

論文の内容の要旨

論文題目 チャンス発見のためのコミュニティマイニング
に関する研究
氏名 松村 真宏

チャンス発見とは、意思決定にとって重要な事象を発見し、その意味を理解することである。では、どのような事象が意思決定にとって重要なのであろうか。多くの事例に当てはまる事象、例えば「ユニクロのフリースはよく売れる」というルールを見つけることは、意思決定にとって重要なのだろうか。

世の中の変化は激しく、流行したと思ってもあっという間に廃れてしまう。例えば、当初大変な人気だった小泉首相の支持率は現在低迷しているし、今年ノーベル賞を受賞した島津製作所の田中フェローを取り巻くフィーバーぶりもそう長くは続かないだろう。現在の傾向が近い将来にまで続く保障はどこにもない。したがって、来るべき将来を予測してただ待ち続けることは愚かであろう。となると、我々の採るべき選択肢は一つしかない。自分たちの手で新しい将来を創っていけばいいのである。そのための意思決定に重要な事象の発見と理解こそチャンス発見が目指すものである。ユニクロがフリースを売り出すより以前に「フリースはよく売れる」というルールを発見したのであれば、そのルールは新しい将来を創る価値がある。しかし、フリースが売れてしまってからそのルールを発見しても、もはやそのルールに価値はない。まだ誰にも気づかれていないほど稀だけれども実は重要な事象を発見し、その意味を理解することがチャンス発見なのである。

チャンス発見が扱うドメインや技術は多岐にわたる。例えば、スーパーのPOSデータや医療カルテなどの膨大なデータから、まだ誰も気づいていない売れ筋や検査項目の傾向を発見することや、会話文のようなテキストデータの中からキラリと光る発言を発見することなどが行われている。これらは与えられたデータに重要な事象が潜んでいることを前提としているので、データからその事象を発見するアプローチと言える。しかし、データに重要な事象が潜んでいるかどうかは分からない。したがって、チャンス発見のもう1つのアプローチとして、重要な事象を生み出す環境を構築することが挙げられる。

自分一人であれこれ考えるよりも、誰かと話しているときの方が新しいアイデアが生まれることは我々のよく経験するところである。このように、他の人とコミュニケーションすることによって新しい価値観に気づき、それまで気づかなかった重要な事象に気づくことがある。例えば、ユニクロは安いから買っていると思い込んでいた人でも、他の人と話をしているうちに、実は生地が厚くて洗濯してもヨレヨレにならないところが気に入って

いることに気づいたりする。このような潜在的な価値観は、関心は共有しているけれども価値観は多様な人たちが集うオンラインコミュニティで表出することが多い。したがって、そのような潜在的な価値観が表出するオンラインコミュニティの特徴がわかれば、重要な事象を生み出す環境を構築することが可能になる。

そこで本論文では、チャンス発見のためのコミュニティマイニングをテーマとして、人々のコミュニケーションを通して創出される様々な価値観の発見と、そのようなコミュニケーションを実現するためのオンラインコミュニティのメカニズムの解明に取り組む。

本論文は3部構成からなり、本論文の構成は以下のようになる。

第1章 序論

本論文の目的について述べる。

第I部 ディスカッション分析

第2章 語の活性度に基づくキーワード抽出法

読者が文書を読み進めるときの記憶の活性状態に着目して、文書を読んだ後に読者の記憶に強い印象を残す語をキーワードとして取り出す手法について述べる。

第3章 議論構造の可視化による論点の発見と理解

議論を適度に分割・構造化することにより、議論の論点となる話題をユーザが直感的に把握することについて述べる。また、議論の発展に強く影響を与えた話題を同定する手法についても述べる。

第4章 興味理解のための文書組み合わせ手法

ユーザが興味を満たすために行う文書検索において、複数の文書を組み合わせることでユーザを深い理解へと導く手法について述べる。

第II部 オンラインコミュニティにおける影響伝播

第5章 オンラインコミュニティにおける影響の普及モデル

ある人の発した語への興味が他の人に伝播していくプロセスに着目して、テキストによるコミュニケーションにおける影響の普及を表すモデルを提案する。実際に電子掲示板に適用し、盛り上がる話題を提供しているコメント、オピニオンリーダー、面白い語を取り

出せることを示す。

第6章 オンラインコミュニティ参加者のプロファイリング

オンラインコミュニティ内でやり取りされたコミュニケーションデータから、参加者に大きな影響を与えた語を参加者の特徴を表すプロフィールとして抽出する手法について述べる。

第III部 オンラインコミュニティのダイナミズム

第7章 WWW から新しいトピックの予兆発見

WWWのリンク構造からWebコミュニティとそれらをつなぐ弱い紐帯(Weak ties)を見つけることによって、世の中の大きな変動の予兆を発見することについて述べる。

第8章 2ちゃんねるが盛り上がるメカニズム

メッセージのサイズや投稿数、返信率、投稿される早さなどの基本的な属性に加え、2ちゃんねるに特徴的な名無しと、2ちゃんねるに特有の定型的な表現技法に注目して2ちゃんねるの5748スレッドを分析し、SEMにより2ちゃんねるが盛り上がるモデルを構築した結果について述べる。

第IV部 結論

第9章 まとめ

第I部ではディスカッション分析として、オンラインコミュニティにおいて繰り返し広げられるディスカッションデータから、語の活性値や議論構造の可視化によって重要な事象を発見する手法について述べる。また、興味理解のための文書を組み合わせ手法について述べ、ディスカッションからチャンス発見へつなげる切り口を開拓する。

第II部ではオンラインコミュニティにおける影響伝播として、オンラインコミュニティ中を広まる影響に注目して、オピニオンリーダ、盛り上がった話題、参加者のプロフィールを発見する手法について述べる。オピニオンリーダを発見するだけでなくそのプロフィールまで獲得できることで、オンラインコミュニティをより深く理解できるようになる。

第III部ではオンラインコミュニティの参加者が織り成すダイナミズムがオンラインコミュニティの盛り上がりにも及ぼす影響について述べる。WWWのリンク構造の変化からセレ

ラ社の台頭を検証し、2ちゃんねるが盛り上がるメカニズムを解明する。ここで得られた知見は潜在的な価値観を創出するオンラインコミュニティのデザインに役立つ。

昨今では、オンラインコミュニティは誰もが気軽に参加できるコミュニティとしての地位を確立し、我々の生活にも浸透しつつある。本研究で得られた様々な知見が、多くの人にとってオンラインコミュニティが価値あるものとして役立つことに貢献できれば幸いである。