データマイニングによる異常検知

東京大学大学院情報理工学系研究科

山西　健司

2015年6月12日

ソーシャルICT研究センター 第3回シンポジウム

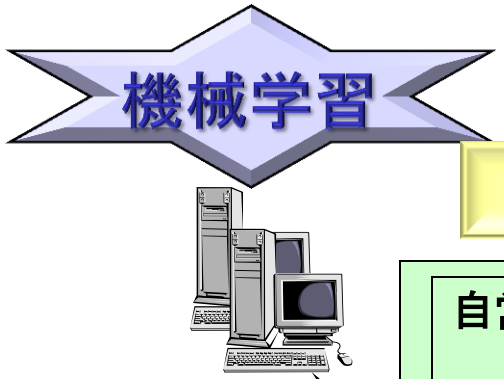
内容

1. データマイニングによる異常検知
2. 変化検知
3. 潜在的構造変化検知
4. 変化兆候検知

1．データマイニングによる異常検知

データマイニングとは

大量のデータに潜む知識を獲得し、将来に向けて活用すること



購買データ

（女、30代、会社員、東京、テニス）、（車種＝セダン）
（男、20代、自営業、千葉、旅行）、　（車種＝バン）
（女、50代、主婦、神奈川、ゴルフ）、（車種＝スポーツ）
（男、30代、自営業、東京、ﾏｰｼﾞｬﾝ）（車種＝なし）
（女、40代、自営業、埼玉、旅行）、　（車種＝バン）
（女、20代、会社員、東京、旅行）、　（車種＝セダン）

