

類似関係データに基づく 関連アイテムの可視化手法の提案

黒田 脩平
Shuhei KURODA
情報メディア学科 2018 年度卒業

1. はじめに

クラウドソーシングから得られたアイテム間の類似関係データに基づき、アイテムを2次元空間に可視化する手法を提案する。本研究では、アイテムのドメインとして映画をとりあげる。

本システムでは、ユーザが選択したアイテムの類似関係を2次元特徴空間上に提示させ、ユーザが選択したアイテムに関連するアイテムを探索することを支援する。

2. システム構成

図1にシステム構成図を示す。以下、各処理について説明する。項目番号は図中の番号に対応する。

(1) クラウドソーシングによりアイテム間の類似関係データを収集する。具体的には、ワークに主観的に類似すると思うアイテム同士を選択してもらう。ワークに類似性を判断してもらうための基準として世界設定が類似しているかという観点に着目してもらう。世界設定とは、ジャンルとは異なり、時代や舞台、登場人物の種族、世界の法則などに関する設定を指すが、類似性の最終的な判断はワークの主観に委ねている。

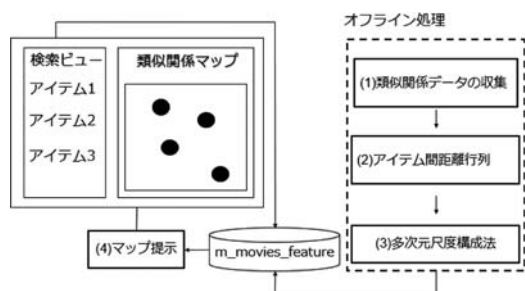


図1 システム構成

$$R = \begin{matrix} & \begin{matrix} 873 & 136 & 822 & 1518 \end{matrix} \\ \begin{matrix} 873 \\ 136 \\ 822 \\ 1518 \end{matrix} & \begin{bmatrix} 1 & 4 & 1 & 1 \\ 4 & 1 & 3 & 4 \\ 1 & 3 & 1 & 0 \\ 1 & 4 & 0 & 1 \end{bmatrix} \end{matrix}$$

図2 アイテム間共起行列

$$D = \begin{matrix} & \begin{matrix} 873 & 136 & 822 & 1518 \end{matrix} \\ \begin{matrix} 873 \\ 136 \\ 822 \\ 1518 \end{matrix} & \begin{bmatrix} 0 & 0.2 & 0.5 & 0.5 \\ 0.2 & 0 & 0.25 & 0.2 \\ 0.5 & 0.25 & 0 & 1 \\ 0.5 & 0.2 & 1 & 0 \end{bmatrix} \end{matrix}$$

図3 アイテム間距離行列

(2) アイテム間距離行列を作成するため、収集した類似関係データを基に、アイテム間共起行列を作成する。ここでの行列の要素は類似関係データから得られた、アイテムとアイテムの共起回数を示す。ただし対角要素はすべて1とする。アイテム間共起行列の例を図2に示す。また、行列名はアイテムIDを示す。

作成したアイテム間共起行列をアイテム間距離行列に変換する。アイテムの間の距離は共起回数の逆数とする。ただし、0による除算を避けるため分母に1を加える。アイテム間距離行列Dの例を図3に示す。

(3) アイテム間距離行列を基に、多次元尺度構成法を用いてアイテム集合を2次元特徴空間に写像し、アイテム間類似関係マップを作成する。この際、得られたデータを `m_movies_feature` に格納する。

(4) 選択されたアイテム周辺のアイテム間類似関係マップをマップビューに表示する。

3. 評価

類似関係マップに提示されている各アイテムの周辺に類似性のあるアイテムが提示されているか確認する。まず、全体を俯瞰した後、例として、グルー



図4 類似関係マップ

「ヒーロー」付近の類似関係マップについて定性的に分析する。

図4は表示された類似関係マップの全体の傾向を俯瞰するために、近傍のアイテム同士の関連性を踏まえ、大まかなグループを図中に示したものである。

各グループに含まれるアイテムの内容について、著者が調べたところ、互いに関連するアイテム同士がそのグループ内で近くに分布していることを確認した。

図5はグループ「ヒーロー」周辺を拡大したものである。『アベンジャーズ (2012年の映画)』はグループが「ヒーロー」であり、ニューヨークが舞台である。同様に、『スパイダーマン：ホームカミング』もグループが「ヒーロー」であり、ニューヨークが舞台である。このことから、グループ、世界設定がともに類似しているアイテムが提示されていることがわかる。また、周辺には『マン・オブ・スティール』など、世界の法則としてヒーローが登場する作品が提示されている。この結果から、『アベンジャーズ (2012年の映画)』の周辺に類似作品を提示できたといえる。

以上の結果から、人の主観により挙げられた類似

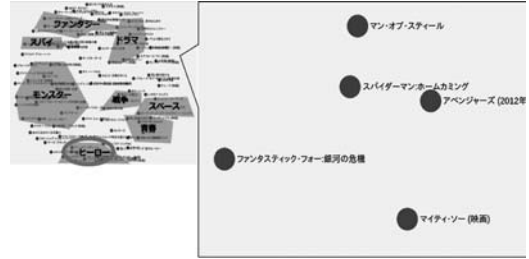


図5 グループ「ヒーロー」の拡大図

関係データを基にすることで、類似関係マップを適切に生成できることを確認した。さらに、世界設定が類似しているアイテムも提示できることがわかった。この類似関係マップを用いることにより、ユーザが興味のあるアイテムに関連するアイテムの探索を支援することが可能であるといえる。

4. おわりに

本稿では、クラウドソーシングから得られたアイテム間の類似関係データに基づき、アイテムを2次元空間に可視化する手法を提案した。今回の実験結果より、このシステムは得られた類似関係データに基づき構築されたアイテム空間において、ユーザが興味のあるアイテムに関連するアイテムを提示するシステムとしては機能していた。しかし、アイテムとの距離が離れるほど類似していないアイテムを提示するといったことができず、アイテムによっては近い距離に類似しない作品が提示され、システムとしては不完全な推薦が確認できた。

今後の課題として、クラウドソーシングによる類似関係データ数の増加や、類似アイテムだけでなく選択したアイテムと極端に類似していないアイテムのデータを適応させ、類似しているアイテム同士の距離を短く、類似していないアイテムの距離を長くすることで各アイテムの類似関係がより確認できるようにすることである。