での信用と不信の分析[2]や不信とミストラストの計算法[3]など、トラストの諸概念についての整理・分類・分析に関する課題が、実問題への応用を見据えつつ、Marshらのトラスト値を用いて研究されていた。

2. 研究の目的

(1) Marshらの手法では、信用と不信は対局に位置する概念であり、不信度は信用度から求められる数値となっていた。しかし一方で、心理学の分野でもトラストの諸概念(信用、不信)を扱う動きはあり、そこでは、Lewickiら[4]の研究のように「人の心の中では、対象への不信度は信用度から決まるわけではなく、独立である」とする結果もある。つまりこの結果は、「相手を信用しているけれども、疑念が全くないわけではない」といった、信用と不信が入り混じった、曖昧なトラストの評価がありうるという指摘をしている。この指摘は自然なものに見えるが、従来のトラスト計算の理論は、こうした曖昧さをうまく扱えないという課題があった。本研究課題の目的のひとつは、信用と不信を独立に扱える新たなトラストモデルを開発し、Marshらの従来型モデルにおけるトラスト分類の成果を、上述の「曖昧なトラスト評価」を許す中でも適用できるようにすることである。

(2) さらに、トラストの評価値は、時とともに絶え間なく遷移する。たとえば、災害発生中のある時点で被災者によって書き込まれた救助要請メッセージが、1時間後にも意味のあるものだとは限らない。その被災者に対する救助が完了した場合には、そのメッセージはもはや無用となり、緊急性の高い救助要請として扱うべきものではなくなる。一方で、その被災者が死亡したままメッセージだけがSNS上に残されてしまった場合においても、同様のことが起こる。本研究課題のもう一つの目的は、対象のトラストの遷移を、分散アルゴリズム理論(システムの状態遷移について盛んに研究されており、豊富な分析手法がある)の知見を活用して分析する手法を開発することである。

3. 研究の方法

本研究課題では、上述の目的(1)を達するためのアプローチとして、ファジィ多項目並列評定法(Fuzzy-set Concurrent Rating法。以下、FCR法)と呼ばれる感性評価・計測の手法を用いた、二次元的なトラストの表現を導入することとした。FCR法は、「好きだが嫌い」という矛盾や「好きでも嫌いでもない」といった無関心を「好き」と「嫌い」の2軸でとらえ、統合値(真の「好き」の度合い)や矛盾度(評価がどれだけ首尾一貫しているか)を計算する理論である。この2軸に「信用」と「不信」を選ぶことで、FCR法で開発された感性評価の手法を、トラストの定義・分析に応用することを試みた。

また、目的(2)については、各時点におけるトラストの値を状態とみなし、I/O-オートマトンと呼ばれる分散アルゴリズム理論の結果を用いて、トラストの遷移を分析するアプローチを試みた。

4. 研究成果

(1) 上述の目的(1)に関連して、予備研究の段階から行ってきた2次元トラスト値の理論の整

備をすすめるとともに、災害時における救助要請メッセージの抽出や大学生への就職支援の問題に対して2次元トラスト値の理論を適用し、有効性を確認した。

まず、前者の事例研究に関する成果から述べる。SNS上のメッセージを分析するには、具体的なトラスト値の決め方を定める必要がある。本研究課題では、水害時の救助要請メッセージの分類（本物の救助要請と信じられるか、それとも疑わしいか）を題材に、信用度と不信度の決め方を検討した。具体的には、信用度を定めるために、メッセージ中の詳細な住所などのプライバシー情報の出現度合いに着目した。平時においてユーザは、郵便物が届くような詳細な自宅住所の情報をSNSに書き込むことはしない。しかし水害時においては、被災者は救助を呼ぶために、そうしたプライバシー情報であってもSNSに書き込むことがある。そこで本研究では、こうした特殊な情報の出現度合いを二次元的トラスト値の信用度として用いた。ただし、実際には詳細な住所情報がメッセージ中に現れたとしても、救助要請ではない場合もある。典型的には、本物の救助要請メッセージを引用したリツイートや、救助完了の報告メッセージなどが、これにあたる。こうしたメッセージには、たとえば「ご心配をおかけしました」などの、救助要請にはふさわしくない文言が現れることがある。そこで本研究では、どのような単語が実際に救助要請メッセージに現れ、どのような単語が現れないかを調査し、不適切な単語を抽出した。不適切な文言が検出されたとき、救助要請とはみなせないと考え、高い不信度を与えて統合値（正味のトラスト値）が低くなるように自動調整した。さらに、曖昧さを含むトラスト情報（信用度と不信度の組）から、曖昧さを取り除いたMarshらによるトラスト値をもとめ、信用・不信などの分類に当てはめてメッセージを分類した。上述の手法を、西日本豪雨時の約20万件のメッセージに適用し、適切に救助要請メッセージが抽出できることを確認した。

