

# テレビサービスの分析と価値創出に向けての検討

\* 福原知宏<sup>(1)</sup>, 武田英明<sup>(2)</sup>

(1) 東京大学人工物工学研究センター, (2) 国立情報学研究所

〒277-8568 千葉県柏市柏の葉 5-1-5

E-mail : {fukuhara, takeda}@race.u-tokyo.ac.jp

**Abstract:** 本論文ではテレビサービスの分析と価値創出に向けた研究について報告する。テレビ番組の価値を探るため、本論文ではテレビ番組出演者のネットワークについて分析を行った。またテレビサービスの改善を目的として、地上波と衛星放送番組表を用いたテレビ番組のジャンル推定について述べる。

## 1. はじめに

近年、テレビ離れが進んでいるという報告がある。2009年3月21日に放送されたNHK番組「テレビの、これから」<sup>1</sup>では、テレビ局関係者、テレビ番組制作者、一般視聴者が参加し、昨今のテレビ離れの原因は何か、テレビは今後どうなっていくかについて意見交換が行われた。この中で、近年はライフスタイルの変化や、携帯電話、インターネットの普及等によりテレビを見ない人が増えているという報告があった。

テレビ離れを示すデータとして、図1に2005年国民生活時間調査報告書[国民生活時間調査]より作成した1995年、2000年、2005年における年代別の平均テレビ視聴時間を示す。図1を見ると、20代、30代でテレビ視聴時間が減少していることが分かる。20代の平均テレビ視聴時間は2時間38分(1995年)、2時間37分(2000年)、2時間25分(2005年)であった。1995年と2005年を比較すると、20代では12分、30代では22分の視聴時間が減少している。このように我々のテレビとの接触は若年層を中心に変化し始めており、これに伴ってテレビサービスも新たな変化を要請されている。すなわち、番組制作側においては、どのような番組を制作するか、その番組はどのようなテーマでどのような出演者を登場させるかといった番組の魅力度・価値に関する問題から、番組情報の効率的生成など、解決すべき課題がある。

本論文ではテレビサービスにおける価値の創出を目的とし、現在放送されている番組について以下の分析を行った。第1に、番組の出演者に注目し、番組出演者間のネットワーク(出演者ネットワーク)の分析を行った。第2に、テレビサービスの改善における人工知能アプローチとして、ケーブルテレビ(以下、CATV)の番組に対するジャンル推定を行った。

本論文の構成は次の通りである。2.ではテレビ番組出演者ネットワークの分析について述べる。3.では機械学習を用いた番組ジャンル推薦について述べる。4.では本論文のまとめと今後の展望について述べる。

---

<sup>1</sup> 番組ホームページ : [http://www.nhk.or.jp/korekara/nk23\\_tv/index.html](http://www.nhk.or.jp/korekara/nk23_tv/index.html)

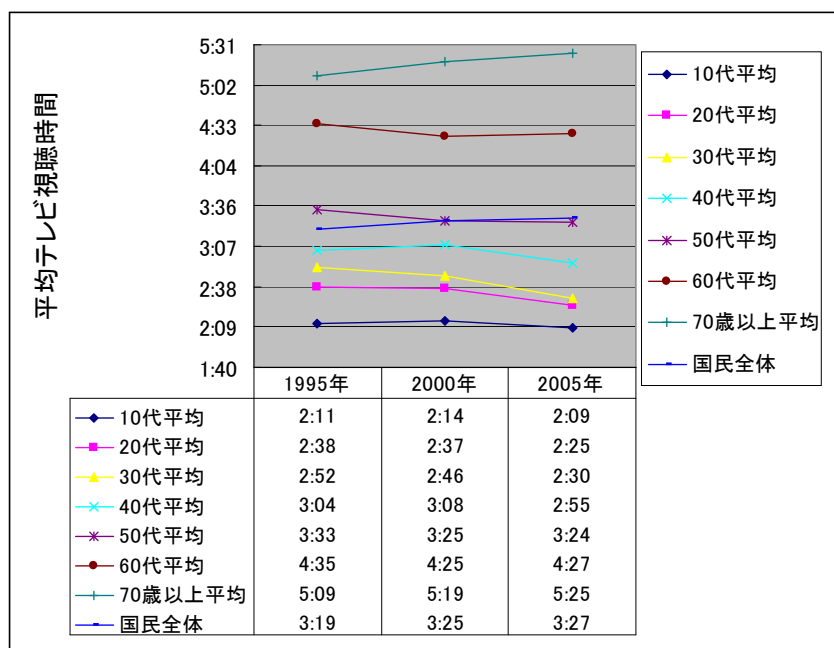


図 1 年代別平均テレビ視聴時間

(2005年国民生活時間調査報告書[国民生活時間調査]より作成)