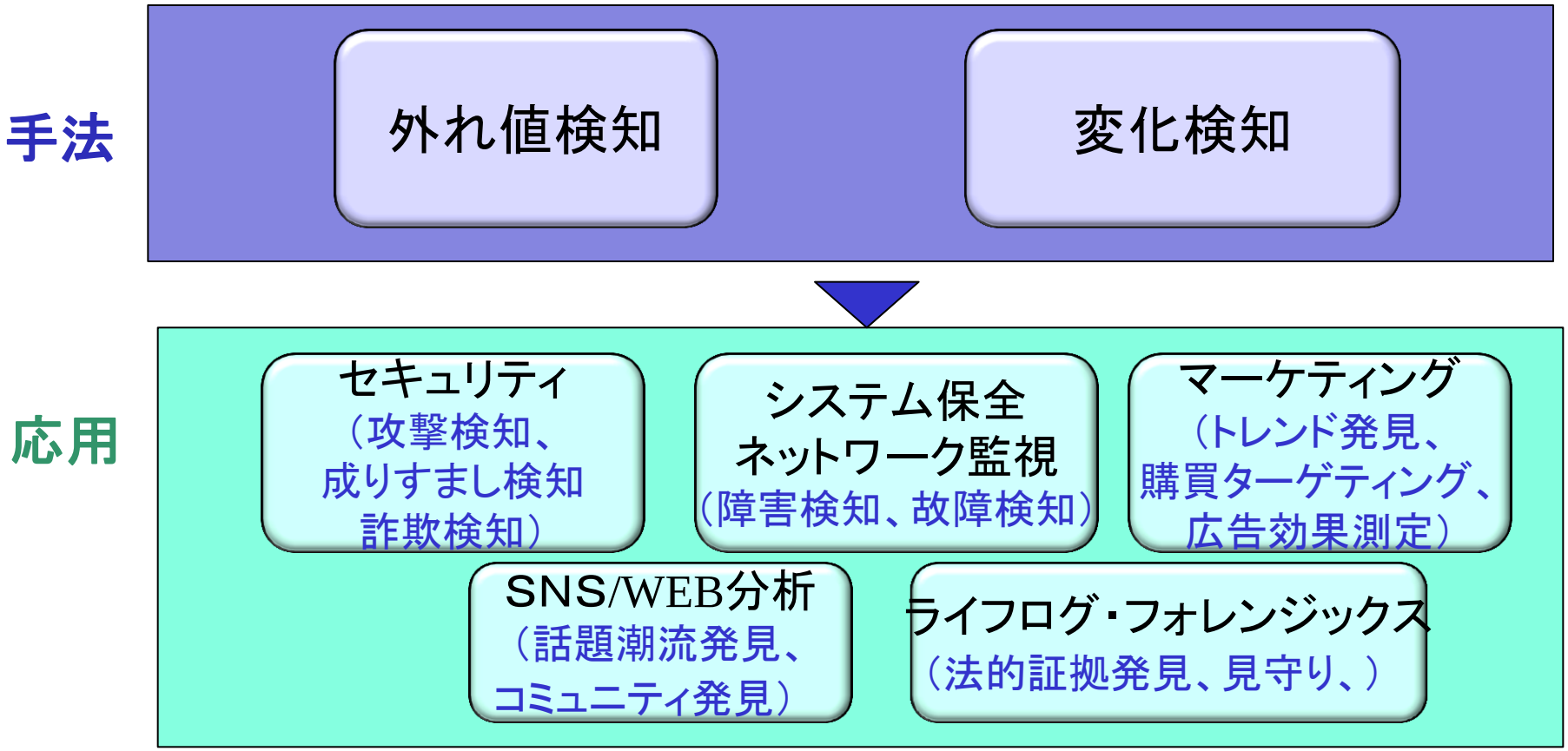
規則の学習

	確率
自営業＆旅行 →（車種＝バン）	0．9
50代会社員＆東京 →（車種＝セダン）	0．8

異常検知

女、70代、車種＝スポーツ

異常検知の体系


手法
外れ値検知
変化検知
応用
セキュリティ
（攻撃検知、
成りすまし検知
詐欺検知）
システム保全
ネットワーク監視
（障害検知、故障検知）
マーケティング
（トレンド発見、
購買ターゲティング、
広告効果測定）
SNS/WEB分析
（話題潮流発見、
コミュニティ発見）
ライフログ・フォレンジックス
（法的証拠発見、見守り、）

外れ値検出

統計的パタンから外れた異常値を抽出

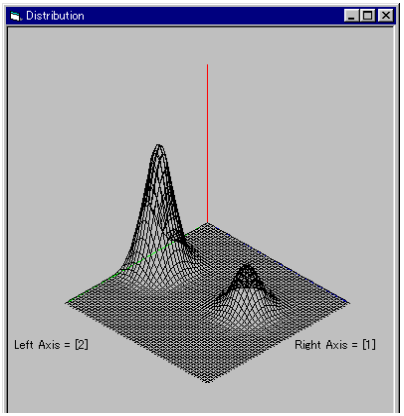
ネットワークのパケットデータ……アクセスログ

NO	Source　IPアドレス X1	接続開始時間 X2	Dest. Portアドレス X3	情報送信量 X4
1	123.136.24.58	11:07:26	206.94.179.38	150
2	127.136.89.45	11:10:34	206.94.179.55	300
3	127.136.89.66	11:34:56	206.94.179.88	328
4	127.136.89.57	11:35:37	206.94.179.52	911
5	127.136.79.551	11:55:19	206.94.179.38	928
	……	·····	·····	·····

オンライン外れ値検知

[Yamanishi et al. KDD2000, DMKD2004]

NO	Source IPアドレス X1	接続開始時間 X2	Dest. Portアドレス X3	情報送信量 X4
1	123.136.24.58	11:07:26	206.94.179.38	150
2	127.136.89.45	11:10:34	206.94.179.55	300
3	127.136.89.66	11:34:56	206.94.179.88	328
4	127.136.89.57	11:35:37	206.94.179.52	911
5	127.136.79.551	11:55:19	206.94.179.38	928
	……	…..	…..	…..



$$P(X) = \sum_{i=1}^{K} \pi_i P(X|\mu_i, \Sigma_i)$$

スコア
＝統計モデルの変化の大きさ
スコア大 ⇔ 異常度合い大

$$-\log P(x_t)$$

ログデータ
トランザクション

↓

統計的モデル
の学習

↓

スコア計算

No	スコア	時間	通信量	ｻｰﾋﾞｽ
1	3.4567	120	34.567	ftp
2	2.3456	110	98,345	http
3	2.1234	9	97.543	http
4	1.2345	23	13.578	telnet
.	…….	…	……	……

