



アクティビティ旅行パターンを用いた場合と
用いない場合でのライフスタイル分類

LIFESTYLE CLASSIFICATIONS

WITH AND WITHOUT ACTIVITY-TRAVEL PATTERNS

Hong-Zhi Lin, Hing-Po Lo, Xiao-Jian Chen: Lifestyle
classifications with and without activity-travel patterns,
Transportation Research Part A, Vol.43A, pp.626–638, 2009.

2010/10/09 論文ゼミ#4 M1 大村

朋之

内容

1. 導入
2. 方法論
 - 1) アクティビティ旅行パターンがある場合のライフスタイル分類
(クラスター分析・因子分析)
 - 2) アクティビティ旅行パターンがない場合のライフスタイル分類
(SVM)
3. 実データへの適用
 - 1) データについて
 - 2) **Phase1**:アクティビティ交通パターンがある場合のライフスタイル分類
 - 3) **Phase2**:アクティビティ交通パターンがない場合のライフスタイル分類
 - 4) センサスデータへの適用
4. 結論と将来の方向性
5. (自分の研究との絡み)

1. 導入

- 古典的なトリップベースなアプローチよりも、**アクティビティベースなアプローチ**は、個人や世帯の交通行動をより現実的に表現できる。
- 活動・交通パターンに影響を与える要因というのは
 - 家庭内での要因(社会人口学特性, 世帯構造, 世帯構成員年齢など)
 - 家庭外での要因(都市形態, 社会環境, 技術, 交通政策など)
- アクティビティベースなアプローチで目指すのは、簡単に土地利用計画や交通計画に反映させられる情報を得ること。
 - セグメンテーションしてそれぞれモデル化が望まれる
 - 単純な性別年代とかで分けるのではなく、ライフスタイルでのセグメンテーションはどうか？

1. 導入

- 従前の研究では、交通や活動からライフスタイル分類を3つの方法で行っている
 1. (Salomon and Ben-Akiva, 1983; Hildebrand, 2003)
個人と世帯の社会経済特性・人口学的特性からクラスター分析
 2. (Krizek and Waddell (2002), Lanzendorf (2002), and Krizek (2006))
短期間の活動・交通データと長期間の社会経済特性・人口学的特性から、因子分析・クラスター分析
 3. (Bagley and Mokhtarian, 1999, 2002)
長期間の活動・交通データ → 但し、データ入手が困難で難しい。

活動・交通データがある2・3が望ましい。ただ多数のデータ入手は困難。サンプリングされたデータに対してのみでしか分類を行わない。抽出されていない人のライフスタイル分類をどのようにすればよいかわからない。

- 国勢調査は全数調査なので、社会経済特性は把握可能。ここから、ライフスタイル分類ができるとうい。

2.1 活動・交通データを用いたクラスタリング手法

○ クラスタ分析

- (本論文の中では,) 様々な変数(社会経済人口情報, 移動情報)をもつサンプルがいる中で, 似た変数を持つ人をグループ化する手法. **K-means**法など複数のアルゴリズム. 距離が近いほど類似度は高い. 距離計算には**Euclidean**距離など.

○ 因子分析

- 観測値(結果)に影響を与えていると考えられる因子(原因)を求め, 変数を減らす手法. (主成分分析では, 観測値が原因, 新成分が結果となるので, 因果関係が逆になっている)

2.2 活動・交通データを用いないクラスタリング手法

○ SVM (Support Vector Machine)

2クラス ($y \in \{-1, 1\}$) の分類問題を解くためにつくられた学習機械.

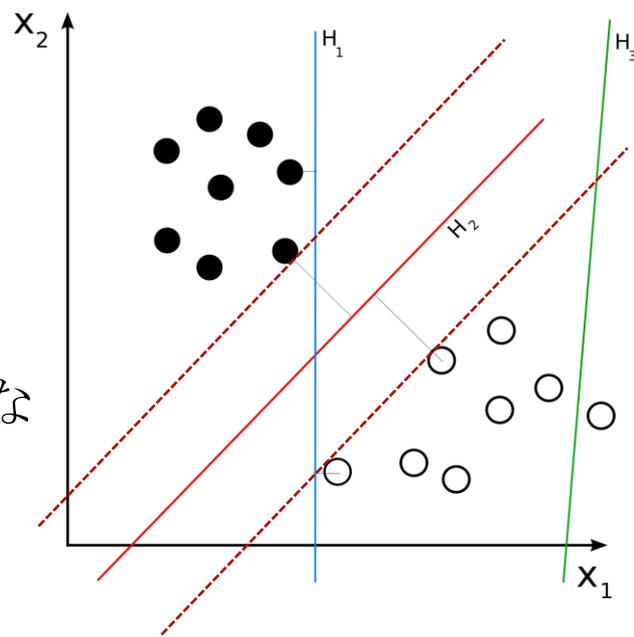
1. マージン最大化を基準とすることで, 高い判別性能を実現.
2. カーネル関数を用いた柔軟なモデリング.
3. パラメータ推定が凸2次最適化問題として定式化されるため, 高速な最適化アルゴリズムが利用可能.

といった, 特徴によって近年注目されている.

2.2 活動・交通データを用いないクラスタリング手法

1. マージン最大化を基準とすることで、高い判別性能を実現。

左図のように点が分布されている場合に、
H1, H2, H3は超平面(境界)候補。
H3は上手く分けられていない→×
H1, H2は点を上手く分けられているが、
点までの距離(マージン)がより大きくなるような
超平面を選ぶアルゴリズムなので、
この場合はH2が超平面になる。



教師データを用いて、上手くこの超平面を生成して、判別していく。

2.2 活動・交通データを用いないクラスタリング手法

1. マージン最大化を基準とすることで、高い判別性能を実現.

