RF-004

テレビ番組に関するコメント解析手法

Message Analysis Algorism about TV Programme

有安　香子†　　藤沢　寛†　　金次　保明†
Kyoko ARIYASU　Hiroshi FUJISAWA　Yasuaki KANATSUGU

1. はじめに

ソーシャルネットワークサービス（SNS）やブログの普及により，個人意見の発信の場としてのインターネット空間は円熟の境に入り，誰もが意見を発信できる表現の場から「意見の共有」の場へと進化し続けている．一方，放送は多くの視聴者が時を同じくして，同じ内容を享受する「視聴経験の共有」の担い手としての特徴を持つ．我々は，「意見の共有」の場としての通信と，「視聴経験の共有」の場としての放送を，一元的に扱う情報空間をコンセプトとした情報還流システムを提案している．本稿では情報還流システムの核となる，放送番組に関する意見共有のための SNS メッセージの解析手法の提案と，実際の SNS メッセージを用いた検証実験の結果について報告をおこなう．

2. 関連研究

放送と通信の利点を活かした，放送通信連携のための研究プロジェクトが，近年各国で盛んにおこなわれている．異なるネットワーク間のコンテンツを一元的に扱う“iNEM4U [1]”，セマンティック Web の概念をテレビに導入した“Notube [2]”，BBC を中心に複数のテレビ局やメーカーが協力関係を築き開発を進める“YouView[3]”，通信を用いて放送サービスを強化する“Hybridcast™ [4]”の研究など，プロジェクトによりその特徴はさまざまである．どのプロジェクトも，視聴者のテレビ視聴環境をより豊かにするための個人向けサービスに力をいれている．本稿で論じる情報還流システムは，Hybridcast 開発の一端として，より豊かな放送通信連携サービスを提供するために用いられる．情報還流システムは，テレビ番組放送中に投稿された視聴者の意見や感想（以下コメントと表記する）を入力とし，そのコメント内容の解析結果から個人向けサービス生成するシステムである．

本稿で提案するコメント解析手法に関連する研究として，トピック抽出と感情分類に関するものがあげられる．SNS メッセージのように短い文章からトピック抽出をおこなう従来研究には，携帯端末から得られるさまざまな情報で足りない情報を補完する[5][6]などがある．提案手法は，放送局の番組関連情報を用いて足りない情報を補完し，コメントのトピック抽出をおこなっている．感情分類に関する従来研究は，映画レビューなどの文章内容を解析する[7]，Twitter[8]のような短い文章を感情分類する[9][10]，フレーズ単位で分類する[11]などがあるが，解析対象を番組に関するコメントに絞ったものはまだない．本稿では，番組に関するコメントを解析する際の独自の問題点を洗い出し，これを解決するための解析手法について論じる．

† NHK 放送技術研究所, Science & Technology Research Laboratories JAPAN BROADCASTING CORPORATION

本論文の構成は以下の内容とした．3 章において情報還流システムの概要とシステム構成を説明し，コメント解析の前提となる背景部分を説明する．4 章において視聴者の番組に関するコメントを言語処理する際の問題点の洗い出をおこない，本稿主題である 5 章では番組に関するコメントの解析手法の提案，6 章では実際の Twitter のつぶやきを用いた実験結果を記した．7 章において実験結果の考察をおこない，8 章にまとめと今後の課題を記した．

3. 情報還流システム

Figure 1. 情報還流システム概念図

Figure 1.に情報還流システムの概念図を示す．情報還流システムは，番組コメントの解析結果を用いて，視聴者・放送局双方に利益をもたらす，情報空間の提供をコンセプトとしている．放送局は，テレビ番組に関する視聴者の反応を，コメントの傾向を示す解析結果表示グラフ[12]から知ることができる．視聴者向けサービスには，投稿数の盛り上がりを基にした漫画風ダイジェスト[13]，視聴者の興味度を基にしたお勧め番組サービス[14]，携帯端末や放送通信連携型テレビなどから気軽に参加するための共感グラフサービス[15]などがある．

コメント解析に求められる要件はサービス毎に異なる．番組制作者が広く視聴者の反応を知るために，番組のどの部分でいくつコメントがあったかを基に，解析結果表示グラフが生成される．盛り上がりをもとにした漫画風ダイジェストでも，番組のどの部分に対していくつコメントがあったかを基にサービスを生成ため，より多くのコメントが処理されることが重要となる．