## 2. テレビ番組出演者ネットワークの分析

本節ではテレビサービスの分析として、地上波および衛星放送の出演者に関する分析を行った。テレビ番組において出演者は番組の価値を左右する1つの要素と考えられることから、本研究では Web 上の番組情報サイト<sup>2</sup>から情報収集を行い、番組名、放送チャンネル、時間、概要、出演者等の情報を得て、出演者間のネットワークについて調査を行った。以下、出演者ネットワークの分析について述べる。

### 2.1 調査手順

2008年7月22日～8月28日の東京地区の地上波番組表ならびに衛星放送番組表から出演者情報を抽出し、同一番組に出演した2人の間に無向リンクを張ることで出演者ネットワークを作成した。作成したネットワークはノード数 18,520、リンク数 195,407、密度 0.0010315 であった。また、このネットワークにおける最長最短距離は 21 であった<sup>3</sup>。

### 2.2 分析結果

出演者ネットワークの各ノードの次数、媒介中心性について分析を行った。図 2 に次数順の出演者リストを、図 3 に媒介中心性順の出演者リストを示す。次数順ではお笑いタレント（久本雅美、中山秀征、青木さやか、今田耕司）が上位に出現することが分かった。これはお笑いタレントが幅広いジャンルの番組に出演していることが背景にあると考えられる。図 4 に久本雅美を中心とする出演者ネットワークを示す。媒介中心性では、アナウンサーや声優といった出演者が浮き彫りとなった。出演者ネットワークにおいてアナウンサーや声優が他の分野の出演者を橋渡ししているものと考えられる。

### 2.3 今後の課題

今後の課題として、番組に出演する出演者と実際の視聴データとの関係について調査し、視聴者にとって魅力的な番組とは何か、出演者の観点から調査を行う必要がある。

<sup>2</sup> <http://www.nikkansports.com/>

<sup>3</sup> [イ・ダヘ]⇔[橋爪ゆか]間の距離

順位	出演者	回数	順位	出演者	回数
1	久本雅美	625	21	高橋英樹	422
2	中山秀征	608	22	宮崎宣子	417
3	青木さやか	579	23	エドはるみ	417
4	今田耕司	549	24	石原良純	409
5	島田紳助	534	25	みのもんた	404
6	小島よしお	510	26	竹山隆範	401
7	タカアンドトシ	505	27	内田恭子	399
8	明石家さんま	501	28	伊東四朗	397
9	西川史子	492	29	アンガールズ	397
10	次長課長	484	30	高橋克実	391
11	柴田理恵	483	31	ギャル曽根	390
12	にしおかすみこ	469	32	内藤剛志	389
13	賀集利樹	467	33	つるの剛士	385
14	ビビる大木	452	34	徳光和夫	378
15	勝俣州和	450	35	優香	378
16	磯野貴理	436	36	ベッキー	375
17	スザンヌ	428	37	高畑淳子	375
18	チュートリアル	427	38	真鍋かをり	372
19	土田晃之	426	39	石田純一	372
20	関根勤	424	40	里田まい	371