攻撃検知への応用

ログ

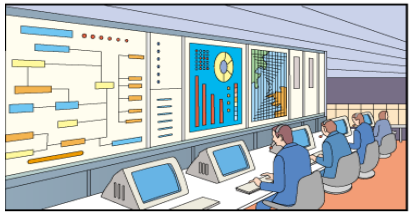
ルールベース・
ポリシーベース

既知の攻撃
の発見

特定ファイルへの
特定操作を監視

IF Source IP Address=XXXX,
Destination Port=80, Protocol=TCP,
Flag=SYN, NoBytes:120
Then 特定マシンへの攻撃

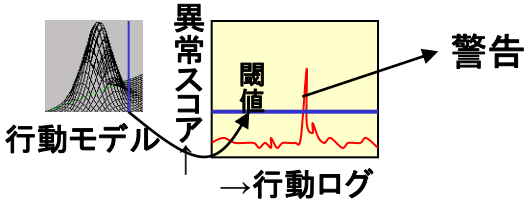
シグネチャ



データマイニング

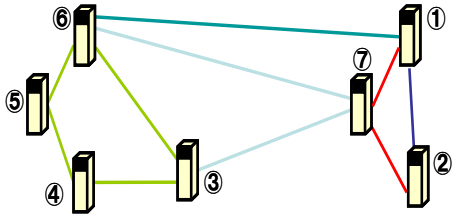
未知の攻撃
の早期発見

不自然なトラフィック、振る舞いを警告

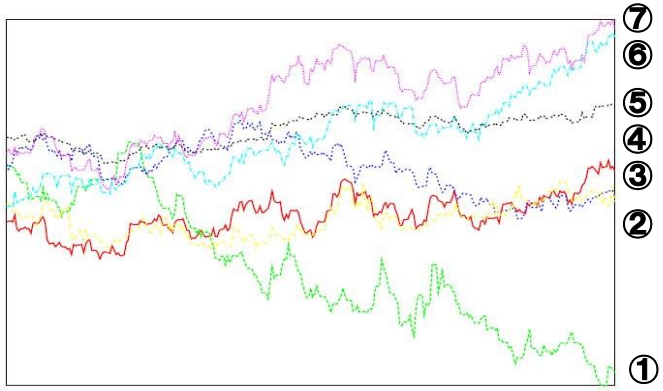


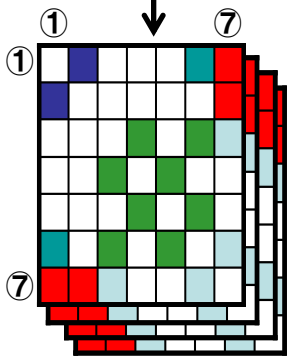
ネットワーク障害検知への応用

[Hirose, Yamanishi, Nakata, Fujimaki KDD2009]



Input(observation) :
Server i's traffic at time t: $u_i(t)$
$(i = 1, 2, \cdots, 7)$
$(t = 1, 2, \cdots)$



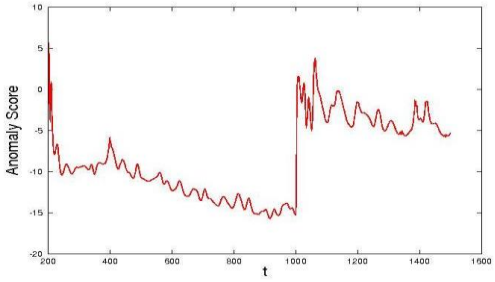


Correlation matrix time series ...