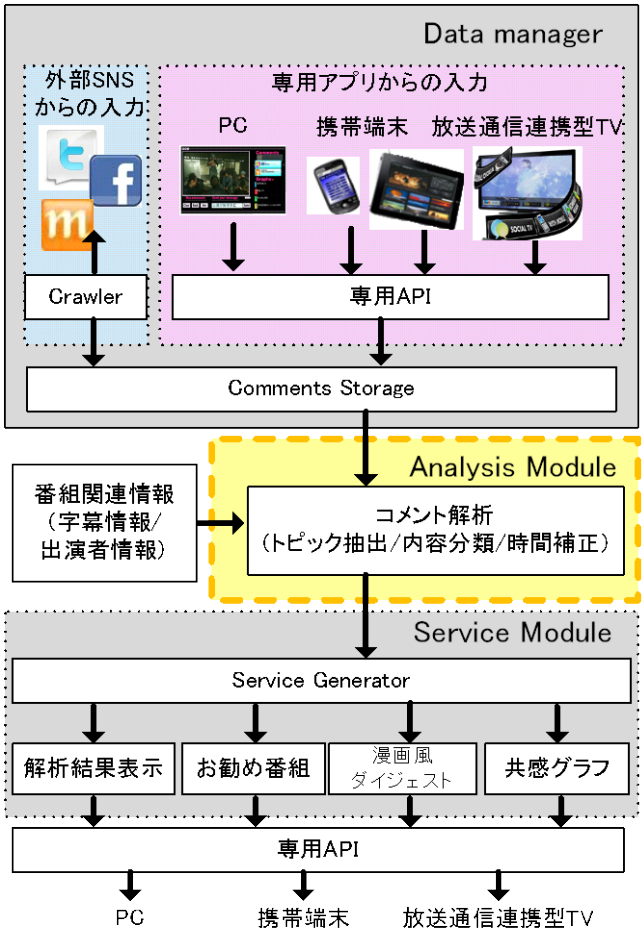


Figure 2. 情報還流システム構成図

一方，コメント内容から興味度を算出するお勧め番組サービスでは，誰に対するどのような内容のコメントかを基にお勧め番組を決定するため，解析精度の高さが重要となる．

また，キーボード入力が困難な携帯端末や放送通信連携型テレビから，意見共有を楽しむための共感グラフサービスでは，自分の気持ちを投票するタイミングを逸しないよう，処理時間の早さ（処理の軽さ）が重要となる．

我々は，情報還流システムとしてコメント解析に求められるこれらの要件と，Figure2.に示すシステム構成を考慮したコメント解析手法を設計することとした．

情報還流システムは，データマネージャー・解析モジュール・サービスモジュールから構成される．データマネージャーは，さまざまなデバイスからのコメント入力を Rest 形式の専用 API[16]を介して受け取る．また，外部 SNS の検索 API を介して入力対象となるコメントの収集をおこなう．各コメントは，時間情報・ユーザ ID・コメント内容を要素としデータベースに蓄積される．

解析モジュールは，収集蓄積された番組コメントを，放送中の番組に付随する字幕情報と関連する出演者情報を補助情報として，5 章に記すコメント解析手法を用いて解析する．

サービスモジュールは，解析モジュールから渡される解析結果を用いて個人向けサービスを生成し，専用 API を介して各種デバイスにサービスをフィードバックする．

4. 番組コメントの特徴

SNS メッセージは元来，コミュニケーションをその主目的としているため，口語調が多く，文法的な誤りや省略を多く含む．SNS ユーザは，それらの文法的な誤りを無意識に修正し，表記方法の醸し出す雰囲気からユーザの人となりを推し量ることができる．しかし，これらの人間の認知を通した微妙な解釈と推論をシステムでおこなうことは不可能であるため，その特徴を加味した処理方法の設計が必要となる．以下に，SNS メッセージの中でも特に，テレビ番組放送中に投稿されるコメントに焦点を絞り，特徴の抽出とコメントを言語処理する際の問題点を記す．

4.1 主語の省略

番組を観ながら投稿されるコメントは，視聴者が同じ時間に同じ画面を見ているため，主語が省略されることが多い．主語を省略しても前後の文脈から，何に関するコメントか容易に推測できるため，入力の時間を短縮するなどのメリットから主語が省略されると考えられる．特に，有名なドラマやスポーツ番組などの人気番組が対象の際には，全体のコメント数が多いので，タイミングを逃さないよう，主語が省略されるケースが多くみられる．

4.2 表記ゆれ

SNS ユーザは，自分の属するコミュニティの雰囲気を醸し出すため，特有の表現を用いてコミュニケーションをとることが多い．また，多少の誤字・脱字があっても，読む側が意図を忖度することが好ましく，訂正などの指摘は雰囲気を壊すとされている．また顔文字などの絵文字や，日本語特有の漢字・ひらがな表記，半角・全角の混在などの表記の種類が多数にあること，コメントを目立たせるために語尾を繰り返し表記すること，などに起因する表記ゆれが他の Web の文章に比べ極端に多く，一般的な言語処理を困難にしている．

4.3 入力時間遅延

コメント入力にかかる時間の分だけ，時間遅延が生じる．番組のコンテキストを理解している視聴者は 2 分から 5 分遅れてもコメントの指し示す内容を理解できるため，入力速度はあまり問題視されないが，システムでは文脈を解釈できないため解析処理における大きな問題点となる．

5. コメント解析手法

この章では，4 章であげた番組コメントを言語処理する際の問題点に関する解決策を述べる．まず 5.1 で主語省略に関する解決法をトピック抽出アルゴリズムとして述べ，5.2 で表記ゆれに関する解決法を内容分類アルゴリズムとして述べる．次に 5.3 で，個々の解決法では解決できないコメントの処理手法として，コメント間の類似性を用いた再処理アルゴリズムを述べ，更に 5.4 で，これらの処理結果を用いた入力時間補正アルゴリズムについて述べる．

提案するコメント解析手法には 2 つの特徴がある．1 つ目は，データ放送に使用する番組関連情報を用いて，足りない情報を補完することである．具体的には，データ放送送出時に放送電波内の Transport Stream に格納される，字幕情報と EPG（Electric Program Guide）内の番組関連情報を用いて情報を補完する．字幕データは，字幕を表示するタイミング，字幕話者，字幕内容テキストから構成されている．また，処理対象番組の EPG 電子番組表に含まれる番組概要と出演者情報を抽出し，出演者情報として使用する．対象番組がドラマの場合，出演者情報として，役名・役者名を含み，対象番組がスポーツの場合，チーム名・選手名・アナウンサー名・試合場所などの情報を含む．

