

D\_014

# 安否確認システムにおける不完全なデータベースのマッチング

経田 拓馬<sup>†</sup>      鈴木 優<sup>‡</sup>      川越 恭二<sup>‡</sup>

<sup>†</sup>立命館大学大学院 理工学研究科

<sup>‡</sup>立命館大学 情報理工学部

## 1 はじめに

近年、地震等の災害発生に伴い、都心部を中心に防災に対する関心が高まっている。大規模な災害が発生した場合、被災者は自分の家族や知人に自己の安否を知らせるため、検索者（家族や知人が被災した可能性のある人）は被災者の安否を知るために安否確認システムを利用している。NHK 伝言サービス、通信事業者の災害用伝言板サービス、IAA 等の Web サイトなどが代表的な安否確認システムである [1][2]。

既存の安否確認システムでは、被災者の登録した安否情報に対し検索者が一方的に検索を行うという方法がとられているため、携帯電話番号など被災者に関する情報を一つでも検索者が知ることが可能であれば、その検索者は登録されている被災者に関する情報をすべて取得できてしまうという問題がある。例えば IAA で検索可能な項目は氏名、生年月日、性別、血液型等である。今後は、個人情報を含む可能性のある安否情報にもプライバシーを考慮する必要があると考える。

そこで本稿では、プライバシーを考慮した安否確認システムの実現のために、双方向マッチングを用いた安否確認システムを提案する。さらに、実際の災害時には登録内容に誤りや欠如などの不完全な情報が含まれる可能性が高いため、不完全な情報が含まれた場合であっても正確性の高いマッチングを可能とする手法を提案する。提案手法により、プライバシーを考慮した安否確認システムが実現できる。

## 2 双方向マッチングを用いた安否確認システム

### 2.1 双方向マッチング

上記の問題は、被災者自身が安否情報を公開しても良い人を限定することで上記の問題を解決できると考えた。そこで被災者および検索者が、自己の情報（以下 ID 情報とする）と相手を特定するための情報（以下条件情報とする）を設定した後、双方向マッチングを行う方法を提案する。すなわち 1) 被災者は自己の安否情報に加え、ID 情報とその情報を公開してもよい人の ID 情報を条件情報として被災者 DB に登録する。2) 検索者は、自己の ID 情報と安否の知りたい人の ID 情報を条件情報として検索者 DB に登録する。これら 2 つがお互いに合致する組合せを見つけ出すことを双方向マッチングと定義する。双方向マッチングにより、被災者は公開してもよい相手を限定できるた

め、プライバシーを考慮した安否確認システムが実現できる。

例えば、被災者の ID 情報が“井上太郎、男、45 才、滋賀県草津市在住”であるとき、検索者が 40 才代男性という検索条件を指定した場合にはこの被災者の ID 情報はすべての検索者に通知されることとなる。しかし、この被災者が例えば大阪にいる知人の鈴木次郎に伝えたいという場合には、知人の ID 情報を条件情報として設定することで安否情報を公開してもよい検索者を限定することができる。この場合には例えば、“鈴木次郎、40 才代、大阪在住”をいう条件情報の設定が考えられる。

### 2.2 双方向マッチングにおける問題点

2.1 節の方法を用いるためには全ての登録が正確に行われている必要がある。しかし、一般に被災者や検索者はお互いの ID 情報を全て正確に認識しているわけではない。そのため、実際の災害時に登録される条件情報には登録者の曖昧な認識や勘違いにより実際とは異なる条件情報が登録される可能性が高い。このような状況下で、既存のデータベースシステムのように完全一致条件の下でマッチングを行った場合には必要な安否情報の伝達が行われず。しかし、一つでも項目が一致すれば情報を取得できるというような曖昧なマッチングが可能であると、被災者の意図した相手以外に情報を公開してしまう恐れがある。

そこで本稿では、双方向マッチングを行う際、登録された情報に不完全な情報（誤りや欠如を含む情報）が含まれる状況であっても正確性の高いマッチングを行うための方式を提案する。

## 3 不完全な情報のマッチング

本章では、不完全な情報が引き起こす以下の問題点の解決手法について述べる。

- キーによるマッチング（結合）ができない。
- 値による一致ができない。
- 一致・不一致だけでなく一致可能性も不明。

### 3.1 想定環境

被災者または検索者が登録する ID 情報または条件情報の項目を名前、カナ、性別、生年月日、住所、携帯電話番号、所属と想定する。また、想定環境の単純化のために、被災者と検索者が各々自己の ID 情報と相手の条件情報を 1 対 1 で登録するものとする。また、2.2 節で説明したように、登録される情報には不完全な情報を含んでいるものとする。

### 3.2 動的キー作成

災害時に安否確認システムに登録される情報は、どの属性が欠如するか不明であるため、マッチングに使用できるキーを定めることができない。例えば、通常の情報システムであれば社員番号や学生番号がなくと

Matching in Incomplete Databases for Safety Confirmation Systems

Takuma KYODEN, Yu SUZUKI and Kyoji KAWAGOE

<sup>†</sup>Graduate School of Science and Engineering, Ritsumeikan University, <sup>‡</sup>Faculty of Information Science and Engineering, Ritsumeikan University