Quantities (eigenvalues) Representing a whole NW

$$\lambda(t) = (\lambda^1, \lambda^2, \cdots, \lambda^m)^\dagger$$

Change-point score/Anomaly Score of $\lambda(t)$

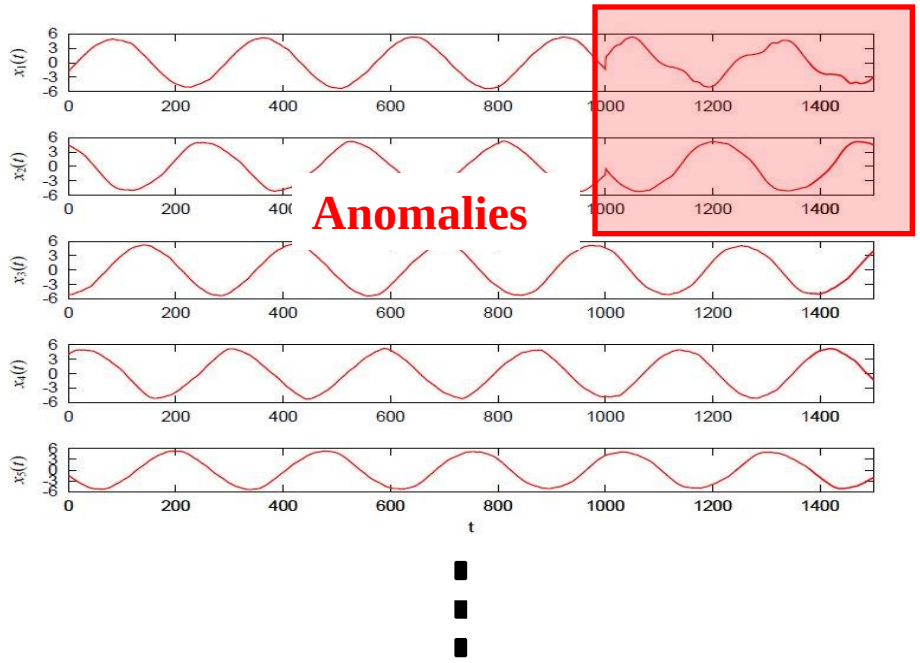


機械系障害検知への応用

Data:

- **Simulating a failure in machinery where several parts broke down,**
  - **currents in 20 different parts of a induction motor.**
  - **failure triggered by the 1st and 2nd parts.**

- **We have used MotorCurrent data**
  (http://www.cs.ucr.edu/~eamonn/TSDMA/index.html).



Changes of correlations

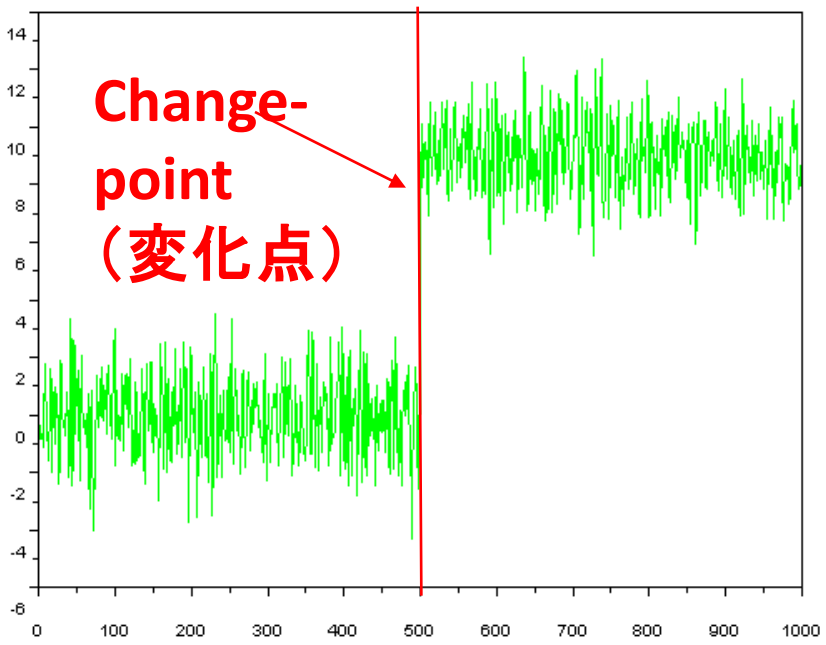
caused by a failure:

Triggered by the 1st and 2nd parts.