先ほどの赤い点線は,

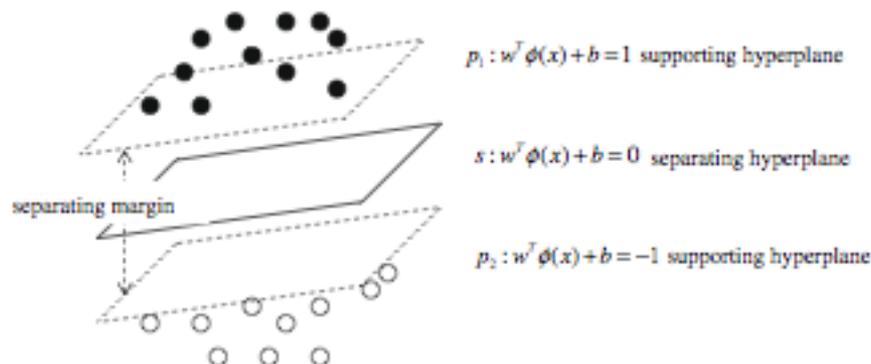
$$P_1: w^T \phi(x_1) + b = 1$$

$$P_2: w^T \phi(x_2) + b = -1$$

という超平面で表される.

但し, w^T は重み付けベクトル, b はパラメータ

$s: w^T \phi(x) + b = 0$ は P_1 と P_2 の中間平面.



マージンは学習データ x_i から, $s: w^T \phi(x) + b = 0$ までの距離

$\rightarrow |w^T \phi(x_i) + b| / \|w\|$ となる. これを最大化する!

この値を2倍した値(P_1 から P_2 までの距離): $w^T[\phi(x_1) - \phi(x_2)] / \|w\|$ を最大化することと同等.

ここで, $w^T[\phi(x_1) - \phi(x_2)] = 2$ および $\min |w^T \phi(x_i) + b| = 1$ という条件を加えると

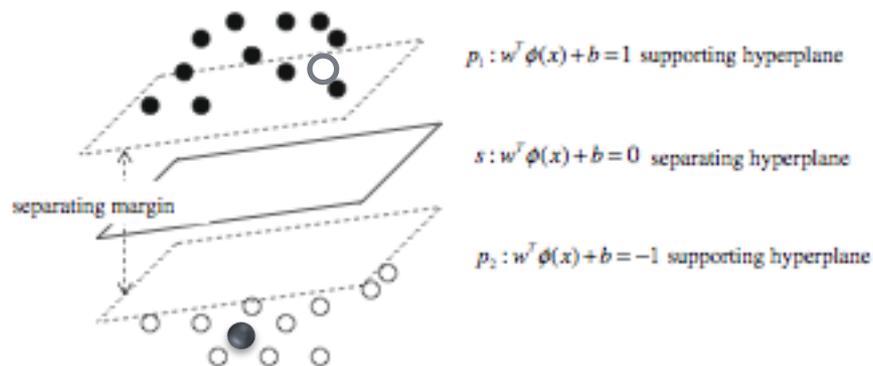
つまりは, $2 / \|w\|$ を最大化

$\Leftrightarrow \|w\| / 2$ を最小化(解きやすくなる)

2.2 活動・交通データを用いないクラスタリング手法

1. マージン最大化を基準とすることで、高い判別性能を実現.

完璧に2つの領域に分けられない場合
白玉 ($y_i=1$) だが, $w^T\phi(x_i) + b \geq 1$ や
黒玉 ($y_i=-1$) だが, $w^T\phi(x_i) + b \leq -1$
にはペナルティを与える.



($y_i=1$) の場合, $w^T\phi(x_i) + b \geq 1 - \xi_i$
($y_i=-1$) の場合, $w^T\phi(x_i) + b \leq -1 + \xi_i$
まとめて書くと, $y_i(w^T\phi(x_i) + b) \geq 1 - \xi_i$

この ξ を用いてペナルティを表現すると,

先ほどの最適化問題 ($\|w\|/2$ を最小化) は,

$\Phi(w, \xi) = w^T w / 2 + C \sum_{i=1}^n \xi_i$ を最小化, 但し $y_i(w^T\phi(x_i) + b) \geq 1 - \xi_i$, $\xi_i \geq 0$ の条件で.

マージンを大きくしようとすると, ペナルティが増加するためトレードオフの関係になる.

2.2 活動・交通データを用いないクラスタリング手法

2. カーネル関数を用いた柔軟なモデリング.

SVMは本来は線形分離不可能な問題には適用できないが、カーネル関数を用いると、パターンを有限or無限次元の空間へ写像でき、その空間で線形分離を行う手法がVapnik(1992)らに提唱された。これで、非線形の問題にも適用できるようになった。また、写像がどのように行われるか意識せずに計算できるので極めて有用。

3. パラメータ推定が凸2次最適化問題として定式化されるため、高速な最適化アルゴリズムが利用可能.

$\Phi(w, \xi) = w^T w / 2 + C \sum_{i=1 \sim n} \xi$ を最小化, 但し $y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i, \xi \geq 0$ の条件で.

という最適化問題は、ラグランジュの未定乗数法を用いて解くのだが、凸二次型となるために初期値に依存せず最適解を求めることができる。

3. 実データへの適用

- 個人属性・世帯の社会経済特性・人口学的特性・アクティビティ交通データなどがわかる大規模な調査.
 - 1: 因子分析とクラスター分析 → ライフスタイル分類
 - 2: 個人属性・世帯の社会経済特性・人口学的特性からSVM
 - 3: 検証.