図 2 回数順で見た出演者

順位	出演者	媒介中心性	順位	出演者	媒介中心性
1	中野美奈子	0.049182669	21	三船敏郎	0.007242776
2	坂上みき	0.025081242	22	ニコラス・ケイジ	0.007240572
3	山寺宏一	0.020360961	23	スマップ	0.007136077
4	戸田恵子	0.019822035	24	ハリソン・フォード	0.006834164
5	大塚明夫	0.012265308	25	浜田雅功	0.006386179
6	八嶋智人	0.011819457	26	内藤剛志	0.006356041
7	役所広司	0.010381984	27	青木さやか	0.006353084
8	押切もえ	0.010326498	28	次長課長	0.006352529
9	小澤征爾	0.010323855	29	タカアンドトシ	0.006344226
10	ビビる大木	0.009497155	30	てらそままさき	0.006113949
11	クリント・イーストウッド	0.009491981	31	ロバート・ミッチャム	0.006004388
12	ショーン・コネリー	0.009296223	32	シルベスター・スタローン	0.005946753
13	黒柳徹子	0.008621664	33	中山秀征	0.005862626
14	明石家さんま	0.008323498	34	キャメロン・ディアス	0.005791928
15	遠藤玲子	0.008027997	35	サイモン・ラトル	0.005744028
16	磯部勉	0.007994096	36	山口達也	0.005696107
17	島田紳助	0.007782808	37	石坂浩二	0.005547831
18	久本雅美	0.007709688	38	中江有里	0.00542942
19	柘植恵水	0.007567788	39	津嘉山正種	0.005404114
20	内田光子	0.007300874	40	内田恭子	0.005396673

図 3 媒介中心性順で見た出演者（背景色が緑：アナウンサー，青：声優）

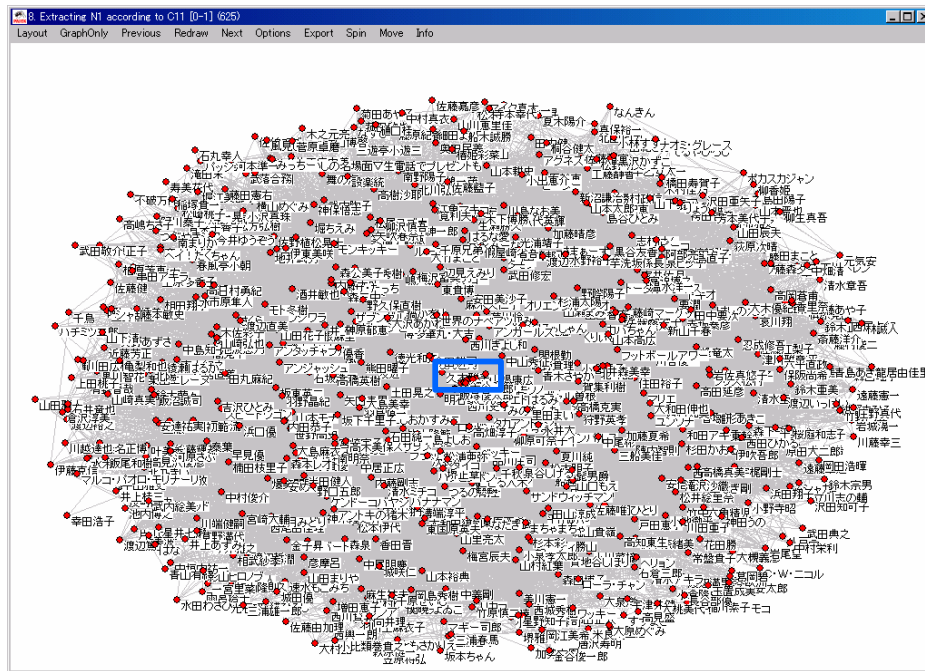


図 4 久本雅美を中心とする出演者ネットワーク  
(625 ノード, 20,225 リンク, 密度:0.1037179)

### 3. 機械学習を用いた番組ジャンル推定

本節では、テレビサービスの改善において人工知能研究が貢献し得る課題として、テレビ番組の EPG(Electronic Program Guide)におけるジャンル付与を取り上げ、機械学習を用いた番組ジャンル推定について述べる。今日、テレビ番組の情報提供手段として、EPG が広く用いられている。EPG では番組名、番組概要、放送開始時刻・終了時刻、チャンネル等の情報が配信されている。ここでは EPG データを訓練データとして教師あり学習を行うことで、番組ジャンルの自動推定を行う。本研究では地上波ならびに衛星放送の EPG データを用いて、CATV 番組のジャンル推定を行った。