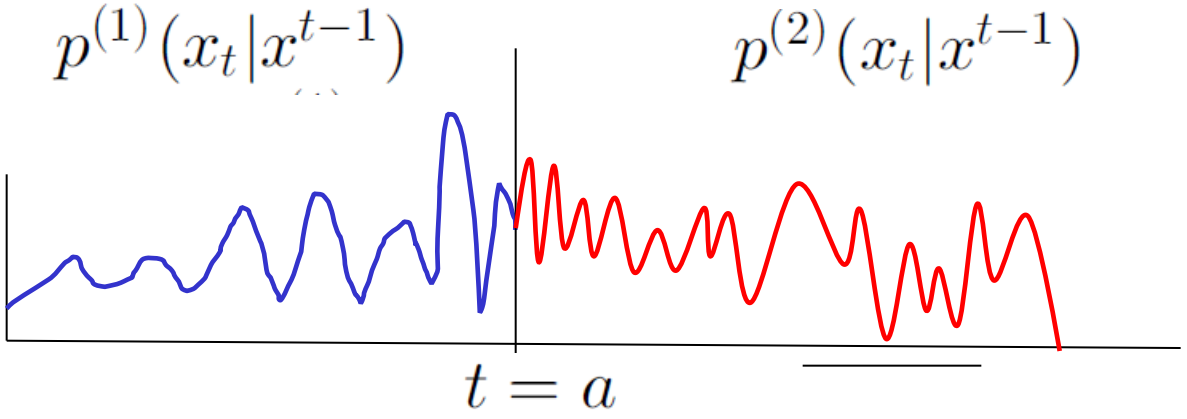
2. 変化点検知

変化点検知とは

時系列データのバースト（塊）的な変化の開始を検出
＝異常の兆候（予兆） を早期に検知



変化点の定義



$$p(x_t|x^{t-1}) = p^{(1)}(x_t|x^{t-1}) \qquad t < a,$$

$$p(x_t|x^{t-1}) = p^{(2)}(x_t|x^{t-1}) \qquad t \geq a.$$

t=a が変化点 ⟺ $D(p^{(2)}||p^{(1)}) \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} E_{p^{(2)}} \log \frac{p^{(2)}(x^n)}{p^{(1)}(x^n)},$

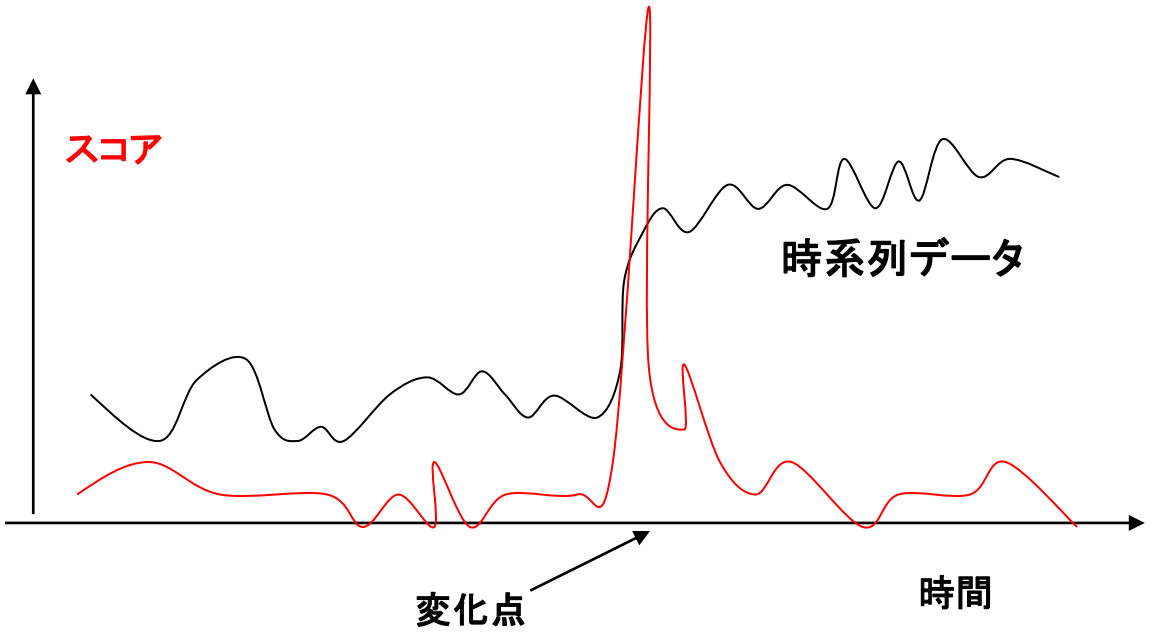
Kullback-Leibler Divergence = Dissimilarity Measure が大