3.1 データ概要

- 香港トラベル特性調査2002 (TCS2002:the Hong Kong Travel Characteristics Survey 2002)
 - 10年おきに香港住民の移動特性を調べ将来の交通計画に活かす, 香港交通局が主体. 地理的に満遍なく抽出. 3歳以上が対象.
 - 主に, **世帯インタビュー調査**・SP調査・ツアー調査・フォーカスグループ.
 - 調査期間:2002年9月10日～2002年12月21日
 - 調査規模:30005世帯, 92500人以上
 - データ内容:世帯と個人の社会経済人口情報, 就業状況, 車所有, **アクティビティ交通データ**(トリップの発着地, 発着時間, 目的, 交通手段)
- 2001年国勢調査 (the 2001 Population Census)
 - 香港世帯の1/7には聞き取り調査も実施, 国勢調査・統計局
 - 調査期間:2001年3月15-27日(13日間)
 - データ内容:世帯と個人の社会経済人口情報, 就業状況,

データ処理:AM3時を一日の境界, 3歳未満は対象外, 不整合・欠損のあるデータ除外

対象地域:TuenMun, 香港の副都心(面積84.45km², 人口502305人, 人口密度6057人/km²),

TCS2002→6463サンプル, 2001Census→4433サンプル

3.2 アクティビティ旅行パターンを用いたライフスタイル分類

○ 個人特性・世帯特性・都市形態・交通システム

- アクティビティ決定に影響を及ぼす
 - 頻度・活動時間・性質・交通機関選択・出発時間選択など相互に影響・合わさることでライフスタイル群が構成される。

TCS2002 (社会人口学特性・アクティビティ行動データ) から **ライフスタイル分類**

◎利用 **変数**

Age: 年齢, Nmember: 世帯構成員数,

Nchild: 12歳以下の人数, Nelder: 60歳以上の人数,

Ntour: ツアーの数,

Ncompulsory: 義務的アクティビティ数,

Dcompulsory: (同) 時間,

Nmaintenance: 生活維持アクティビティ数,

Dmaintenance: (同) 時間,

Ndiscretionary: 自由アクティビティ数,

Ddiscretionary: (同) 時間,

Nwalkonly: 徒歩のみのトリップ数,

Dwalkonly: (同) 時間,

Npub: 公共交通1回利用トリップ数,

Dpub: (同) 時間,

Nprivate: 自家用車利用トリップ数,

Dprivate: (同) 時間,

N2pub: 公共交通2回以上トリップ数,

D2pub: (同) 時間,

HHdensity: 住宅密度, POPdensity: 人口密度, RETdensity: 小売店密度

3.2 アクティビティ旅行パターンを用いたライフスタイル分類

○ 因子分析 → クラスタ分析

Rotated component matrix (factor loadings).

Variable measured	Factors								
	Urbanform	Walkonly	Regular	Pub2	Privatecar	Maintenance	Discretionary	HHstructure	Age
HHdensity	0.997	0.059	0.001	-0.019	-0.043	0.001	-0.002	-0.015	0.006
POPdensity	0.997	0.058	0.001	-0.019	-0.045	0.001	-0.002	-0.014	0.005
RETdensity	0.997	0.059	0	-0.019	-0.045	0.001	-0.003	-0.014	0.005
Nwalkonly	0.095	0.906	-0.222	-0.137	-0.094	0.08	0.151	0.039	-0.018
Dwalkonly	0.068	0.841	-0.236	-0.153	-0.114	0.015	0.097	0.002	0.003
Ntour	0.042	0.776	0.288	0.132	0.157	0.195	0.107	0.09	-0.064
Npub	-0.017	-0.128	0.929	-0.145	-0.1	0.044	0.123	-0.002	-0.052
Dpub	0.005	-0.117	0.912	-0.126	-0.096	-0.016	0.079	-0.062	-0.016
Dcompulsory	0.05	0.054	0.509	0.312	0.163	-0.546	-0.334	-0.009	-0.169
Ncompulsory	0.019	0.227	0.495	0.246	0.287	-0.563	-0.127	0.014	-0.214
N2pub	-0.027	-0.079	-0.09	0.974	-0.037	-0.013	0.021	-0.039	-0.017
D2pub	-0.029	-0.076	-0.087	0.969	-0.041	-0.017	0.018	-0.04	-0.011
Nprivate	-0.075	-0.024	-0.054	-0.031	0.961	0.017	0.069	-0.008	-0.015
Dprivate	-0.047	-0.033	-0.062	-0.042	0.955	-0.012	0.043	-0.029	-0.018
Nmaintenance	0.016	0.548	-0.022	-0.03	0.053	1	-0.047	0.048	0.061
Dmaintenance	0.01	0.117	0.084	0.076	0.057	0.789	-0.066	-0.032	-0.016
Ddiscretionary	-0.011	0.031	0.043	0.022	0.036	0.034	0.923	-0.036	0.036
Ndiscretionary	0.009	0.297	0.103	0.019	0.093	-0.067	0.849	-0.113	0.079
Nmember	-0.015	-0.043	0.015	-0.021	-0.028	-0.065	-0.06	0.867	0.056
Nchild	-0.022	0.146	-0.078	-0.051	-0.001	0.074	-0.056	0.794	-0.192
Nelder	0.009	-0.006	-0.063	-0.046	-0.059	-0.048	0.019	0.081	0.899
Age	0.008	-0.037	-0.06	0.023	0.038	0.207	0.118	-0.36	0.727

\mathbf{y} : 既知の変数, \mathbf{x} : 因子

\mathbf{y}_1

$$\mathbf{y} = \mathbf{A}\mathbf{x} \rightarrow \mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$$

3. 2 アクティビティ旅行パターンを用いたライフスタイル分類

因子の解釈

- Urbanform: 建物密集度合い(大きいほど密集)
- Walkonly : 徒歩度合い(大きいほど徒歩依存)
- Regular : 日常的に行われる移動(公共交通1回利用)
- Pub2 : 公共交通2回以上利用
- PrivateCar: 自家用車利用
- Maintenance: 生活維持目的
- Discretionary: 娯楽目的
- HHstructure: 世帯構造
- Age : 年齢

3. 2 アクティビティ旅行パターンを用いたライフスタイル分類

○ 因子分析 → クラスタ分析

- 9因子を基底にしてK-means法でクラスタ分析.