提案するコメント解析手法の 2 つ目の特徴は，コメント同士の類似性と時系列性を使用した再処理アルゴリズムである．投稿されたコメントは情報還流システムの中で時間順にソートされ蓄積される．同じ番組を観ながら同じ時間に投稿されたコメントは，その内容が類似する可能性が高いという類似性を使い，コメントに含まれる単語をキーとして，解析をおこなうこととした．

5.1 トピック抽出アルゴリズム

トピック抽出アルゴリズムは，番組関連情報からトピックとなり得る候補リストを作成し，この候補リストを基にコメントの特徴に応じてトピック抽出をおこなう．主語が表記されているコメントは 5.1.2 に記載の処理，字幕情報との類似性があるコメントは 5.1.3 に記載の処理，両者により処理できなかったコメントは 5.3 に記載の処理によりトピック抽出がおこなわれる．

5.1.1 トピック候補リストの作成

EPG の番組関連情報を用いてトピックの候補となるリストを作成する．番組概要文のすべての名詞を抽出し，出演者情報と合わせ，トピック候補リスト作成する．このリストがない場合，コメントに出てくるすべての名詞をトピック候補として扱わなければならないため，トピック候補リストを用いることで，処理時間の大幅な短縮が可能となる．

5.1.2 トピック候補を含むコメント

コメントがトピック候補リスト内の名詞を含む場合，その単語をトピックとして抽出する．複数の候補がある場合には，最初の一致名詞をトピック候補とした．

5.1.3 字幕情報を用いたトピック抽出

コメントがトピック候補リスト内の名詞を含まない（5.1.2 によって処理されなかった）場合，字幕内容テキストとコメントの類似性を用いてトピック抽出をおこなう．字幕の時間情報より後に投稿されたコメント内の単語と，字幕内容テキスト内の名詞・形容詞・形容動詞・副詞・動詞を比較し，一致したら字幕話者情報をトピックとして補完する．この処理はドラマや情報番組のように，動作主と話者が一致する類の番組において，有効である．

スポーツ番組では，字幕話者はそのほとんどがアナウンサーと解説者であるため，字幕内容中に 5.1.1.で作成したトピック候補リスト内の単語が出てきた場合，その単語を話者に置き換える前処理をおこない，動作主と話者データを可能な限り一致させることとする．

5.2 コメント内容分類アルゴリズム

SNS に投稿されるメッセージは，一般的な文章に比べ 4.2 に記した表記ゆれを多く含むため，一般的な言語解析に適さない．SNS メッセージの内容を解析する研究がいくつかなされてはいるが，単語の特徴量を用いて「肯定」「否定」の 2 値に振り分けるものが多い[9][10].

一方[17]では，ドラマ番組の感想における特徴語の抽出実験をおこなっており，ドラマに関するコメント文中の頻出単語として「辛い」「考える」「泣く」「ハラハラする」「ドキドキする」など，肯定・否定に分けにくい単語を頻出単語として挙げている．本アルゴリズムでは「辛い」を“同情”，「泣く」を“悲嘆”，「ドキドキする」を“期待”に分けるなど，対象番組のジャンルに合わせ，コメント内容をいくつかのカテゴリに分類することとした．対象番組がドラマであれば，“肯定”“感嘆”“悲嘆”“失笑”“否定”“同情”“期待”など，スポーツであれば“状況解説”“肯定”“応援”“不安”“否定”“要望”など，番組内容に見合ったカテゴリに分けるアルゴリズムである．

本アルゴリズムでは，はじめにコメント内容をカテゴリに分類するための辞書を作成する．辞書作成用の学習データとなる番組コメントを集め，コメントに含まれる単語の特徴量をカテゴリ毎に算出する．単語の特徴量は，単語の出現頻度と重要度を表す TF-IDF を使用する．次に，特徴を示す単語を列方向に並べ，行方向にカテゴリを並べたマトリクスの各要素を，行方向に正規化した各単語の TF-IDF 値で埋めた分類辞書マトリクスを生成し，これを分類辞書として使用する．このようにして得られた分類辞書を用いて，入力コメントをその構成単語によりカテゴリに分類する．5.1 のトピック抽出アルゴリズムと同様に，5.2 により処理できなかったコメントは 5.3 に記載の処理により再処理される．

5.2.1 TF-IDF の計算

分類辞書を作成するため，学習データを形態素解析器にかけ，単語ごとに分割する．過学習を避けるため，すべての出現単語のうち，出現回数の極端に少ない単語（出現回数 2-5 回）を除去し，残った単語のうち固有名詞や場所を除く名詞・形容詞・形容動詞・副詞・動詞を選び，下記の要領で TF-IDF を計算する．