も名前や生年月日等の属性の組合せであらかじめ候補キーを設定することができるが、本稿が対象とする安否確認システムでは、名前に誤りがあるとか、生年月日の値が欠如しているということも起こりうる。

そこで、登録されている情報を元に、動的にいくつかの情報の組合せをキーとして扱いマッチングを行う。これにより、少量の情報が欠如した情報が存在する場合でもマッチングを行うことが可能となる。またこの際、どの項目の情報が欠如しているかという情報を3.4節で説明する一致可能性の算出に使用する。

例えば、携帯電話番号は他の人と重複しない値であるため、キーとして用いることにより、確実にマッチングできる。しかし、携帯電話番号が欠如していても、名前や生年月日など他の情報を組合せをキーとして扱うことが可能である。

### 3.3 値の集約化

被災者または検索者が検索者または被災者を想定して登録した条件情報に含まれる誤りには、全く異なる情報が登録される場合と、わずかな違いの情報が登録される場合の2種類がある。このうち、後者のような誤りを含む情報に対して集約化を行ったのちマッチングを行う。ここで、集約化とは、登録されている値(1点)から範囲に拡張しグループとして扱う処理である。この集約化により、曖昧な認識の下で登録された、少しの誤りを含む情報でもマッチングすることが可能となる。しかし、誤った情報を正しい情報と見なすため正確性の向上の妨げになる。そこで、3.4節で説明する重み付けの算出の際、誤った情報であったという情報を伝えることで、この問題を解決できると考える。

年齢を例にすると、検索側が認識していた被災者の年齢が実際より少し若かった場合、登録された年齢は、実際の年齢と異なるため、マッチングすることができない。しかし、年齢を一つの数値ではなく、登録された値を中心にした範囲として扱うことにより、少量の誤りがある場合でもマッチングすることが可能となる。

### 3.4 一致可能性と重み付け

上記で説明した動的キー作成と値の集約化を用いることにより、情報に誤りや欠如が含まれる場合でもマッチングを行うことが可能となる。さらに、上記の方法に加えて、マッチングの正確性の向上を図るために、3.1節で想定した情報の各項目ごとに設定した重み付けを使用して一致可能性を示すスコアを算出する。この一致可能性を示すスコアと閾値との大小関係によりマッチングが正確であるかのチェックを行う。各項目ごとの重みの値は、その項目が一致したときに個人を特定できる可能性の高い項目順に高い値を設定する。例えば、性別 < 年齢 < 誕生日 < 所属 < 住所 < カナ < 名前 < 電話番号という順で個人を特定できる可能性が高いと考えて値を設定する。この値を基準とし、3.2節と3.3節で取得した情報を基に重みを算出する。これにより、マッチングの際どの情報に誤りや欠如が含まれるかわからない状況でも適切な重みの設定が可能となると考える。

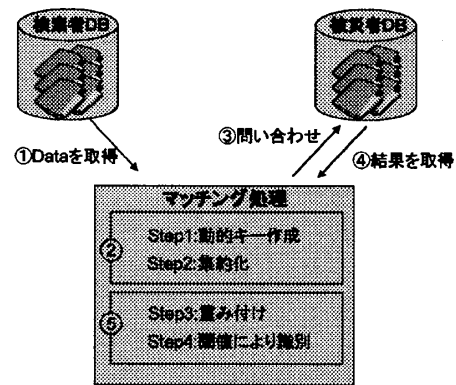


図1: マッチング処理手順の概略図

### 3.5 マッチング処理手順

提案したマッチング処理の概要を図1に示す。検索者DBから取得した情報を基にStep1, Step2で問合せキーを作成し、誤りと欠如の程度を表す情報を取得する。その後、被災者DBに問い合わせを行い、得られた結果を基に、Step3とStep4で信頼性のあるマッチングだけを識別する。以下に各Stepの説明を行う。

**Step1** 検索者DBから取得した検索者が検索したい被災者の情報に対し、動的キー作成を用いて、欠如した情報以外を利用し組み合わせキーを作成する。また、どの項目が欠如しているかという情報も取得する。

**Step2** 作成したキーに対し集約化を用いて、キーとして扱う値をグループに拡張する。また、どの情報の誤りを集約化したかという情報も取得する。

**Step3** 被災者DBへの問合せで得た結果を基に、各項目に設定された基準となる重みの値とStep1とStep2で取得した欠如と誤りの情報を基に、一致可能性を示すスコアを算出する。

**Step4** スコアと閾値の大小関係から正確性の高いマッチングのみを識別する。

## 4 おわりに

本稿では、被災者のプライバシーを考慮した安否確認システムを実現するために、被災者と検索者が登録した情報を双方向からマッチングさせる手法を提案した。またマッチングの際に、災害時に起こりうる、誤りや欠如を含む情報間での正確性を向上させるマッチング方式を提案した。今後の課題を以下に述べる。

- 名前等の非連続値項目の適切な集約化方法。
- 欠如や誤りの項目に応じた動的重み付け方法。
- マッチングの正確性向上方式に加えて、実際の災害時に求められる効率的なマッチング処理の実現。

### 参考文献

- [1] IAA Alliance. IAA システム. <http://www.iaa-alliance.net/>.
- [2] NTT ドコモ. i モード災害用伝言板サービス. <http://www.nttdocomo.co.jp/info/disaster/>.