○ K-means法

- ① 各データに対してランダムにクラスタを割り振る
- ② 割り振ったクラスタごとにクラスタ重心を求める
- ③ 各データを一番近くにあるクラスタ重心のクラスタで再割り振り
- ④ ②と③を繰り返す
- ⑤ クラスタの割り振りに変更が生じなかったら終了

初期値に依存するので何回かやる? → 統計ソフト(SPSSなど)の利用

- 適切なクラスタ数を設定 → 6クラスタ

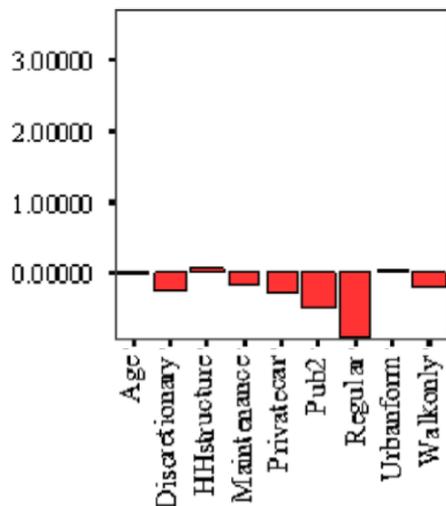
3.2 アクティビティ旅行パターンを用いたライフスタイル分類

○ ライフスタイル分類

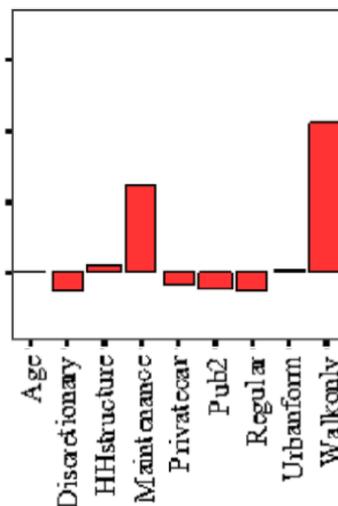
Cluster size and centers for each lifestyle.

Cluster						
	A	B	C	D	E	F
	Home-centered persons	Household makers	Discretionary activity participants	Urban travelers	Regular travelers	High-income residents
Size	2155	536	607	697	2108	360
Percent	33.3%	8.3%	9.4%	10.8%	32.6%	5.6%
Urbanform	0.03939	0.03219	-0.01550	-0.07486	0.02868	-0.28057
Walkonly	-0.19696	2.10682	-0.04131	-0.19199	-0.22589	-0.19371
Regular	-0.89498	-0.24717	0.08471	-0.34111	1.12421	-0.33982
Pub2	-0.48810	-0.20073	-0.19598	2.63828	-0.22727	-0.22608
Privatecar	-0.29695	-0.13658	-0.23220	-0.14842	-0.15143	3.54651
Maintenance	-0.14564	1.22499	0.12366	-0.06909	-0.15520	-0.11796
Discretionary	-0.25827	-0.25802	2.31242	-0.16308	-0.27792	-0.02568
HHstructure	0.08633	0.09871	-0.11725	-0.11991	-0.01829	-0.12678
Age	-0.02764	0.00478	0.46391	-0.02536	-0.09107	-0.04151

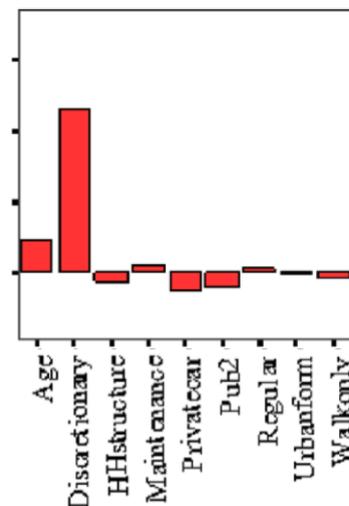
A: Home-centered Persons



B: Household Makers



C: Discretionary Activity Participants



3. 2 アクティビティ旅行パターンを用いたライフスタイル分類

Summary statistics of each lifestyle.

Indicator	A Home-centered persons	B Household makers	C Discretionary activity participants	D Urban travelers	E Regular travelers	F High-income residents
Age 平均	32.75	36.47	46.34	35.73	31.44	37.18
Male (%)	42.14	18.89	47.33	63.72	53.26	72.36
Female (%)	57.86	81.11	52.67	36.28	46.74	27.64
Employed (%) 割合	24.60	21.67	27.99	82.95	70.34	85.4
Unemployed (%)	7.01	3.38	10.68	1.40	0.41	0.93
Students (%)	36.47	20.68	8.29	11.47	25.26	7.45
Homemakers (%)	17.65	41.95	21.36	2.02	2.33	4.35
Retired (%)	11.71	10.34	30.20	1.86	1.45	1.24
Nmember	3.85	3.67	3.40	3.52	3.73	3.40
Nchild	0.76	0.94	0.39	0.44	0.52	0.51
Nelder 平均	0.36	0.28	0.59	0.2	0.21	0.13
Nstudent	1.2	1.36	0.73	0.9	1.06	0.8
Nworker	1.54	1.33	1.41	1.94	2	1.95
Publichouse (%)	39.73	42.35	37.57	35.5	29.92	11.49
Subsidizedhouse (%) 割合	29.41	26.44	30.02	31.32	36.96	23.91
Privatehouse (%)	29.95	31.21	31.68	32.71	32.76	62.73
Houseincome (HK\$) 平均	13 450	15 575	13 100	20 050	19 300	27 650

3.2 アクティビティ旅行パターンを用いたライフスタイル分類

各クラスター

- **A: 家中心型**→シェア**33.3%**. 全ての因子で負なのは, そもそもトリップ自体少ない. 働いてない人や学生が多い.
- **B: 主婦型**→シェア**8.3%**. 娯楽や義務のトリップが少ない. 女性率, 主婦率, **12歳以下の子供がいることが多い**
- **C: 娯楽型**→シェア**9.4%**. 年齢を層が高い. 働いてない人も多い.
- **D: 公共交通複数利用型**→シェア**10.8%**. 労働者が多い. 世帯収入も高め.
- **E: 公共交通1回利用型**→シェア**32.6%**. 通勤に公共交通**1回**利用する人. 労働率が一番高く, 非労働率は一番低い.
- **F: 高収入型**→シェア**5.6%**. 自家用車をよく使う. 個人宅を持っている. 世帯収入高い. 働いている.

