

モビリティ社会の安全安心に関わる社会課題の俯瞰的評価

○橋本泰一(東京工業大学), 村上浩司(東京工業大学), 乾 孝司(東京工業大学)
内海和夫(東京工業大学), 石川正道(東京工業大学)

1. はじめに

社会の問題が複雑化する中で、俯瞰的なアプローチが必要であることが提案されている[1]。しかしながら、俯瞰的アプローチにおいては、専門性の異なる複数の分野にわたって膨大なテキスト群から有意な情報を取得する作業が必要となり、分析者の主観が強く作用することになる。我々は、情報処理の専門家および STS 分析者との緊密な協力により、汎用性の高いテキストマイニングツール“RiverStone”を開発し、約 10 万件のモビリティ（自動車・交通）分野に関する新聞記事を、語彙分布に基づいて系統的にクラスタリングし、自動要約の手法によりクラスタからトピックを抽出することによって効率的かつ客観的に社会の事象を俯瞰することを可能とした[2]。本報告では、日本社会の深く浸透するモビリティ分野に関わる社会課題を対象として、テキストマイニングの技術によって生成される Disjoint な記事集合群（互いに記述内容は補完的ではあるが、記事の重複がない記事集合群）をもとに社会の対立軸を見出し、これらの社会的な関心の強さを計量的に比較評価することによって、課題解決のための仮説を形成する新しいアプローチについて報告する。

2. テキストマイニングツールによる新聞記事の分析

分析対象となる新聞記事は、2000 年から 2006 年の日研新聞記事のうち同社が提供する日経シーラス¹に含まれる自動車に関連する用語(359 単語)を含むもの(約 10 万件)である。そして、記事が取り扱う話題の類似性を次に述べる方法によって体系的に評価した。1)「類似した内容をもつ記事は、似たような語彙の使用によって特徴付けられる」という考え方のもとに、全記事に含まれる複合語を含む名詞(約 1 万語)について記事中での出現頻度をもとに重み付けし、記事間における語彙出現の類似性をもとに記事をクラスタに分類した(クラスタリング)。さらに、クラスタを階層構造化することでクラスタグループを形成する。2) 構造化された記事クラスタを要約するキーワードを抽出する。3) クラスタ個々について中心性(当該クラスタの他クラスタとの類似性)および密度(クラスタ内における記事内容の類似性)を算出する。4) 次いで、構造化したクラスタグループに所属するクラスタのうち中心性および密度が高く多くの記事数をもつ重要クラスタを特定する。5) 重要クラスタについて重要なキーワードを自動的に抽出し、その周辺文脈から記事の内容を把握し、トピックを抽出する。6) 得られたトピックは補完的ではあるが相互の内容は個性が高いという観点から、記事群の Disjoint な特性に依存する課題(仮説)を分析者の視点を加えて発見する。以上の分析は、我々が開発したテキストマイニングツール“RiverStone”によって予備的な情報を利用することなく効率的に行うことができた。

¹ 日本経済新聞デジタルメディアが 1982 年から作成している新聞記事検索のための用語集。詳しくは、次の URL を参照。http://telecom21.nikkei.co.jp/help/contract/price/00/help_KIJI_thes.html

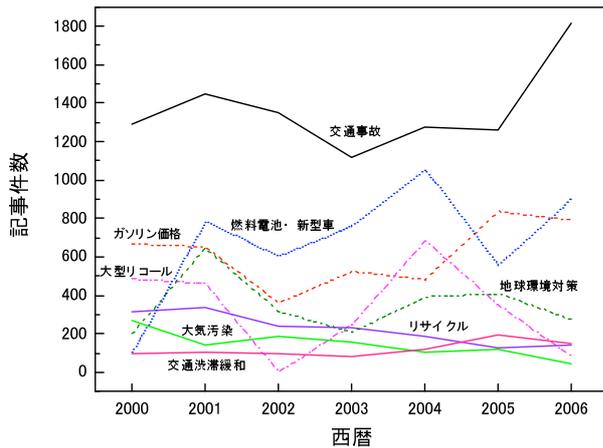


図1 クラスタリングによって得られたトピック毎にみた記事件数の推移(2000-2006年)

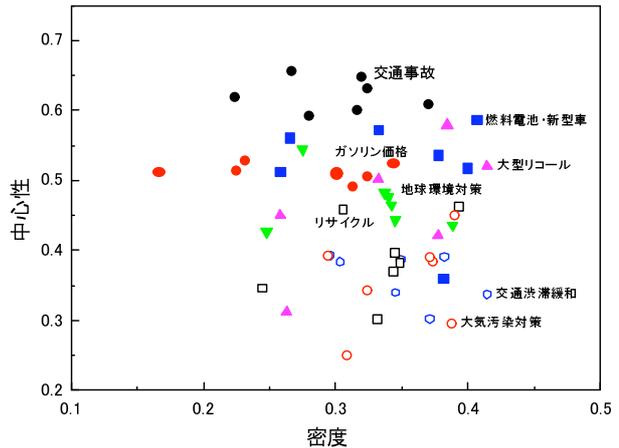


図2 記事クラスタの密度-中心性ダイアグラム(プロットデータは図1に対応)

3. モビリティ分野における社会課題のトレンド評価

2000 から 2006 年の記事を年ごとに 200 のクラスタに階層構造化し、密度および中心性指標に基づいて、交通事故、燃料電池・新型車、ガソリン・貴金属価格、大型リコール、地球温暖化対策、リサイクル、交通渋滞緩和、大気汚染対策のトピックを含む 8 個のクラスタを重要クラスタとして判別した。各トピックについて記事件数の年次トレンドを図 1 に、対応する記事クラスタの密度-中心性ダイアグラムを図 2 に示す。これらの図から、報道件数の増加・減少傾向および取り扱われたトピックの中心性あるいは類似性を評価した結果、次のことがわかった。第 1 に、交通事故に関する記事はモビリティ分野において中心性および報道記事件数が共に高く、最近 7 年間で圧倒的に強い関心事項となっている。第 2 に、燃料電池・新型車の開発に関する記事が交通事故に次いで多く、2002 年以降ガソリン価格および貴金属等の資源価格および地球環境対策に関する記事件数の増大に連動して増加する傾向にあった。この他、大型リコールの話題は年毎に報道の変動幅が大きく、事件が起きた年に大きくクローズアップされ速やかに緩和する傾向があり、大気汚染対策および交通渋滞緩和に関する報道は相対的に関心が低いことが分かった。新聞報道は、事件や話題が発生すると集中して報道記事が増加する傾向にあるため、大型リコールについてはこの傾向が強く現れていた。これに対して、交通事故はほぼ一定して深刻な社会問題と話題にのぼり、自動車の安全対策は重要な社会課題であることが改めて認識された。また、ガソリンおよび貴金属等、素原材料の価格上昇に関する問題が増加しており、日本の自動車産業の持続的発展において新たな社会問題として急浮上しつつある。本発表では、さらにトピック別に情報抽出を行い、これをもとに社会課題の構造化を試みた結果についても報告する。

参考文献

- [1] 堀井秀之 (2006) 安全安心のための社会技術. 東京大学出版会.
- [2] 橋本泰一, 村上浩司, 乾孝司, 内海和夫, 石川正道 (2007) 文書クラスタリングによるトピック抽出および課題発見. 情報処理学会誌. (投稿中)