対象番組を k 個のカテゴリC_jに分類する場合，単語 i のカテゴリC_jにおける出現数をn_i，全カテゴリにおける単語 i の出現数の合計を$\sum_k n_k$，全コメント数を｜D｜，単語 i を含む全カテゴリ中のコメント数を$|\{d: d \ni t_i\}|$とした時カテゴリC_jの単語 i の TF-IDF は(1)式により算出される．

$$tf_{C_{ji}} = \frac{n_i}{\sum_k n_k}$$

$$idf_i = \log \frac{|D|}{|\{d: d \ni t_i\}|}$$

$$\text{TF-IDF} = tf_{C_{ji}} * idf_i \qquad \cdots(1)$$

$tf_{C_{ji}}$はカテゴリC_jでの単語 i の出現頻度を表し，IDF はコメント全体での単語 i の重要度を表す．

5.2.2 TF-IDF 辞書の作成

辞書作成の対象となるすべての単語の TF-IDF を計算し，単語×各カテゴリの行列を作る．過学習を防ぎ処理を軽量化するため，すべてのカテゴリの TF-IDF が 0.1 を下回る

単語を除外する．単語ごとの TF-IDF の総和が 1 となるように正規化した辞書を生成する．

$$dic(C_{ji}) = \left|tf * idf_{(C_{ji})}\right| = \frac{tf*idf_{(Cji)}}{\sum_{j=1}^{k} tf*idf_{(Cji)}} \quad \cdots(2)$$

5.2.3 コメントの内容分類アルゴリズム

(2)式の結果を用いてマトリクスを作り分類辞書とする．入力コメントが a 個の単語からなる時，入力コメントを構成する各単語を W_h (h=[1-a])と表現することとする．単語 W_hに対応する各カテゴリの$dic(C_{jW_h})$を辞書から選び，入力コメントを構成する全ての単語に関する$dic(C_{jW_h})$ を，カテゴリ毎に合算する．

$$sum(C_jW_h) = \sum_{h=1}^{a} dic(C_{jW_h}) \quad \cdots(3)$$

(3)式の値が最も大きいカテゴリ C_jを，入力コメントの分類結果｜C_j｜とする．

$$|C_j| = \max(sum(C_jW_h))_{j=1\ to\ k} \quad \ldots(4)$$

5.3 類似性による補完アルゴリズム

5.1 に述べたアルゴリズムによりトピックが決まらなかったコメント，および 5.2 に述べた手法により内容分類できなかったコメントに関する再処理手法を Figure3.に示す．

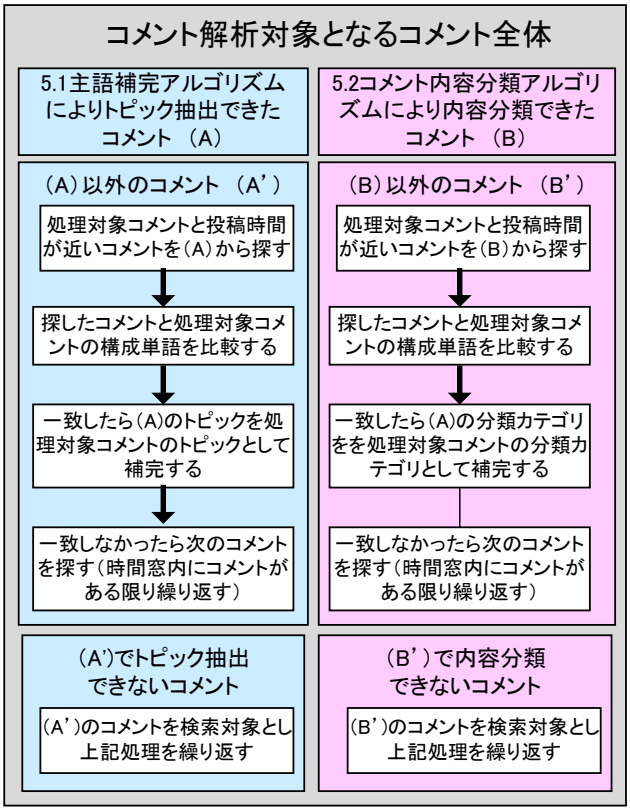


Figure3.　類似性による補完アルゴリズム

このアルゴリズムは，時系列に並んだコメントの特徴を活かし，時間的に近く，コメント構成単語が似ているコメント同士は，同じトピックもしくは同じ内容を含有している確率が高いという時系列性と類似性を基に補完をおこなうものである．5.1 および 5.2 でトピック抽出や内容分類ができたコメントと投稿時間の近いコメントを探し，構成単語が一致したらトピックおよび分類カテゴリを補完する．本アルゴリズムで用いる「時間的に近い」とは，処理対象コメントの投稿時間を中心とした時間窓内に投稿されたコメントとする．係数 b を，直前 m 秒間に投稿されたコメント数/n （n＝定数）とし，

$$\text{時間窓}\ \alpha\ (\text{秒}) = \begin{cases} \frac{m}{2} & 1 \geqq b\ \text{とき} \\ m & 0.4 \leqq b < 1\ \text{のとき} \\ 2m & b < 0.4\ \text{のとき} \end{cases} \quad \cdots(5)$$

のように時間窓を決定し，処理対象コメントの投稿時間から前後 α(秒)内のコメントを検索対象とする．定数 m と n は解析モジュールを実装するサーバの処理速度に応じて，適宜設定するものとする．

5.4 入力時間遅延補正アルゴリズム

入力時間の補正には字幕データの時間情報を使用する．処理対象とするコメントは，トピック抽出または内容分類の処理がおこなわれているものとする．Figure4.に時間補正アルゴリズムを記す．

