

## バースト解析を用いた小説テキストの物語理解手法の提案

## D-4 Proposal of Techniques for Comprehending Narratives of Novels by Burst Analysis

金子 航太<sup>†</sup> 荒川 達也<sup>†</sup>Kota KANEKO<sup>†</sup> Tatsuya ARAKAWA<sup>†</sup><sup>†</sup> 群馬工業高等専門学校 生産システム工学専攻<sup>†</sup> Advanced Production Systems Engineering Course, National Institute of Technology, Gunma College

## 1. はじめに

時系列のテキストデータにおいて、あるキーワードを含む文書が一時的に急激に増加する現象を「バースト」と呼ぶ[1]. これまでにニュース記事や Twitter など、いくつかのメディアを対象にバーストを検知することによりトレンドの動向を分析するという手法が研究されている[2][3]. それに対し、本研究ではバースト解析の方法を小説テキストに適用することにより、作品内のストーリーや場面の流れを大まかに把握する手法を提案する. 小説テキストに対する「物語理解」の研究は近年、盛んに行われているが([4]など), バースト解析を用いることで比較的手軽に登場人物の活躍度や場面転換等を大づかみに抽出することができ、小説読書支援([5]など)への応用も視野に入ると考えている.

本稿では以下、提案手法の概要と簡単な試作システムによる実行例を述べる.

## 2. 提案手法

本提案手法では、対象とする小説テキストの各章を時系列文書と考え、指定したキーワードの出現頻度が特定の章において急激に増加する現象を「バースト」と定義する. キーワードとしては人名や地名、アイテムなどさまざまなものが考えられるが、今回の試作システムでは人名(登場人物名)のみ対象とした. 各登場人物のバーストを検知することにより、その人物が活躍する章を抽出することが本システムの目的である. バースト解析の手法として本システムでは、Kleinberg のバースト検知アルゴリズム[1]を使用した.

## 3. 実行結果

現段階ではオルコット作『若草物語』(青空文庫版)とバーネット作『小公女』(青空文庫版)にて実行結果が得られている. 実行結果の妥当性を確かめるため、試作システムにより得られた各登場人物のバースト箇所と目視により選定した各登場人物の活躍箇所の比較を行った.

以下に『若草物語』を使用した実験結果を示す. バースト箇所と目視で選定した活躍箇所を表 1 に示す. また、登場人物の章ごとの登場頻度(一部抜粋したもの)を図 1 に示す. 表 1 と図 1 を見ると、「デビス」や「アンニイ」は物語全体では滅多に活躍しない脇役だが、それぞれ活躍する章で確かにバーストが検知されていると分かる. 次に、実際に各登場人物のバースト箇所と活躍箇所とを比較して F 値を算出したところ『若草物語』では平均して 0.93 となった. 同様に『小公女』で実験したところ、F 値は平均して 0.89 となった. これらの値は良好な数字を示しており、本提案システムにより登場人物の活躍シーンを検出できていると考えられる.

表 1 実行例:『若草物語』(全 23 章)

登場人物	バーストした章	活躍する章
メグ	9, 20, 23	9, 20, 23
ジョウ	3, 8, 11, 14	3, 5, 8, 11, 14
ベス	1, 6, 17, 18, 22	6, 17, 18
エミイ	7, 8, 19	7, 8, 19
デビス	7	7
アンニイ	9	9

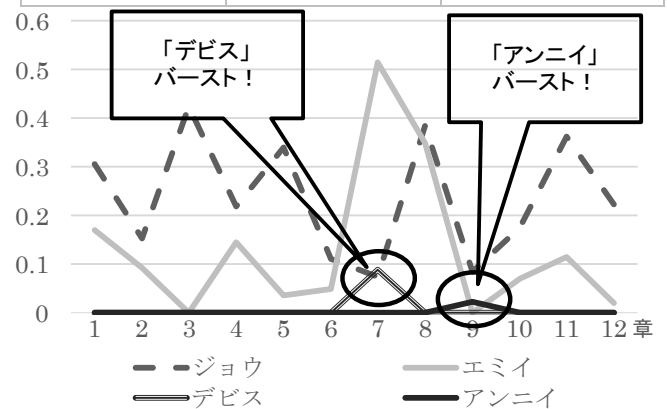


図 1 登場人物の登場割合(一部抜粋)

## 4. まとめ

Kleinberg のバースト検知アルゴリズムを用いて、小説テキストの解析により、登場人物の活躍シーンを検出する方法を提案した. 現在までに 2 つの小説で実験を行い妥当な結果を得た. 今後はより多くの小説にて実験を行い、また、アンケートなどを行って実行結果の評価を行っていききたい. 参考文献

[1] J. Kleinberg, “Bursty and Hierarchical Structure in Streams,” in Proceedings of the eighth ACM SIGKDD International Conference on Knowledge Discovery, pp.91-101, 2002

[2] 高橋佑介ほか, 「ニュースにおけるバースト キーワードの話題への集約」, 第 3 回 DEIM フォーラム論文集, 2011

[3] 佐々木謙太郎ほか, 「Twitter におけるキーワードの出現周期の自動判別手法に関する検討」, 2013 年度人工知能学会全国大会論文集, 2013

[4] 山本聖也ほか, 「推理小説の難易度評価のための犯人推定」, 言語処理学会 第 21 回年次大会 発表論文集, 2015

[5] 岡田悟ほか, 「質問応答技術を用いた 小説読書支援システムの提案」, 知能と情報(日本知能情報ファジィ学会誌)Vol.27, No.2, pp.608-615, 2015