

## 口コミ情報解析による情報遷移の把握に関する一考察 Grasp Method of Information Transition by Analysis of Word-of-mouth Information

菅野 元之† 菊池 寿徳† 佐野 直美† 齊藤 典明† 高見 一正†

Motoyuki SUGANO Toshinori KIKUCHI Naomi SANO Noriaki SAITO Kazumasa TAKAMI

### 1. はじめに

インターネットの普及に伴い、口コミ情報に代表されるユーザから配信される主観的な情報は、マーケティング等の分野では量的にも影響力の面からも重要視されつつある。

本稿では、インターネット上に存在する口コミ情報を収集・解析することで、世間の評判やトレンドの把握や変遷の追跡等を可能とする口コミマイニングを提案し、実現例として口コミマイニングの機能構成例と解析例を示す。

### 2. 研究の背景

従来、実世界において意見や感想、評判等の主観的な情報は口コミ情報として、主に友人や知人間のコミュニケーションを介して世間に流通していた。しかし、各個人が不特定多数の人に対して情報を発信できるインターネット環境の普及に伴い、口コミ情報の伝播方法も変化し、友人や知人以外の第三者を含む特定多数のユーザが配信した口コミ情報がインターネット上に流通していると考えられる。

口コミ情報の中でも、商品に関する意見や感想、評判等の情報はユーザの購買行動に影響を与える情報として特に重要視されている。商品情報には、カタログ等に記載されている商品スペックや機能等の客観的な情報と、口コミ情報に代表される主観的な情報の主に二種類が存在する。しかし、インターネットショップにおいて、ショップによるレコメンドと比べて、ユーザからの生の声である口コミ情報は約7倍の効果があるという報告[1]があるように、主観的な情報である口コミ情報はショップが行う宣伝等の情報とは異なるユーザが感じた生の情報であり、スペックからは読み取ることができない商品の使い勝手や快適さ等の商品の感想を含む情報としてその有用性が認知されつつある。

企業やメーカーもまたインターネット上における口コミ情報を重視しつつあり、ユーザの関心やトレンド等を把握するために、BBS等のWebサイト上の情報に関する調査・分析も行われつつある。例えば、ガーラでは人手によるインターネット上の主要サイトに関するインターネット情報モニタリングサービスがすでに行われている。しかし、これらの調査はあくまで、一般のWebサイトを対象とした調査を目的としているため、自動化は特定の商品に関連するページ数等の単純な統計情報のみであり、テキストマイニング等の文章の内容まで踏み込んだ分析や解析の自動化は行われていない。

以上より、インターネット上に流通する口コミ情報は世間を表わす指標の一つであると考えられる。さらに、この口コミ情報の解析は、特定商品に関する悪評が盛り上がるの傾向を把握し、悪評伝播の前兆として知らせる悪評検出サービス等を可能とするなど、重要である考えられる。そこで、我々はインターネット上の多くサイトにおいて、ユ

ーザがリアルタイムに配信している口コミ情報を収集し、その中に埋もれているユーザの最新の関心やトレンド等を分析、さらにその変遷情報を解析する口コミマイニングという概念を提案する。

本稿では、口コミマイニングの一例として、インターネット上に存在する口コミ情報に対してテキストマイニングにより抽出された商品情報や評価情報等と口コミ情報が持つ時間、ユーザ、サイトとの付加情報を元に、商品に関する評判情報の変遷、話題の盛り上がりを解析する口コミマイニングの機能構成を明らかにする。

### 3. 関連研究

口コミ等の評判情報の検索に関して、立石らの評判情報検索[2]、二本木らの口コミ評価の分析・検索[3]の研究が行われている。評判情報検索では、インターネットから収集した情報から近接演算や適正値判定処理を用い評判情報の抽出を行い、発見された評判情報の肯定・否定を元に統計処理を行うことで、重要と考えられる評判情報を発見することを可能としている。しかし、検索した評判情報の解析は肯定・否定にとどまり、評判情報の変化等の解析については行われていない。また、口コミ評価の分析・検索についても主眼が評判情報の検索に置かれており、評判情報自体の分析は行われていない。

口コミ情報の分析に関して、立石らの意見収集分析システム[4]の研究が行われている。意見収集分析システムでは、抽出された意見情報に対してテキストマイニングを行っている。しかし、時間情報やユーザ情報、サイト情報等の評判情報に付随する情報が解析に考慮されておらず、本研究が目的としている情報の変遷や動向の把握等を解析にそのまま適用することはできない。

### 4. 口コミマイニング

#### 4.1 解析情報

口コミマイニングでは、解析対象である口コミ情報を以下の3つの情報にて捕らえる。

##### (1) 文章情報

口コミとして配信されているテキスト情報であり、形態素解析を行うことで単語や品詞情報を抽出、さらにテキストマイニングを行うことで話題や評価情報を抽出することが可能である。

##### (2) 時間情報

口コミ情報が配信された日時等を表わす情報である。現在時刻やイベント情報との比較により口コミ情報の新鮮度等を分析することが可能となる。

##### (3) ユーザ情報

口コミ情報を配信したユーザとそのサイトを特定する情報である。サイト情報やユーザプロフィール、口コミの履歴等を用いることで口コミ情報の信頼度等を分析することが可能となる。

