

平成 24 年度  
修 士 論 文

クラスタリングを利用した KJ 法支援システムの開発

Development of Support System  
for KJ Method by Using Clustering

梅林 洋

神戸情報大学院大学  
情報技術研究科 情報システム専攻

## 内容梗概

近年様々な場面で発想法が利用されている。その中でも創造的なアイデアや全体の把握をする際に有効な手法として KJ 法がある。しかし、KJ 法を利用する場合、数多くの手順を要するため「作業に多くの時間が必要となる」といった問題点がある。これらの作業時間を短縮するために、各々のユーザが作成したラベルを自動的にグループ編成する KJ 法支援システムの開発を目的とする。

グループ編成の自動化は、特徴語抽出処理、出現頻度の算出処理、ラベルの分類処理の 3 つで構成した。特徴語抽出処理では、形態素解析と代表表記の取得をする。手順は、まず形態素解析ツールの `mecab` を用いてラベルを形態素解析し名詞と動詞を取得する。次に、「動詞の自立語」と「名詞」を形態素解析ツールの `juman` に入力して代表表記を一定のルールのもと取得する。「取得した代表表記」と「代表表記取得のルール外の単語」を特徴語とする。出現頻度の算出処理では、ラベル中の単語とその出現頻度で構成される単語ベクトルを作成する。なお、代表表記を複数持つ特徴語における出現頻度は、特徴語の出現数をその特徴語が持つ代表表記の総数で割った値とする。ラベルの分類処理では、単語の出現頻度を重みとした単語ベクトルを使用する頻度分析でのクラスタリングを行う。クラスタリングの際のラベル間の類似度の評価基準は、本システムでは、ベクトル空間モデルを用いているためコサイン類似度を採用した。クラスタリングのアルゴリズムは繰り返しクラスタを 2 分割することでクラスタリングを行う **Repeated Bisection** を用いた。このアルゴリズムは初期値の設定によるクラスタリングの偏りがなく、繰り返し 2 分割することによる計算量の減少といったメリットがある。クラスタ数を事前に決定出来ないため、閾値を設けてクラスタリングを行う **leader-follower** 法を用いた。なお、KJ 法のラベルの分類処理を行う際に、①短文のため特徴を取得しづらい、②口語的な表現が多いため表記揺れが発生する、といった問題が存在する。この問題に対応するために、本システムでは代表表記を取得する。

本システムの有効性を確認するために、実証実験を 2 つ行った。実証実験 1 では数値で評価し、実証実験 2 では人による評価を行った。実証実験 1 で有効性を確認する項目は、①人手とのグループ編成結果との比較、②特徴語を抽出する際に、`mecab` で形態素解析したのみの結果と、形態素解析した単語を `juman` に入力し、代表表記を取得した結果との比較、の 2 つとした。方法は、手動で KJ 法を実施して作成したラベルを本システムに入力し、それぞれのグループ編成結果を比較することとした。評価の指標は  $F$  値を用いた。なお、 $F$  値が 0.7 以上で①の有効性はあるとした。実証実験 2 では、数値ではなく、実際に人が見た場合で確認するために、実証実験 1 に協力して頂いた方々に本システムのグループ編成結果を評価して頂いた。実験の結果、 $F$  値が 0.7 をほとんど上回らず、手動によるグループ編成結果に比べて若干の相違が出た。

今回の研究の成果として、グループ編成は出来たものの、まだ実際に KJ 法を支援できるだけのグループ編成は行えていない。今後は、各機能の向上をするための開発を行う必要がある。