#### 3.1 実験方法

本研究では SVM(Support Vector Machine) [Vapnik1998]を用いてジャンル推定を行う。学習ツールには LIBSVM<sup>4</sup>を用い、線型カーネルを用いた。用いた特徴量として番組長さ(分)、番組タイトルならびに概要に形態素解析を施して得た形態素(名詞、動詞、形容詞、副詞、未定義語)、出演者名を用いた。分類項目として、ARIB(社団法人電波産業会)の規定する大項目 13 ジャンル [ARIB STD-B10]への分類を行った。

#### 3.2 訓練データ

訓練データには、2008年7月1日～10月31日の期間に放送された地上波番組表(東京地区)と衛星放送番組表を用いた。訓練データは計 30,000 件の番組情報を含む。テストデータにはある CATV で放映されている 3 チャンネル A, B, C の番組データを用いた。

#### 3.3 実験結果

訓練データを用いて分類器の学習を行った結果、テストデータにおける正解率は Channel A, B, C でそれぞれ 59.9%, 29.7%, 40.8%であった。また、図 5 に地上波ならびに衛星放送番組表から推定したジャンル別 F 値を示す。これを見ると、ドラマジャンルでは概ね良い正答率を維持しているが、その他のジャンルでは正答率にばらつきがあることが分かる。

<sup>4</sup> <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

### 3.4 今後の課題

今後の課題として、分類に失敗したデータの分析を行い、更なる正答率向上を目指すと共に、番組や出演者に関する付随情報を Wikipedia 等から抽出し特徴量として用いることについて検討する。

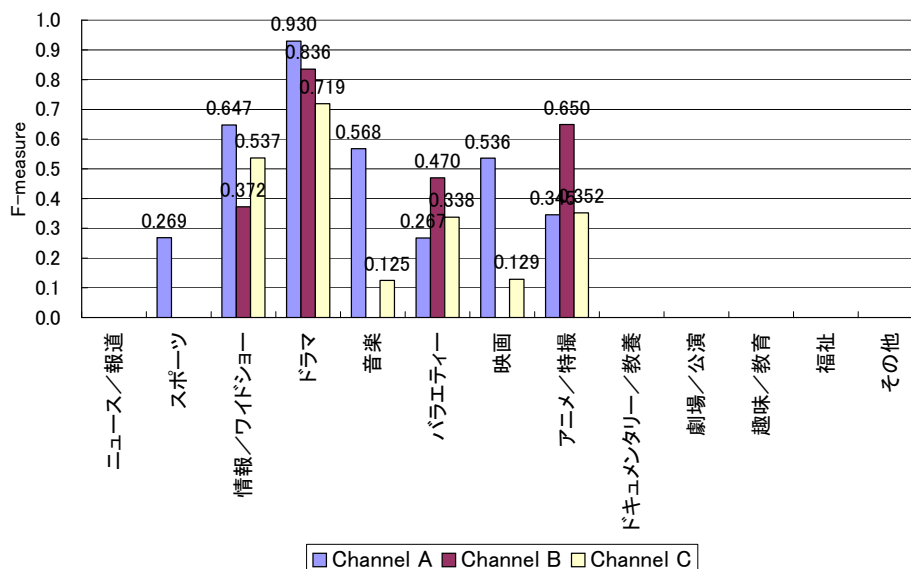


図 5 地上波+衛星放送番組表から推定したジャンル別 F 値

## 4. まとめと今後の展望

本論文ではテレビサービスの分析と価値創出に向けた取り組みとして、番組出演者に注目した番組の分析と、機械学習を用いた番組ジャンル推定について述べた。テレビサービスにおける価値創出というテーマは非常に広く、本研究だけで捉えきれものではないが、EPGをはじめとする番組情報や視聴データなどを用いてテレビサービスにおける価値創出に向けて研究を進めていきたい。今後の展望として、実際の視聴データを用いた番組価値の測定、視聴データに基づく番組推薦について研究を進める。

### 参考文献

- [国民生活時間調査] 日本放送協会編; 2005 年国民生活時間調査報告書, 日本放送出版協会 (2005).
- [ARIB STD-B10] 社団法人電波産業会編, デジタル放送に使用する番組配列情報 (ARIB STD-B10), 4.6 版 (2008).
- [Vapnik1998] Vladimir N. Vapnik, The Statistical Learning Theory. Wiley (1998).