後者の事例研究では、2次元トラスト値の理論を、大学生の希望内容と求人情報のマッチングに応用した。就職先企業に関する大学生の主な興味・関心としては、「職種」と「勤務地」が挙げられる。本研究課題では、これらの条件について「希望する事項」と「希望しない事項」をユーザに自由記述形式で回答させ、企業のプロフィール文面とのマッチングをとって、2次元トラスト値を算出した。なお、「希望する事項」と「希望しない事項」のそれぞれから信用度と不信度を求めるが、これらの事項は独立に記述されるため、首尾一貫しない場合がある。たとえば、信用度と不信度の両方が高い状態になることもあれば、どちらの度合いも低くなる場合もある。本研究課題では、ファジィ理論における「統合値」の考え方をを用いて、学生の希望内容から取り出した曖昧さを含む2次元トラスト値から、矛盾のないMarshらのトラスト値を求め、学生と企業のマッチングを判定した。なお、一般に学生は、「希望しない事項」を含む企業を敬遠する傾向がきわめて強い。これは、大学生の就職活動における企業の選択が、人によっては一生に一度のことでもあり、失敗を避けたいと当事者が思うためだと考えられる。上述の水害時のツイート分析でも、「失敗が許されない」という意味では、共通点がみられる。本研究課題で扱ったこれらの事例を通じて、失敗が許されない事態における不信度の決め方について、基本的な指針を得ることができた。

(2) 本研究課題では、上述の目的(2)に関して、分散アルゴリズム理論の結果を適用する形で、トラスト遷移の検証のためのシミュレーション関係の自動生成手法を開発した。また、トラスト遷移に関する性質の定義について、さらなる考察・検討をおこない、ファジィ・シミュレーション関係を用いた新たなトラスト分析法の基礎づけを行った。

トラストの移り変わりに関する性質としては、「ある条件を満たしている間は、信用を突然失うことはない」という意味での安全性や、「いつか信用してもらえる状態になる」という意味の活性を考えることができる。分散アルゴリズム理論では、これらの性質は、状態遷移機械の動作系列である「トレース」の集合を用いて扱われる。本研究課題では、トラストの移り変わりに関するこれらの性質を効率的に分析するため、トレース集合の包含を自動判定する手法を検討した。具体的には、トレース包含の十分条件となる状態集合間の二項関係を自動で発見するツールの開発に向け、プロトタイプを製作しながら、基礎的な検討を行った。

上述のトレース包含に基づくトラストの検証技術では、観測対象（人や情報など）の振る舞いのみからトラスト値を決定できるという仮定を設けていた。救助要請メッセージの判定の事例研究でも同様の考え方をを用いており、モデリングの工夫をすることで、実用上は問題ないように扱うこともできる。しかし一般的には、対象の動作のみからトラスト値を決定できるとする仮定は強すぎると考えられる。そこで本研究課題では、トラストの移り変わりについての分析法について、さらなる検討を行った。具体的には、観測対象の振る舞いのみからトラスト値を決定するのではなく、トラストが「事実確認できない環境下で、対象が期待どおりだと観測者が信じる度合い」であることを踏まえた、観測者の主観を採り入れた新たなモデリングを提案した。上述の研究ではトラスト値を状態とみなしたが、新たな手法においては、観測対象の動作のほか観測者の動作もモデル化し、状態間の対応関係（シミュレーション関係）にトラスト値を割り当てている。これにより「観測者の振る舞いに対して、観測対象が観測者の期待する通りの反応をするか」を分析できるようになった。