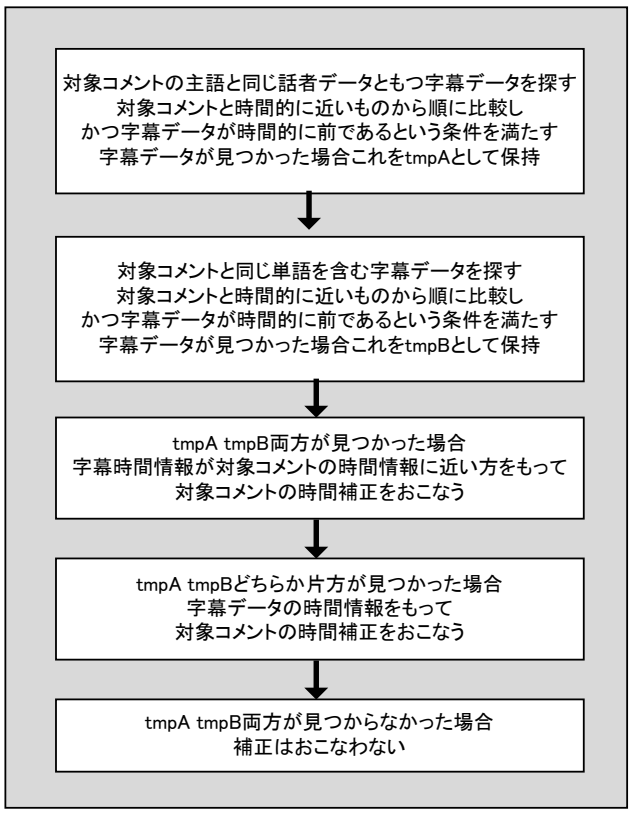


Figure4.　時間補正アルゴリズム

6. Twitter を用いたコメント解析手法検証実験

本章では，実際の Twitter のつぶやきを入力として集め，5 章で提案手法の検証実験をおこなった結果を記す．

実験対象番組として 2010 年 6 月におこなわれたワールドカップ本戦・日本戦 4 試合分のつぶやきを，TwitterAPI を用いて収集した．まず，#worldcup ハッシュタグを用いてつぶやきを集め，集めたつぶやきに併記されたハッシュ

タグを用いて再帰的につぶやきを収集した．Twitter ユーザの中には，より多くの人にコメントを見てもらうために，複数のハッシュタグを並列表記する人が少なからず存在する．（例：ごーーーーーーーーーる！#worldcup #Japan #FIFA2010） これらの並列表記されたハッシュタグを使い，再帰的に検索をかけることで，地域によるテレビ局名表記の違いなどを気にせず，ワールドカップに関するつぶやきを集めることができる．その後，複数ハッシュタグの並列表記によるつぶやきデータの重複を除去し，RT(Retwitte)や QT(QuotedTweet)と呼ばれるつぶやきも除去した．RT や QT には公式・非公式により様々な使われ方をしているが，誰かのコメントに対する返信を追加する形で使用しているものが，収集したつぶやきに多く含まれた．これらのつぶやきは私信を多く含み，番組に内容から逸脱した内容を多く含むため，今回の実験対象から除外した．さらにＥＰＧデータから得られた番組開始時間と終了時間の中に含まれるつぶやきだけを残し，実験に使用した．

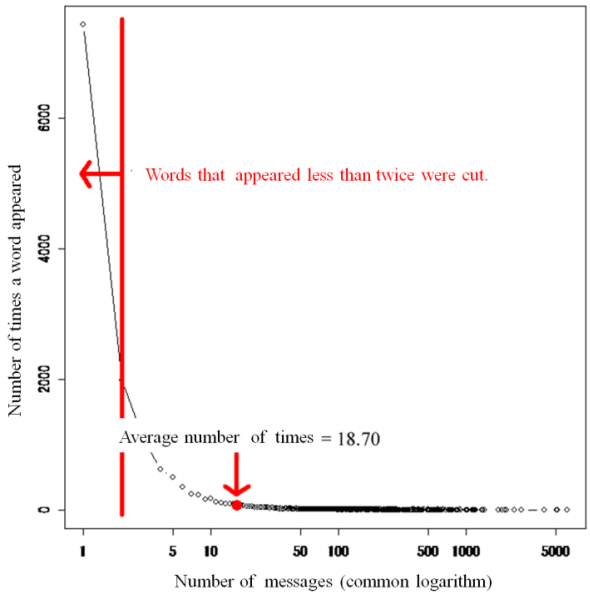
このようにして集めたコメントのうち，3 試合分を使い，TF-IDF 辞書を作成し，残りの 1 試合のコメントで実験をおこなった．番組関連情報は，番組概要・番組放送時間・字幕情報をデータ放送の Transport Stream から抽出し，使用した．また，時間窓の定数パラメータは m=20，n=50 とした．

6.1 コメント内容分類辞書の事前作成

6.1.1 TF-IDF の計算

Table 1. 学習データに用いた試合に関する情報

対戦相手	対戦日	勝敗	つぶやき数
オランダ	6/19	勝	73,061
デンマーク	6/25	負	54,567
パラグアイ	6/29	負(PK)	55,196



Graph 1. 学習データ中の各単語出現回数と出現コメント数の関係

Table1 に記したワールドカップ日本戦に関するデータを用いて TF-IDF 辞書を作成した．3 試合・182824 コメントから 5 人のアノテータが 50000 コメントをランダムに抽出し，各コメントを「状況解説」「肯定」「応援」「不安」「否定」「要望」の 6 つに分類した．試合の流れや結果によりコメントの内容により生じる偏りを防ぐため，試合結果の異なる 3 つの試合から平均的に学習データを抽出した．