When both $p^{(1)}$ and $p^{(2)}$ are i.i.d.

$$p^{(1)}(x_{a+1}^{a+m}) \approx \exp(-mD(p^{(2)}||p^{(1)})).$$

リアルタイム変化点検出

リアルタイムに変化らしさをスコアリング



⇒　時系列モデルのオンライン学習
＋　モデルの変化度スコアリング

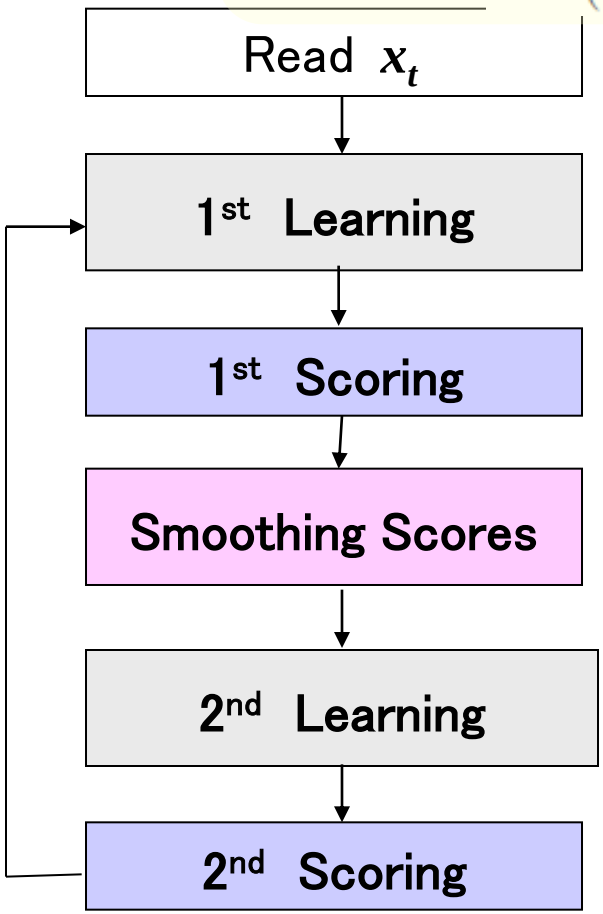
2段階学習による変化点検知

[Yamanishi and Takeuchi KDD 2002]
[Takeuchi and Yamanishi TKDE 2006]

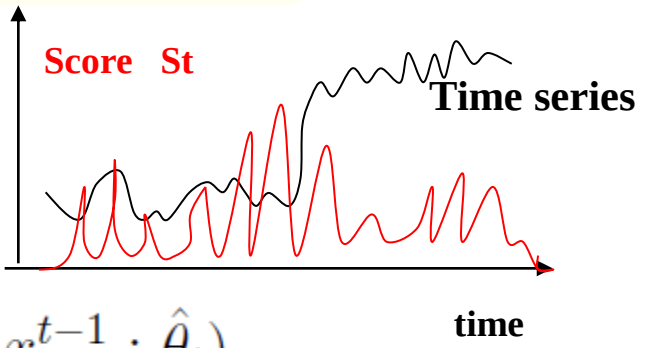
k-th AR(auto-regression) model

$$p(x_t|x^{t-1}:\theta) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2\sigma^2}\left(x_t - \sum_{i=1}^{k} A^{(i)}x_{t-i}\right)^2\right)$$