3. 3 アクティビティ旅行パターンを用いないライフスタイル分類

○ 社会人口学特性のみ(活動交通パターンなし)での分類

- TCS2002とCensus2001に共通なデータ
 - Age, Gender, Status, Nmember, Nchild, Nelder, Nstudents, Nworker, Housetype, Houseincomeの10変数.
 - 教師データとしてTCS2002の5815サンプル.
- パターン認識の識別機には統計分類・決定木・ニューラルネットワーク・Multinomial Logit Model・SVMがある.

3. 2 のクラスタ結果を真値として, TCS2002データをSVM等にかけて識別機の性能を比較.

識別機	統計分類	決定木	ニューラルネットワーク	MNL	SVM
検知率	45.7%	50.4%	48.6%	49.7%	56.2%

- SVMの識別率が他の識別機よりも5.8%以上すぐれている.
- ただ, それでも56.2%の識別率であった.

3.3 アクティビティ旅行パターンを用いないライフスタイル分類

○ 検証1: 誤検知の理由考察

Cross tabulation of individuals using cluster analysis and SVM.

		Original using cluster analysis		クラスタ分析で分けたライフスタイル			Total	
		A Home-centered persons	B Household makers	C Discretionary activity participants	D Urban travelers	E Regular travelers	F High-income residents	
SVM	A	1306	198	242	69	309	34	2158
	B	60	175	22	6	17	7	287
	C	26	10	110	4	8	2	160
	D	7	0	7	49	9	4	76
	E	465	120	158	502	1582	232	3059
	F	6	0	4	15	7	43	75
Total		1870	503	543	645	1932	322	5815

- 分類上多数を占めるA, Eは比較的上手く分類できている.
- その他もAやE(多数群)に分類されてしまうことが多い.
- クラスタB・CなのにAに分類される・クラスタD・FなのにEに分類される
→ 仕事持っている割合:A・B・Cは30%以下, D・E・Fは70%以上
A・B・Cは主婦・学生・定年後, D・E・Fは労働者

3.3 アクティビティ旅行パターンを用いないライフスタイル分類

○ 検証2: クラスター特性を比較1

Cluster size and centers for each lifestyle.

TCS2002 20変数	Cluster					
	A Home-centered persons	B Household makers	C Discretionary activity participants	D Urban travelers	E Regular travelers	F High-income residents
Size	2155	536	607	697	2108	360
Percent	33.3%	8.3%	9.4%	10.8%	32.6%	5.6%
Urbanform	0.03939	0.03219	-0.01550	-0.07486	0.02868	-0.28057
Walkonly	-0.19696	2.10682	-0.04131	-0.19199	-0.22589	-0.19371
Regular	-0.89498	-0.24717	0.08471	-0.34111	1.12421	-0.33982
Pub2	-0.48810	-0.20073	-0.19598	2.63828	-0.22727	-0.22608
Privatecar	-0.29695	-0.13658	-0.23220	-0.14842	-0.15143	3.54651
Maintenance	-0.14564	1.22499	0.12366	-0.06909	-0.15520	-0.11796
Discretionary	-0.25827	-0.25802	2.31242	-0.16308	-0.27792	-0.02568
HHstructure	0.08633	0.09871	-0.11725	-0.11991	-0.01829	-0.12678
Age	-0.02764	0.00478	0.46391	-0.02536	-0.09107	-0.04151

Cluster size and centers for each lifestyle.

TCS2002 10変数	Cluster					
	A Home-centered persons	B Household makers	C Discretionary activity participants	D Urban travelers	E Regular travelers	F High-income residents
Size	2158	289	160	76	3059	75
Urbanform	0.00712	-0.05674	-0.08781	-0.19036	0.05469	-0.63322
Walkonly	0.00321	1.44893	-0.02501	-0.12202	-0.08465	-0.04987
Regular	-0.47482	-0.29291	-0.30679	-0.2345	0.4408	-0.17531
Pub2	-0.30644	-0.2313	-0.19116	1.62697	0.23516	0.31529
Privatecar	-0.21514	-0.06521	-0.20041	0.08079	0.12789	2.07909
Maintenance	0.13835	1.53277	0.21191	-0.00278	-0.24682	-0.30311
Discretionary	0.04954	-0.01709	1.37795	0.09548	-0.12304	0.24035
HHstructure	0.24749	0.66676	-0.11539	0.03565	-0.2058	-0.23566
Age	0.00164	-0.17669	1.76892	0.10695	-0.10459	0.00746

3.3 アクティビティ旅行パターンを用いないライフスタイル分類

○ 検証2: クラスター特性を比較2

Summary statistics of each lifestyle.