# 目次

1. 序論	1
2. KJ法の概要	2
2.1. KJ法とは	2
2.2. KJ法の実施手順	2
2.2.1. 概要	2
2.2.2. ラベル作り	3
2.2.3. グループ編成	3
2.2.4. 図解化	3
2.2.5. 文書化	3
2.3. KJ法の問題点	4
2.4. 既存のKJ法支援システム	4
2.4.1. ネットワークを用いたKJ法支援システム	4
2.4.2. 計算機上で複数の参加者が同時に作業できるKJ法支援システム	4
2.4.3. ラベル作りを支援するKJ法支援システム	4
2.4.4. 紙面上での作業結果の保存を支援するシステム	4
2.4.5. グループ編成を支援するKJ法支援システム	4
3. 分類に関する既存方法	6
3.1. 教師有り分類	6
3.1.1. SVM(サポートベクターマシーン)	6
3.1.2. ナイーブベイズ	6
3.2. 教師なし分類	6
3.2.1. 多次元尺度構成法	6
3.2.2. クラスタリング	6
4. クラスタリング	8
4.1. 階層的クラスタリング	8
4.2. 非階層的クラスタリング	8
5. システムの設計	10
5.1. システムの概要	10
5.2. 本システムの作業及びデータの流れ	10
6. グループ編成方法	12
6.1. グループ編成結果を出力するまでの処理の概要	12

6.2.	グループ編成結果を出力するまでの処理の詳細	12
6.2.1.	特徴語抽出処理	12
6.2.2.	出現頻度の算出処理	13
6.2.3.	ラベルの分類処理	13
6.3.	代表表記について	14
7.	実装	18
7.1.	システム構成	18
7.1.1.	システムの全体構成	18
7.2.	使用するツール及び運用環境	18
7.3.	画面遷移と各機能	19
7.3.1.	画面遷移の概要	19
7.3.2.	ログイン	20
7.3.3.	メニューの操作	20
7.3.4.	トピックの新規作成及び編集	21
7.3.5.	ラベルの登録及び確認	21
7.3.6.	グループ編成結果の確認	22
8.	実証実験	24
8.1.	実証実験の概要	24
8.1.1.	実証実験 1	24
8.1.2.	実証実験 2	24
8.2.	実証実験 1 の実験方法	24
8.2.1.	手順	24
8.2.2.	手動で実施した KJ 法の概要	24
8.2.3.	評価方法	24
8.3.	実証実験 1 の実験結果	26
8.3.1.	実験テーマ 1 (運動習慣の仕組み検討)	26
8.3.2.	実験テーマ 2 (夏を快適に過ごす仕組み検討(グループ A))	29
8.3.3.	実験テーマ 3 (夏を快適に過ごす仕組み検討(グループ B))	31
8.3.4.	実験テーマ 4 (秋をもっと快適に過ごす仕組み検討)	35
8.4.	実証実験 2 の実験方法	40
8.4.1.	手順	40
8.4.2.	評価方法	40
8.5.	実証実験 2 の実験結果	40
8.6.	実験結果のまとめ	40
8.6.1.	実証実験 1 の各条件の $F$ 値について	40

8.6.2.	実証実験 1 と実証実験 2 について .....	42
9.	考察 .....	43
9.1.	本システムと手動でのグループ編成結果について .....	43
9.1.1.	本システムと手動でのグループ編成結果の相違 .....	43
9.1.2.	本システムと手動でのグループ編成結果の相違の原因と解決策 .....	43
9.2.	「mecab のみ」と「mecab と juman」の比較 .....	43
9.3.	「名詞と動詞」と「名詞のみ」の比較 .....	44
9.4.	$F$ 値による評価と人による評価の相違について .....	44
10.	まとめと今後の課題 .....	45
10.1.	研究成果 .....	45
10.2.	今後の課題 .....	46
10.2.1.	手作業のメリットがシステムを利用することで減少することへの課題 .....	46
10.2.2.	グループ編成に関わる課題 .....	46
10.2.3.	検証に関わる課題 .....	46
	謝辞 .....	47
	参考文献 .....	48
	付録 1 : 電子情報通信学会 総合大会 論文 (2013.3) .....	49
	付録 2 : 修了発表会 発表資料 .....	50