学習データは 597,606 単語からなり，14,837 種類の単語を含み，各単語の平均出現確率は 18.7 回であった．Graph 1.に各単語の出現回数と出現コメント数（X 軸は対数）を示す．過学習を避けるため，出現頻度の極端に少ない単語（本稿では 2 回以下）9309 種類を計算対象から除外することとする．TF-IDF 計算対象となる 5528 種類の単語から，名詞（固有名詞・場所を除く）・形容詞・副詞・形容動詞・動詞（2241 種類）に関して，カテゴリ毎に割り当てられたコメントに含まれる単語の TF-IDF を算出した．カテゴリ毎の出現率の差が少ない単語（各カテゴリの特徴的な単語でない単語）の TF-IDF は小さくなるので，算出した TF-IDF が閾値（本稿では 0.1 以下）の単語を除外し．これらの処理をおこない特徴的な 424 単語の TF-IDF を用いて辞書を作成した．

6.1.2 TF-IDF による分類辞書の作成

Table 2. TF-IDF 辞書例

	状況説明	肯定	応援	不安	否定	要望
攻める	0.7608	0	0	0	0.2391	0
危ない	0.1472	0.2182	0.1065	0.4162	0.0761	0.0353
かっこ	0.1008	0.7478	0.0525	0.0588	0.0147	0.0252
うるさい	0	0	0	0.1040	0.8959	0

6.1.1 により得られた 424 単語の TD-IDF 値を行要素に，カテゴリを列要素にとり正規化し，TF-IDF を用いた分類辞書行列を作成した．Table 2.に辞書の例を記す．

6.2 トピック抽出結果

Table 3. 処理工程別正解率

処理工程	正解数	処理数	正解率
Step 1	14,678	15,462	0.9492
Step 2	273	876	0.31
Step 3	18,022	30,551	0.5898
合計	**32,973**	**46,889**	**0.70**

2011 年 6 月 13 日におこなわれたワールドカップカメルーン戦の放送中のツイッターのつぶやき（54,767 個）を用いてトピック抽出アルゴリズムの実験をおこなった．5 人のアノテータが放送番組を観ながらつぶやきを読み，文脈からトピックが一意に決められる 46,889 個のつぶやきの正解データを作成した．Table 3.に処理工程別の処理つぶやき数と正解数，正解数を処理つぶやき数で割った正解率を記す．表中の正解数とは各処理ステップで処理されたつぶやき中，正解データとトピック抽出結果が一致したつぶやきの数を記したものである．

Step 1 は，5.1.2.に記したトピック候補リストの単語がつぶやき中に出現する場合，その単語をトピックとして抽出する処理をおこなう工程，Step 2 は，5.1.3.に記した字幕と単語が類似するつぶやきのトピックを，字幕話者で補完する例である．Step 3 は，5.3 に記した時間が近接して投稿さ

れたつぶやきとの類似性を用いてトピックを推定した結果である．

6.3 TF-IDF を用いた分類辞書による内容分類結果

Table 4. カテゴリ別適合率及び再現率

カテゴリ	処理	正解	一致	適合率	再現率
状況解説	8640	9935	7156	0.828	0.720
肯定	24928	21685	21679	0.8696	0.9972
応援	3662	5686	3275	0.894	0.575
不安	5006	7616	4131	0.825	0.542
否定	273	2665	160	0.58	0.06
要望	1350	3085	1238	0.917	0.401
合計	**43859**	**50672**	**37639**	**0.85**	**0.74**

ワールドカップカメルーン戦の放送中のツイッターのつぶやきを用いて TF-IDF を用いた分類辞書による内容分類アルゴリズムの実験をおこなった．5 人のアノテータが放送番組を観ながらつぶやきを読み，文脈から内容が一意に分類できる 43,859 個のつぶやきの正解データを作成した．Table 4.に分類結果別の処理つぶやき数と正解データ数，一致数を示す．表中の正解データとはアノテータが作成した，各カテゴリの正解データの数，処理数とは提案手法により各カテゴリに分類されたつぶやきの数，一致数とは処理数中正解データと分類が一致したつぶやきの数を示す．一致数を処理数で割ったものを適合率，一致数を正解データ数で割ったものを再現率として記した．

7. 考察

7.1 トピック抽出結果考察

本実験でトピック抽出実験の対象とした 46,889 個のつぶやき中，トピックとなりうる単語が明記されていたつぶやきは 15,462 個あり，その大半（14,678 個のつぶやき）は，トピックとして使えるものであったことがわかる．ステップ 1 の処理の結果，正解と認められなかった 784 個のつぶやきに関しては，トピックとなりうる単語が複数個つぶやき中に存在し，正解データのトピックと一致しなかったことが不正解の原因となっている．