$$\theta = (A^{(1)}, \cdots, A^{(k)}, \sigma^2)$$



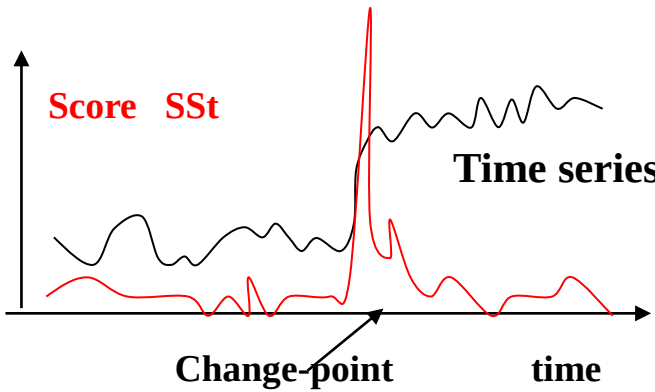
$\hat{\theta}_t$



$$S_t = -\log p(x_t|x^{t-1}:\hat{\theta}_t)$$

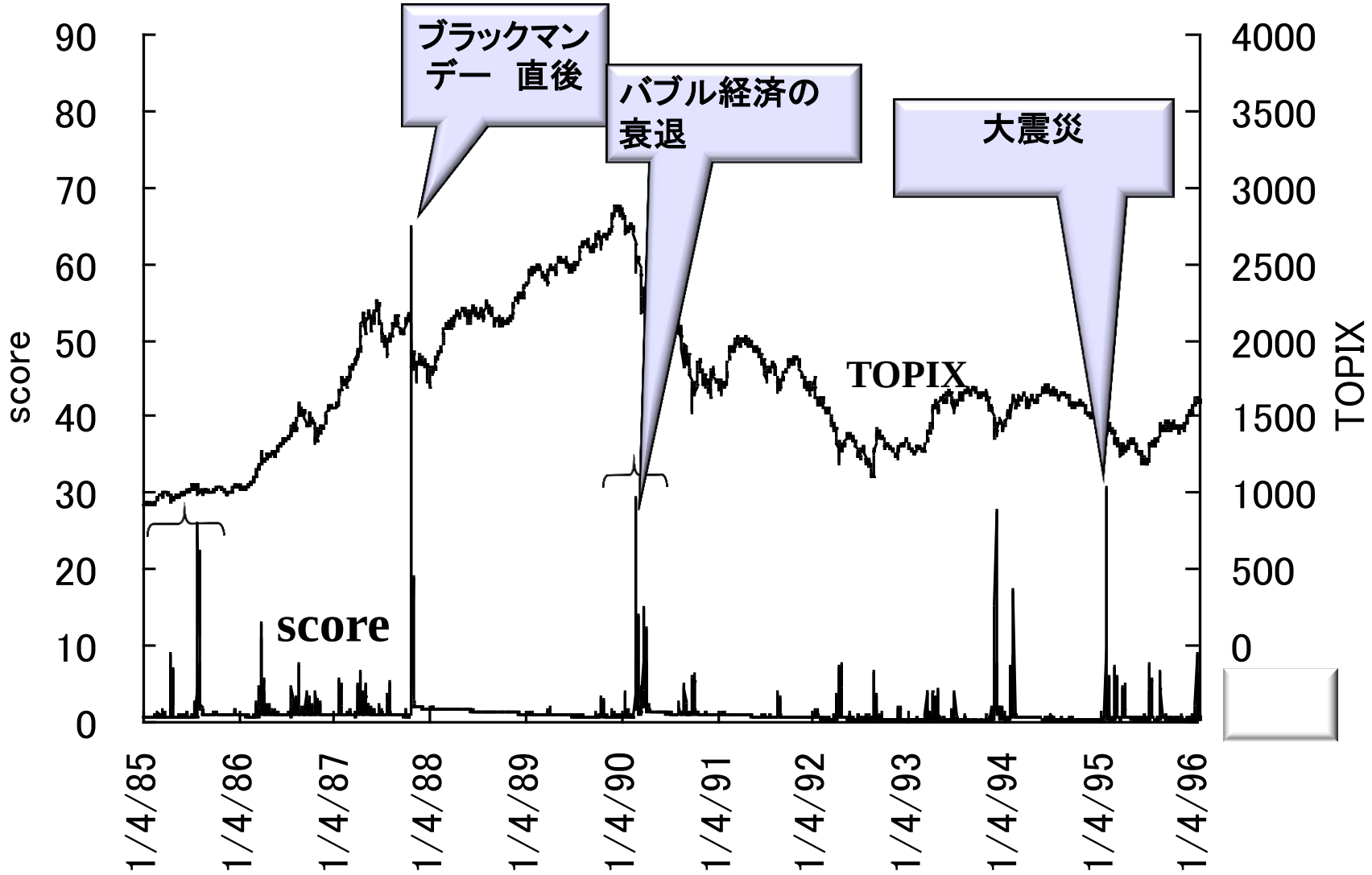
$$y_t = \frac{1}{T}\sum_{j=T-W+1}^{t} S_t$$

