

放送コンテンツにおけるチャット内容を誘導するエージェント

山本裕貴[†] 渡邊哲[‡] 尾形正泰[‡] 今井倫太[†]

慶應義塾大学理工学部[†] 慶應義塾大学大学院理工学研究科[‡]

1. はじめに

インターネットで個人用のストリーミング放送サービスが提供され、個人が無料で不特定多数に番組を配信することができるようになった。本稿では放送内容について自由にチャットできる機能が付いているサービス（以下「生放送」とする）において、視聴者として参加しながらコメントを投稿するエージェントを設計して、チャットの内容を誘導すること、また誘導の結果として得られた視聴者のコメント（以下「視聴者コメント」とする）を解析して、特定の話題の情報を得ることを目的とする。エージェントを設計することにより、番組の主権者などがエージェントを通して欲しい話題を収集することが可能になる。本エージェントはチャットに参加している視聴者からは隠蔽され、視聴者のひとりとして振る舞う。

これまで、特定のコメントがチャットに参加している視聴者の話題を誘導出来るかどうかを明らかにする方法はなかったが、本稿では話題の方向を変える意図を持って送信したエージェントのコメント（以下「エージェントコメント」とする）に対して、その後視聴者コメントがどの程度影響を受けたかをコメント間の同一単語の出現数を調べることで、コメント間の関連性を数値化することができた。数値化によって話題を誘導出来たコメント、出来なかったコメントが判明するため、視聴者コメントの関連度をフィードバックとして用いることで、影響度の高いコメントを生成するための要素を決定できると仮定し、視聴者の話題を誘導しやすいコメントの生成に利用出来る。また、視聴者コメントとエージェントコメントの関連性を数値化することで、チャットの時系列データを可視化することが可能になった。以上から、エージェントの操作者が話題を誘導し、欲しい話題を収集するために、生放送の視聴者から気付かれな

うに特定の情報を得ることが可能になる。本稿では設計するエージェントの内コメント生成を除いた、コメントの関連付け及び可視化を行った。

2. 関連研究

エージェントを設計するにあたり、チャットシステムやソーシャルネットワークでの事例をウェブサービス上のコメント分析、オンラインユーザの行動分析に分けて紹介する。

コメントの相関を取る手法として、Chen ら [1] のビデオの動画のトピックを関連タグや検索エンジンのよく検索されているキーワードなどから検出する手法がある。本エージェントでは生放送という話題が移り変わるコンテンツで今この話題が何なのかを適宜分析し話題の方向を誘導しやすくする点で異なる。西田ら [2] の "On-Air Forum" は生放送中でのコメントの支援として、ユーザが手動でコメントの関連付けをした上で発言したり、直前のコメントと関連付けたりする方法を示した。本稿ではコメントの関連付けはコメント同士の関連度を時間・共通の単語から計算する。桐戸ら [3] の "言うストリーム" は生放送中において予め用意された話題を適切なタイミングで提示するシステムである。本稿では配信者の支援ではなく、エージェントが生放送に対して話題を変更・誘導することを目的として、関連付けを行った。ストリーミング放送におけるユーザの行動分析に関する Qiu ら [4] の研究では、動画の再生、巻き戻し、早送りなどに対する行動分析であるが、本稿ではユーザの生放

Designing agent for chat system in video streaming service

[†] Yuki YAMAMOTO, Michita IMAI

Faculty of Science and Technology, Keio University

[‡] Satoshi WATANABE, Masa OGATA

Graduate School of Science and Technology, Keio University



図1 ニコニコ生放送

送に対するコメントを用いて発言の内容を分析の対象にしている。

3. エージェントの設計

本稿では生放送としてニコニコ生放送¹⁾のコメントシステムを使用し、エージェントはコメントを受け取り、茶釜²⁾により形態素解析をする。形態素解析によって得られた名詞、形容詞によってコメントの相関を判定する。コメントはエージェントコメントと視聴者コメントに分け、相互のコメントの相関を計算する。エージェントコメントと相関がある視聴者コメントは紐付けられ関連コメントとしてコメント閲覧システムに登録される。