(3) トラスト計算は人の心の状態を扱うテーマと言え、これを適切に取得・数値化できるかは、その後の分析を大きく左右する。そこで、上述の目的(1)(2)の成果に加え、本研究課題では、トラスト表現の基礎となるFCR法自体にも着目し、心的状態の取得に関する検討を行った。具体的には、あるクラシック音楽の特定の演奏が「鐘の音」に聞こえるとされる現象（聴覚パ

レイドリアの一種)を題材に,新たな統合値の定義の提案やファジィ推論における出力のデファジィ法に関する新手法の提案に関する検討を行った.従来では,F C R法は2項目(信用,不信)に基づく評価が専らであり,本研究課題で取り組むトラスト評価への応用においても,2項目の場合に限った議論をしていた.しかし一般には,入力評価は3軸以上になる場合もある.本研究課題では,3軸を超える場合(とくに4項目の場合)について,並列尺度応答の分析に関する検討を行い,基礎的な結果を得た.さらに実験と検討を通じて,3項目以上の場合における統合値の計算方法を見つけるための足がかりを得ることができた.

参考文献

- [1] S. Marsh and M. R. Dibben. "Trust, untrust, distrust and mistrust - an exploration of the dark(er) side." In Proc. of iTrust'05, pp. 17-33, Berlin, Heidelberg, Springer-Verlag, 2005.
- [2] G. Primiero, F. Raimondi, M. Bottone, and J. Tagliabue, "Trust and distrust in contradictory information transmission," Applied Network Science, vol. 2, p. 12, 2017.
- [3] G. Primiero, "A calculus for distrust and mistrust," in Trust Management X (S. M. Habib, J. Vassileva, S. Mauw, and M. Muhlhauser, eds.), pp. 183-190, Springer, 2016.
- [4] R. J. Lewicki, D. J. B. McAllister, and R. J. Bies, "Trust and distrust: New relationships and realities," Academy of Management Review, vol. 23, pp. 438-458, 1998.

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計8件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 小田 哲久, 河辺 義信, 山田 洋巳
2. 発表標題 F C R 法応答の統合値に関する新提案—オーケストラ音響刺激の認知を実験課題としたデータの処理—
3. 学会等名 第50回東海ファジィ研究会（蒲研2021）
4. 発表年 2021年

1. 発表者名 河辺 義信, 小泉 佑揮, 大木 哲史, 西垣 正勝, 長谷川 享
2. 発表標題 SNSにおけるトラスト可能な救助要請メッセージの抽出法
3. 学会等名 第37回ファジィシステムシンポジウム
4. 発表年 2021年

1. 発表者名 小田 哲久, 河辺 義信, 山田 洋巳
2. 発表標題 カールシューリヒト指揮「未完成」交響曲録音における鐘の音の認知について
3. 学会等名 日本心理学会第85回大会
4. 発表年 2021年

1. 発表者名 河辺 義信
2. 発表標題 二次元的トラスト表現法の大学生の就職支援への適用
3. 学会等名 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32回ソフトサイエンス・ワークショップ
4. 発表年 2022年

1. 発表者名 小田 哲久, 河辺 義信, 山田 洋巳
2. 発表標題 ファジィ推論における出力のデファジィ法に関する提案－交響楽における特殊音型の認知を実験課題として
3. 学会等名 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32回ソフトサイエンス・ワークショップ
4. 発表年 2022年

1. 発表者名 小田哲久, 山田洋巳
2. 発表標題 新しい代表値指標の特性と心理学的実験
3. 学会等名 第65回日本経営システム学会全国研究発表大会
4. 発表年 2020年

1. 発表者名 小田 哲久, 山田 洋巳
2. 発表標題 交響楽における特殊音型の認知について
3. 学会等名 日本知能情報ファジィ学会第31回ソフトサイエンス・ワークショップ
4. 発表年 2021年

1. 発表者名 河辺 義信
2. 発表標題 トラスト遷移の検証のためのシミュレーション関係の自動生成
3. 学会等名 日本知能情報ファジィ学会第31回ソフトサイエンス・ワークショップ
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担 者	小田 哲久 (Oda Tetsuhisa) (60131132)	愛知工業大学・経営学部・非常勤講師 (33903)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------