ステップ 2 における正解率は 0.31 と低いものとなった．サッカーなどのスポーツ番組の中継の際，試合状況が膠着している際などに時間をつなぐため，アナウンサーが試合に関連した別の話題を説明することが主な原因としてあげられる．実験の例では，「南アフリカ文化に関する話題」「芝の状況や天候・気温などの情報」など，無生物がトピックであった場合にトピックをアナウンサーと誤って抽出した結果が多くみられた．5.1.3 に記したスポーツ番組における前処理の際に，トピック候補リストに一致する単語がない場合には，その字幕情報を使わないなどの処理をおこなうと，（処理数は少なくなるが）正解率は上がると考えられる．別の改良手法として，[18] における字幕情報の SVM(Support Vector Machine)を用いた分類があげられる．スポーツ番組に関する字幕情報を，試合情報を説明する「実況」と試合に関連した情報を説明する「解説」にわけ，実況部分のデータのみを使うというものである．ただし，このような前処理をおこなうと，正解率は向上すると考えられるが，トレードオフとしてコメント解析のリアルタイム性が失われることとなる．

別の観点からの考察として，ステップ 2 の字幕情報を用いたトピック抽出手法は，スポーツ番組ではあまり正解率の向上や処理数の増加に寄与していないことが，実験結果から読み取ることができる．[19]におけるドラマ番組に関するコメント解析では，全体の 22%のコメントのトピック抽出に字幕情報を用いた推定は寄与しており，正解率も 88%と高い．ドラマ番組では，動作主と発話者が一致していることが多いため，字幕情報からのトピック抽出が正解率の向上や，処理数の増加に寄与するものと考えられる．これらの結果より，どのような種類の番組を対象とするか，コメント解析結果をどのようなサービスで使用するか，リアルタイム性が必要か否かによって，処理手法を変えることが望ましいと考えられる．

また，ステップ 4 では，全体の 65%のつぶやきがこの処理によってトピック抽出された．正解率は 0.59 と決して高くはないものの，この処理をおこなうことで多くのつぶやきのトピックを抽出することが可能となった．

SNS メッセージを対象としたトピック抽出の正解率に関する類似研究がないため，文字列マッチングによりトピック抽出が可能な 14,678 個のつぶやきと，本手法により正しくトピック抽出できたつぶやき(32,973 個)を比べると，約 2.25 倍のつぶやきのトピックが本手法により抽出できたといえる．

7.2 内容分類結果考察

状況説明や肯定にカテゴライズされたつぶやきの適合率再現率が高く，否定にカテゴライズされたつぶやきの適合率・再現率が低いことが実験結果から読み取れる．考えられる主な原因は 2 点ある．

1 点目は，視聴者の大半が日本チームの勝利を願っているような状況下で，試合進行中に否定的な意見が言いづらく，不安や要望に言い換えてしまうため，否定にカテゴライズされる学習データが極端に少なかったことである．本実験では，学習データに関しても「放送時間中のつぶやき」に限定してしまったため，否定データが全体の 5.2%程度と極端に少なかった．この問題点に関しては，学習データを試合後のつぶやきにもひろげ，否定カテゴリに分類される学習データを増やすことで解決できると考えられる．

2 点目の問題点は，文脈の関係から否定単語を含まずに否定的な意見を表している場合や，助動詞に否定表現が含まれている場合に否定と分類できない文法的問題点である．助動詞に否定表現が含まれる例は，その用例をピックアップし，TF-IDF 辞書に追加することで回避できるが，文脈に依存して否定意見と判断される例は，背景知識との符合などの要素が入るため，システマチックな解決は難しいと考えられる．

[10]では Twitter のつぶやきを Naive Base・最大エントロピー法・SVM を用いて，肯定または否定に分ける研究をおこない，各手法において 80%近い正解率を得ている．これに対し本実験は，5.2 で述べたように，テレビ番組関するコメントを肯定/否定の 2 値に分けることは難しいので，6 つのカテゴリに分類することとした．正解データ内での割合を観ても，肯定 43%・否定 5%であり，2 値への振り分けでは残りの 52%のデータに関する処理をおこなうこと

はできない．本手法では，全データを 6 つのカテゴリに分類し，適合率 0.85 と再現率 0.74 という一定の有効性を示す値を得たと考えられる．

7.3 サービス面から解析処理考察

3 章で述べたように，コメント解析に求められる要件はサービス毎に異なるため，それぞれのサービスで必要とされる要件と，それを満たすための解析処理プロセスを検討する．

解析結果表示グラフでは，コメントの全体的な傾向を把握するため，解析処理される数の多さが重要となる．盛り上がりをもとにした漫画風ダイジェストでも，番組のどの部分に対して幾つコメントがあったかを基にサービスを生成する．トピック抽出における Step3 の処理数が多いことから，より多くのコメントを処理し，これらのサービスを生成するためには，類似性による補完アルゴリズムを複数回実行することが望ましいと考えられる．

一方，コメント内容から誰に関する肯定意見が多いかを算出し，それを基にお勧め番組を算出するサービスでは，視聴者の興味のない内容に対する推薦を繰り返すと，不快感を引き起こしかねないため，特にトピック抽出の解析精度の高さが重要となる．実験結果より，トピック抽出を Step1 だけおこなえば，約 95%の正解率を保つことができる．このような解析精度を重視するサービスを生成する際には，トピック抽出を Step1 まで実行することが望ましいと考えられる．

また，直近のコメントから代表的なものを選択グラフとして表示し，ボタン入力で気軽に参加するための共感グラフサービスでは，リアルタイム性が重視されるため，処理時間の早さが重要となる．このようなサービスを生成する際には，処理の重い前処理はせず，類似性による補完アルゴリズムをおこなわないことで，処理時間を短くすることが可能となる．