3.1 計算方法

ある視聴者コメントの関連度 r_t ($0 \leq r_t \leq 100$)の計算について、コメントの話題性は時間で減衰していくとみなし、先に述べた形態素解析によって得られた名詞、形容詞のうち、視聴者コメントとエージェントコメントの両方に出現している単語数 W_{ta} に正の相関、及びエージェントコメントの投稿時間 t_a 秒と対象の視聴者コメントの投稿時間 t_t 秒の差の絶対値に負の相関を持つように計算した。式(1)に使用した計算式を示す。係数 k に $k = 1000, 5000$ の各値を使用し、 $r_t \geq 50$ であるコメントをエージェントコメントに関連した視聴者コメントとした。

$$r_t = \min \left\{ \frac{k \cdot W_{ta}}{|t_a - t_t|}, 100 \right\} \quad (1)$$

3.2 コメントの関連度の視覚化

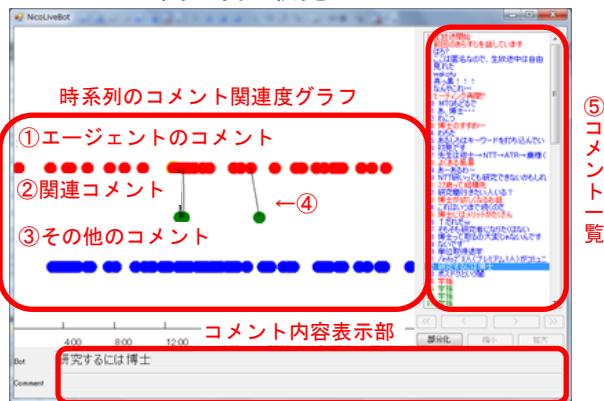


図2 コメント閲覧システム

コメント閲覧システムは図2に示す。全コメントを①エージェントコメント、②エージェントコメントに関連している視聴者コメント、③それ以外の視聴者コメントに分類し色付けた丸によって表し、④エージェントコメントに関連しているコメントは線で結んでいる。同様に

⑤色付けしたコメントのリストを右部に表示し、リストを選択すると選択したコメントが対応している丸をハイライトする。また、それぞれのコメントは選択すると下部にコメントの内容が表示される。関連付けされたコメント同士は選択された際にどのコメントと関連付けされたかが一緒に表示されるようになっている。関連付けによってエージェントのどのコメントが視聴者コメントに影響を与えたかを計ることができる。以上のシステムにより影響の高かったコメントを抽出することが出来る。そのコメントをコメント生成にフィードバックすることで影響の高いコメントを生成することが出来るのではないかと推測した。

4. 考察

エージェントの代わりに1ユーザが発言したときの視聴者コメントとの関連度をシステムで視覚化し、図2に見られるようにいくつかの関連しているコメントを判別することに成功した。係数 k の値が1000と低い時は30秒程度までのコメントが関連していると判別され、5000と高い時は1分ほど離れていても関連していると判別された。この結果より、関連度の高い視聴者コメントを多く引き出すことができたエージェントコメントを抽出することで、長すぎる文よりある程度短めの文、もしくは単語のみの発言の方が誘導しやすいことがわかった。また話題を誘導しようと同じ単語を含む単語を何度も投稿しても関連しているコメントが多くなるとは限らないことがわかった。本稿ではコメントの相関判定に単純なキーワードマッチングを使ったため、人間が判定すれば関連があっても、本システムでは関連付けできてないと思われるコメントが見受けられた。

5. 結論

本稿ではエージェントコメントと視聴者コメントの相関を取り、エージェントコメントが視聴者コメントにどれだけ影響を与えたかを判定、視覚化した。

参考文献

- [1] Tianlong Chen, An Effective Multi-Clue Fusion Approach for Web Video Topic Detection, MM '12, 2012.
- [2] 西田 健志, On-Air Forum: リアルタイムコンテンツ視聴中のコミュニケーション支援システムの設計とその実証実験, コンピュータソフトウェア, Vol28, No.2(2011), 2011
- [3] 桐戸 創也, 言うストリーム: 話題候補を実時間提示する個人生放送支援システム, 情報処理学会 インタラクシオン 2011, 2011
- [4] Fan Qiu, An analysis of user behavior in online video streaming, VLS-MCMR '10, 2010

1) <http://live.nicovideo.jp/>
 2) <http://chasen.naist.jp/hiki/ChaSen/>