TCS2002

20変数

Indicator	A Home-centered persons	B Household makers	C Discretionary activity participants	D Urban travelers	E Regular travelers	F High-income residents
Age	32.75	36.47	46.34	35.73	31.44	37.18
Male (%)	42.14	18.89	47.33	63.72	53.26	72.36
Female (%)	57.86	81.11	52.67	36.28	46.74	27.64
Employed (%)	24.60	21.67	27.99	82.95	70.34	85.4
Unemployed (%)	7.01	3.38	10.68	1.40	0.41	0.93
Students (%)	36.47	20.68	8.29	11.47	25.26	7.45
Homemakers (%)	17.65	41.95	21.36	2.02	2.33	4.35
Retired (%)	11.71	10.34	30.20	1.86	1.45	1.24
Nmember	3.85	3.67	3.40	3.52	3.73	3.40
Nchild	0.76	0.94	0.39	0.44	0.52	0.51
Nelder	0.36	0.28	0.59	0.2	0.21	0.13
Nstudent	1.2	1.36	0.73	0.9	1.06	0.8
Nworker	1.54	1.33	1.41	1.94	2	1.95
Publichouse (%)	39.73	42.35	37.57	35.5	29.92	11.49
Subsidizedhouse (%)	29.41	26.44	30.02	31.32	36.96	23.91
Privatehouse (%)	29.95	31.21	31.68	32.71	32.76	62.73
Houseincome (HK\$)	13 450	15 575	13 100	20 050	19 300	27 650

クラスタ分析で浮き上がった全ての特徴点(太字)について,

- 全てが特徴としてとらえられている.
- より誇張された結果になった.



Summary statistics of each lifestyle.

TCS2002

10変数

Indicator	A Home-centered persons	B Household makers	C Discretionary activity participants	D Urban travelers	E Regular travelers	F High-income residents
Age	32.21	36.57	62.09	39.76	34.1	43.09
Male (%)	35.82	1.39	70.0	80.26	58.52	97.33
Female (%)	64.18	98.61	30.0	19.74	41.48	2.67
Employed (%)	5.79	12.2	1.88	85.53	84.64	97.33
Unemployed (%)	8.85	0.35	5.63	3.95	0.69	1.33
Students (%)	44.9	2.09	0.0	7.89	14.25	0.0
Homemakers (%)	21.36	81.53	15.63	1.32	0.26	0.0
Retired (%)	16.36	2.09	73.13	0.0	0.1	0.0
Nmember	3.83	3.95	3.43	3.29	3.6	3.24
Nchild	0.85	1.56	0.28	0.88	0.36	0.63
Nelder	0.41	0.11	1.34	0.32	0.18	0.12
Nstudent	1.22	1.73	0.49	1.14	0.93	0.65
Nworker	1.36	1.15	1.06	1.33	2.09	1.84
Publichouse (%)	42.96	40.07	26.88	72.37	28.24	0.0
Subsidizedhouse (%)	27.34	22.3	35.0	14.47	36.48	2.67
Privatehouse (%)	28.87	37.63	36.88	7.89	35.01	90.67
Houseincome(HK\$)	12 125	13 225	11 325	20 860	18 100	28 500

SVMでとても似ているライフスタイル群が生成できた.

3.4 センサスデータを用いてのライフスタイル分類

○ センサスデータをSVMで分類したときのクラスター特性

Summary statistics of each lifestyle.

TCS2002
10変数

Indicator	A Home-centered persons	B Household makers	C Discretionary activity participants	D Urban travelers	E Regular travelers	F High-income residents
Age	32.21	36.57	62.09	39.76	34.1	43.09
Male (%)	35.82	1.39	70.0	80.26	58.52	97.33
Female (%)	64.18	98.61	30.0	19.74	41.48	2.67
Employed (%)	5.79	12.2	1.88	85.53	84.64	97.33
Unemployed (%)	8.85	0.35	5.63	3.95	0.69	1.33
Students (%)	44.9	2.09	0.0	7.89	14.25	0.0
Homemakers (%)	21.36	81.53	15.63	1.32	0.26	0.0
Retired (%)	16.36	2.09	73.13	0.0	0.1	0.0
Nmember	3.83	3.95	3.43	3.29	3.6	3.24
Nchild	0.85	1.56	0.28	0.88	0.36	0.63
Nelder	0.41	1.34	1.34	0.32	0.18	0.12
Nstudent	1.22	1.73	0.49	1.14	0.93	0.65
Nworker	1.36	1.15	1.06	1.33	2.09	1.84
Publichouse (%)	42.96	40.07	26.88	72.37	28.24	0.0
Subsidizedhouse (%)	27.34	22.3	35.0	14.47	36.48	2.67
Privatehouse (%)	28.87	37.63	36.88	7.89	35.01	90.67
Houseincome(HK\$)	12 125	13 225	11 325	20 860	18 100	28 500