以上に述べたように，サービスに必要なコメント解析の要件を考え，適切な処理段階を選ぶことができるよう，解析モジュールを設計し，情報還流システムに実装した．

8. まとめ

本稿では，放送中の番組に関する番組コメントの特徴の洗い出しをおこない，その特徴に合わせたコメント解析手法の提案をおこなった．番組に関するコメントを解析する際に問題となる，主語の省略や表記ゆれを解決するため，放送中の番組の字幕情報や番組概要情報を補助情報として用い，コメントの時系列性と類似性を用いて解析する手法を提案した．

提案手法の検証をおこなうため，ワールドカップ日本戦における実際の Twitter コメントを収集し実験をおこなった．トピック抽出の実験結果は，全体の 70%のトピックを正確に抽出できた．個々のコメントが何に関するコメントであるかを文字列マッチングで抽出するのに比べ，約 2.25 倍のコメントのトピックを正しく抽出できた．TF-IDF を用いたコメント内容分類結果では，適合率 0.85・再現率 0.74 の結果を得た．Twitter などの不完全な文章を感情分類する手法の多くは，肯定/否定などの 2 値に振り分ける研究が多いため，本手法の 6 つのカテゴリへの感情分類と一概に比較はできないが，一定の有効性を示す値を得たと考えられる．

今後は，実験考察で述べた適切な前処理法の導入や，ドラマやスポーツ番組以外ジャンルに関するコメント特徴の洗い出しをおこない，汎用的なコメント解析手法を提案していく予定である．また，放送通信連携サービスの一形態として，より魅力的なサービスの提案することで，より豊かな視聴環境を実現する情報共有空間の構築を進めていきたい．

参考文献

[1] S. Smit et al.,“An Open Service Infrastructure for Enriching Networked Interactive Multimedia Experiences in a Converged World”, In Proceeding of the NEM Summit 2008. Saint-Malo, France. (2008).

[2] S. Dietze, A. Gugliotta, J. Domingue,”Exploiting Metrics for Similarity-based Semantic Web Service Discovery”. In Proceeding of the 7th IEEE International Conference on Web Services. Los Angeles, USA. (2009).

[3] youview : http://www.youview.com

[4] H. Kato, “Hybridcast System and Technology Overview”, Broadcast Technology No.43, Winter 2011, pp.6-11, (2010).

[5] J. Han, X. Xie, W. Woo, "Context-based Local Hot Topic Detection for Mobile User", Proceeding of Adjunct Pervasive Computing Conference, (2010).

[6] K. Chen, L. Luesukprasert, T. Chou, "Hot Topic Extraction Based on Timeline Analysis and Multidimensional Sentence Modeling", IEEE Transactions on Knowledge and Data Engineering, Volume 19, Issue 8, pp.1016-1025, (2007).

[7] M. Gilad, “Experiments with Mood Classification in Blog Posts”, In Proceedings of ACM SIGIR 2005 Workshop on Stylistic Analysis of Text for Information Access, (2005).

[8] Twitter : http://twitter.com

[9] P. Alexander, P. Patrick, “Twitter as Corpus for Sentiment Analysis and Opinion Mining”, In Proceedings of 7th Conference on Language Resources and Evaluation, (2010)

[10] A. Go, R. Bhayani, L. Huang, “Twitter Sentiment Classification using Distant Supervision”, CS224N Project Report, Stanford, (2009).

[11] T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing contextual polarity in phrase-level sentiment analysis", In Proceedings of Human Language Technologies Conference/Conference on Empirical Methods in Natural Language Processing, pp.348-354, (2005).

[12] 有安, 妹尾, 鹿喰, “情報還流プロトタイプシステム”, 電子情報通信学会技術研究報告書, vol108, no.278, CQ2008-60, pp.5-9, (2009).

[13] 有安, 妹尾, 鹿喰, “コメント解析結果を反映した漫画風番組ダイジェスト”,電子情報通信学会技術研究報告書,第 8 回情報科学技術フォーラム一般講演論文集, no4, K-069, pp.687-688, (2009).

[14] 有安, 金次, 浜田, “ソーシャルテレビに関する一提案 - 番組コメント解析に基づいたコンテンツ推薦 - ”, 電子情報通信学会 HCG symposium 2009, A8-4, (2009).

[15] 有安, 藤沢, 金次, “ソーシャルテレビサービスのための共感グラフ生成手法”, 電子情報通信学会 2 種研究会 サイバーワールド第 17 回研究会, CW2010-25, pp39-44, (2010)

[16] 有安, 金次, 浜田, “テレビ番組に関する感想共有のためのソーシャルテレビサービス”, 映像情報メディア学会 2010 年年次大会予稿集, 10 - 4, (2010).

[17] 妹尾, “テレビドラマの構造化と評価要因分析”, Keio SFC journal 7(2), pp.110-125, (2007).

[18] 山田, 佐野, 住吉, 柴田, 八木, “アナウンサーと解説者のコメントを利用したサッカー番組セグメントメタデータ自動生成”, 電子情報通信学会論文誌. D, 情報・システム J89-D(10), pp.2328-2337 , (2006).

[19] 有安, 妹尾, 鹿喰, “コメントの類似性に基づく視聴者クラスタリング手法の提案”, 情報処理学会 IPSJ Symposium Series, DBWEB2007, Vol.2007, No3, 7C-2, (2007).