† 日本電信電話株式会社

NTT 情報流通プラットフォーム研究所

時間情報とユーザ情報については、サイトによって該当情報が存在しない場合がある。その場合、時間情報は情報の取得時間にて代替し、ユーザ情報は各サイト間もしくは各口コミ情報間で一意となるユーザを代替として割り合てることで対応する。

口コミマイニングでは、文章情報、時間情報、ユーザ情報を3つの情報を軸とし、各情報から分析可能な様々な情報を組み合わせることで解析を行う。表1に各情報から分析可能な情報の例を示す。

表1: 解析情報例

文章情報	
単語	
品詞	
記事数	
単語数	
商品情報	
評価情報	
ポジティブ/ネガティブ発言数	
時間情報	
特定期間(一週間等)	
全期間	
イベントからの経過時間	
ユーザ情報	
ユーザ配信件数	
配信ユーザ数	

4. 2 口コミマイニング機能構成例

口コミマイニングの機能構成例を図1に示す。口コミマイニングは、口コミ情報収集、文章情報解析、口コミ情報解析の3つのフェーズから構成される。

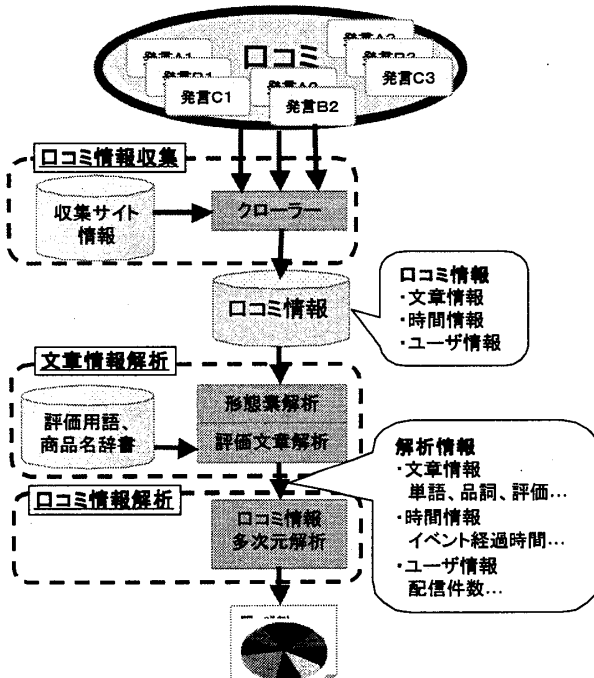


図1: 口コミマイニング機能構成例

4. 2. 1 口コミ情報収集フェーズ

口コミ情報収集フェーズでは、予め指定されている収集サイト情報を元に、クローラを用い BBS 等の Web サイト

から口コミ情報の収集することを想定する。また、口コミ情報の時間情報、ユーザ情報も合わせて収集する。

4. 2. 2 文章情報解析フェーズ

文章情報解析フェーズでは、収集した文章情報から解析情報の抽出を行う。形態素解析では、文章情報から名詞を抽出、複合語化し単語を抽出する。評価文章解析では、名詞、形容詞等から構成される評価表現パターンとマッチングを行うことで、評価情報の抽出し、ポジティブ/ネガティブ判定を行う。

4. 2. 3 口コミ情報解析フェーズ

口コミ情報解析フェーズは、3つの解析処理から構成される。以下に各処理にて可能な解析例を示す。

(1) 一次元解析処理

口コミマイニングの3つの軸に関するユーザ数や商品名の出現回数などの基礎的な統計情報を出力する。

(2) 二次元解析処理

・認知度

配信ユーザと商品名の情報を元に、各商品について話題としたユーザ数を分析することで、商品の認知度を解析する。

・話題の盛り上がり

商品名の出現回数と時間情報を元に、各商品について話題が集中した時期を解析することで、定常的な話題を除去し、話題の盛り上がりを検出する。

・流行語

商品名の出現回数と時間情報を元に、高い出現回数を長い期間連続して継続した商品名を抽出することで、特定期間の流行語を抽出する。

(3) 三次元解析処理

・話題の浸透度合い

一つの話題に関する複数の盛り上がりに対してユーザ情報を加え時間軸にて解析し、各話題の盛り上がりに関するユーザ構成の変化を解析することで、多くのユーザに浸透しつつある話題、限定されたユーザ内での盛り上がり等を解析する。

・重要発言者

複数の話題の盛り上がりについて、その発端となる記事の配信ユーザの集合を特定し、過去の履歴と比較することで話題の発端への関与率を解析し、重要発言者を抽出する。

5. おわりに

本稿では、口コミマイニングの概念を提案し、実現例として、口コミマイニングの機能構成例と解析例を示した。今後は、口コミマイニングの有効性について検証を行う。

参考文献

[1] goo リサーチ, <http://research.goo.ne.jp/>, 2001.  
 [2] 立石 賢二ほか, “インターネットからの評判情報検索”, 情報処理学会研究会報告, NL-144-11, pp.75-92, 2001.  
 [3] 二本木 智洋ほか, “文の構造化による口コミ評価の分析・検索”, インタラクシオン 2002 論文集, pp.175-176, 2002.  
 [4] 立石 賢二ほか, “Web 上の自動意見分析 -情報抽出とテキストマイニングの融合-”, 情報処理学会第 64 回全国大会, pp.3-19 - 3-20, 2002.