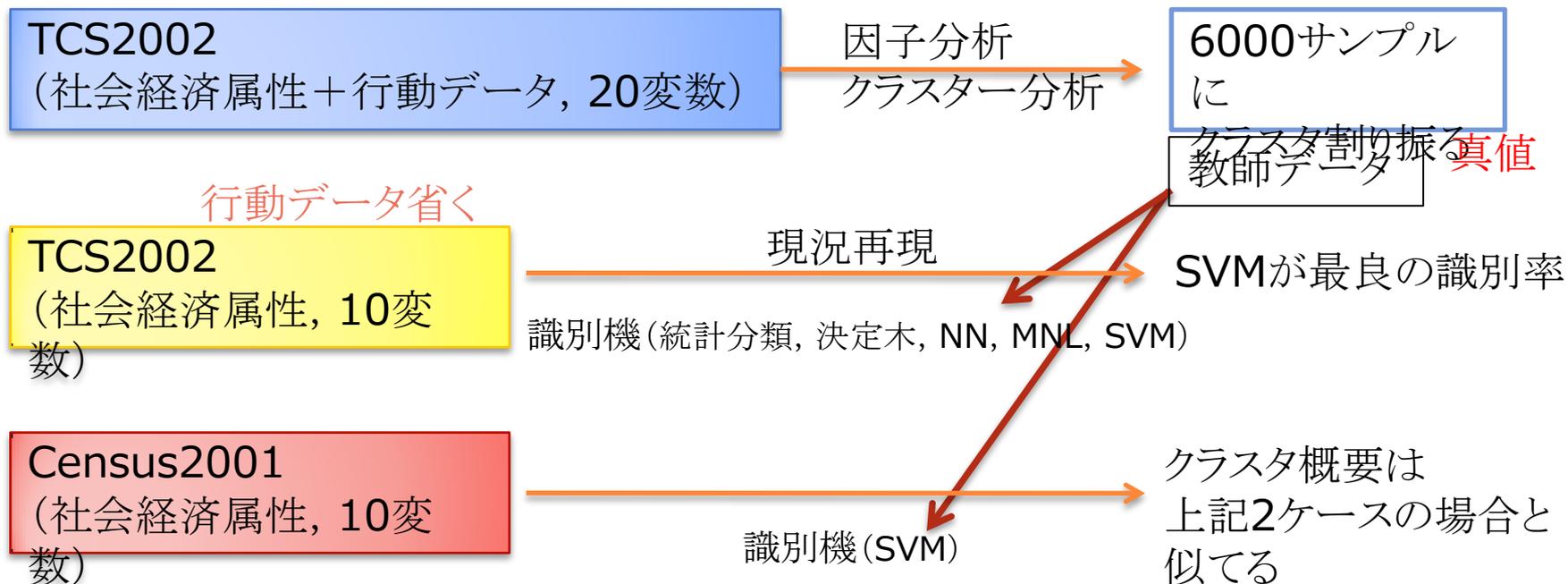
アクティビティ旅行
データをもたない
Census2001から
6グループにわけ
ても、よく似た社会
経済人口学特性
が観測された。

Summary statistics and group size of lifestyle groups using census data.

Census2001
10変数

Indicator	A Home-centered persons	B Household makers	C Discretionary activity participants	D Urban travelers	E Regular travelers	F High-income residents
Size	1660	218	119	48	2322	66
Age	30.87	36.86	63.33	32.15	34.16	41.77
Male (%)	39.04	1.38	77.31	75.00	58.14	98.48
Female (%)	60.96	98.62	22.69	25.00	41.86	1.52
Employed (%)	6.08	7.80	0.00	72.92	84.24	96.97
Unemployed (%)	6.87	0.92	2.52	4.17	0.47	3.03
Student (%)	46.69	0.46	0.00	22.92	14.56	0.00
Homemaker (%)	20.36	88.99	6.72	0.00	0.26	0.00
Retired (%)	14.16	0.46	72.27	0.00	0.17	0.00
Nmember	3.98	3.99	3.66	4.40	3.76	3.50
Nchild	0.94	1.70	0.32	1.67	0.40	0.86
Nelder	0.36	0.05	1.27	0.29	0.20	0.02
Nstudent	1.28	1.78	0.45	1.98	0.91	0.88
Nworker	1.45	1.01	1.29	1.71	2.14	1.94
Publichouse (%)	43.86	37.16	34.45	77.08	35.01	0.00
Subsidizedhouse (%)	25.24	25.69	29.41	12.50	34.80	1.52
Privatehouse (%)	28.98	36.70	35.29	4.17	29.37	83.33
Houseincome(HK\$)	14 325	16 825	17 250	22 550	20 750	28 850

3. 実データへの適用のまとめ



4. 結論と将来の方向性

結論

- 古典的なトリップベースの調査 → 粗雑なデータ(あまり有用でない)
→ アクティビティベースドアプローチの必要性.
- アクティビティ ← 社会経済属性(個人属性・世帯での役割)から決まる
 - アクティビティデータは詳細に必要 → 調査コストの制約
 - アクティビティベースな交通需要のモデリングは複雑すぎる(Shiftan and Ben-Akiva(2006))

アクティビティを無視したトリップベースドアプローチ VS 複雑すぎるアクティビティベースドアプローチ

- 本論文では, 活動(activity)・交通(trip)のパターンを現実味を持たせて, かつシンプルに記述できる代替アプローチを提案・実践(SVM). →そこそこ良いもの.

今後の課題

1. 各クラスターごとに交通需要モデルを作成, 結果を予測に使えるような形で集計.
2. TCS2002から因子を考えたが, より詳細なデータを用いることで, 活動や交通機関選択における意思決定構造をより明確に.
3. 都市形態変数(地域特性や地域の交通政策)を考慮すべき →土地利用・交通政策
4. SVMでの誤分類(44%)の理由を探る. →有用な情報につながる可能性

5. 自分の研究との絡みについて

- BCALs の加速度センサを用いた行動判別

BCALs-1では, 16Hzでの加速度測定や温度・湿度までの細かいデータ取得が可能. この行動データを教師データとする.

BCALs-onlineでは, より少ない変数しかデータ送信されていないが, SVMを用いて行動識別する.

Cross tabulation of individuals using cluster analysis and SVM.

		Original using cluster analysis						Total
		A Home-centered persons	B Household makers	C Discretionary activity participants	D Urban travelers	E Regular travelers	F High-income residents	
SVM	A	1306	198	242	69	309	34	2158
	B	60	175	22	6	17	7	287
	C	26	10	110	4	8	2	160
	D	7	0	7	49	9	4	76
	E	465	120	158	502	1582	232	3059
	F	6	0	4	15	7	43	75
Total		1870	503	543	645	1932	322	5815

○ Table6

Cluster size and centers for each lifestyle.

	Cluster					
	A Home-centered persons	B Household makers	C Discretionary activity participants	D Urban travelers	E Regular travelers	F High-income residents
Size	2158	289	160	76	3059	75
Urbanform	0.00712	-0.05674	-0.08781	-0.19036	0.05469	-0.63322
Walkonly	0.00321	1.44893	-0.02501	-0.12202	-0.08465	-0.04987
Regular	-0.47482	-0.29291	-0.30679	-0.2345	0.4408	-0.17531
Pub2	-0.30644	-0.2313	-0.19116	1.62697	0.23516	0.31529
Privatecar	-0.21514	-0.06521	-0.20041	0.08079	0.12789	2.07909
Maintenance	0.13835	1.53277	0.21191	-0.00278	-0.24682	-0.30311
Discretionary	0.04954	-0.01709	1.37795	0.09548	-0.12304	0.24035
HHstructure	0.24749	0.66676	-0.11539	0.03565	-0.2058	-0.23566
Age	0.00164	-0.17669	1.76892	0.10695	-0.10459	0.00746

Summary statistics of each lifestyle.

Indicator	A Home-centered persons	B Household makers	C Discretionary activity participants	D Urban travelers	E Regular travelers	F High-income residents
Age	32.21	36.57	62.09	39.76	34.1	43.09
Male (%)	35.82	1.39	70.0	80.26	58.52	97.33
Female (%)	64.18	98.61	30.0	19.74	41.48	2.67
Employed (%)	5.79	12.2	1.88	85.53	84.64	97.33
Unemployed (%)	8.85	0.35	5.63	3.95	0.69	1.33
Students (%)	44.9	2.09	0.0	7.89	14.25	0.0
Homemakers (%)	21.36	81.53	15.63	1.32	0.26	0.0
Retired (%)	16.36	2.09	73.13	0.0	0.1	0.0
Nmember	3.83	3.95	3.43	3.29	3.6	3.24
Nchild	0.85	1.56	0.28	0.88	0.36	0.63
Nelder	0.41	0.11	1.34	0.32	0.18	0.12
Nstudent	1.22	1.73	0.49	1.14	0.93	0.65
Nworker	1.36	1.15	1.06	1.33	2.09	1.84
Publichouse (%)	42.96	40.07	26.88	72.37	28.24	0.0
Subsidizedhouse (%)	27.34	22.3	35.0	14.47	36.48	2.67
Privatehouse (%)	28.87	37.63	36.88	7.89	35.01	90.67
Houseincome(HK\$)	12 125	13 225	11 325	20 860	18 100	28 500

Summary statistics and group size of lifestyle groups using census data.

Cluster	A Home-centered persons	B Household makers	C Discretionary activity participants	D Urban travelers	E Regular travelers	F High-income residents
Size	1660	218	119	48	2322	66
Age	30.87	36.86	63.33	32.15	34.16	41.77
Male (%)	39.04	1.38	77.31	75.00	58.14	98.48
Female (%)	60.96	98.62	22.69	25.00	41.86	1.52
Employed (%)	6.08	7.80	0.00	72.92	84.24	96.97
Unemployed (%)	6.87	0.92	2.52	4.17	0.47	3.03
Student (%)	46.69	0.46	0.00	22.92	14.56	0.00
Homemaker (%)	20.36	88.99	6.72	0.00	0.26	0.00
Retired (%)	14.16	0.46	72.27	0.00	0.17	0.00
Nmember	3.98	3.99	3.66	4.40	3.76	3.50
Nchild	0.94	1.70	0.32	1.67	0.40	0.86
Nelder	0.36	0.05	1.27	0.29	0.20	0.02
Nstudent	1.28	1.78	0.45	1.98	0.91	0.88
Nworker	1.45	1.01	1.29	1.71	2.14	1.94
Publichouse (%)	43.86	37.16	34.45	77.08	35.01	0.00
Subsidizedhouse (%)	25.24	25.69	29.41	12.50	34.80	1.52
Privatehouse (%)	28.98	36.70	35.29	4.17	29.37	83.33
Houseincome(HK\$)	14 325	16 825	17 250	22 550	20 750	28 850

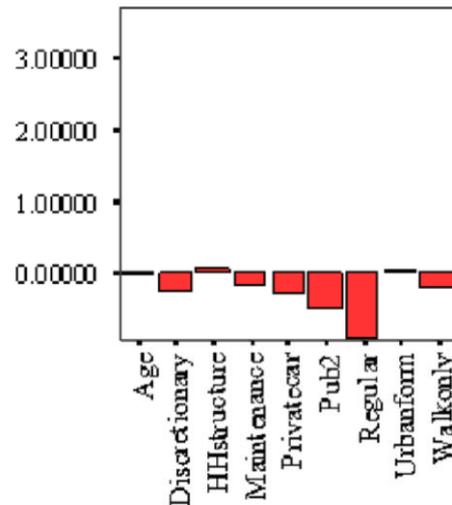
3.2 アクティビティ旅行パターンを用いたライフスタイル分類

○ ライフスタイル分類

Cluster size and centers for each lifestyle

	Cluster
	A
	Home-centered persons
Size	2155
Percent	33.3%
Urbanform	0.03939
Walkonly	-0.19696
Regular	-0.89498
Pub2	-0.48810
Privatecar	-0.29695
Maintenance	-0.14564
Discretionary	-0.25827
HHstructure	0.08633
Age	-0.02764

A: Home-centered Persons



Summary statistics of each lifestyle.

Indicator	A
	Home-centered persons
Age	32.75
Male (%)	42.14
Female (%)	57.86
Employed (%)	24.60
Unemployed (%)	7.01
Students (%)	36.47
Homemakers (%)	17.65
Retired (%)	11.71
Nmember	3.85
Nchild	0.76
Nelder	0.36
Nstudent	1.2
Nworker	1.54
Publichouse (%)	39.73
Subsidizedhouse (%)	29.41
Privatehouse (%)	29.95
Houseincome (HK\$)	13 450

A: 家中心型→シェア33.3%. 全ての因子で負なのは, そもそもトリップ自体少ない. 働いてない人や学生が多い.