$\hat{\xi}_t$



$$SS_t = -\log q(y_t|y^{t-1}:\hat{\xi}_t)$$

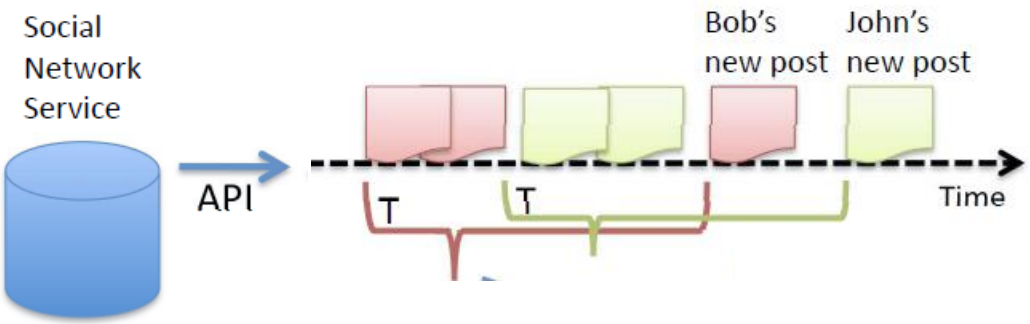
東証株価指数の変化点検出



SNSからの話題出現検知

[Takahashi Tomioka Yamanishi ICDM2011, TKDE2014]

■目的:SNS（Twitterなど）の流れから新しい話題の出現を検知



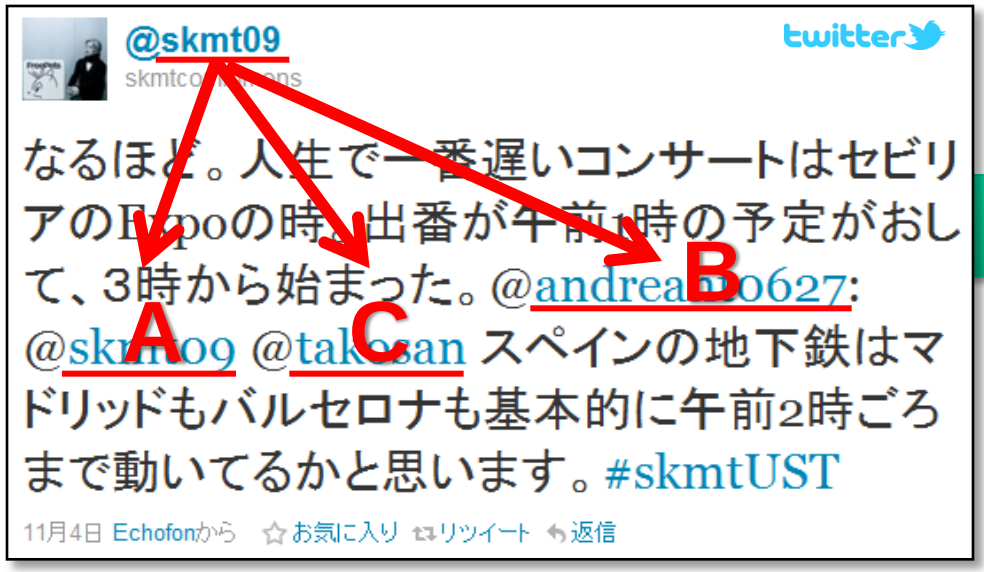
■従来技術の問題点:

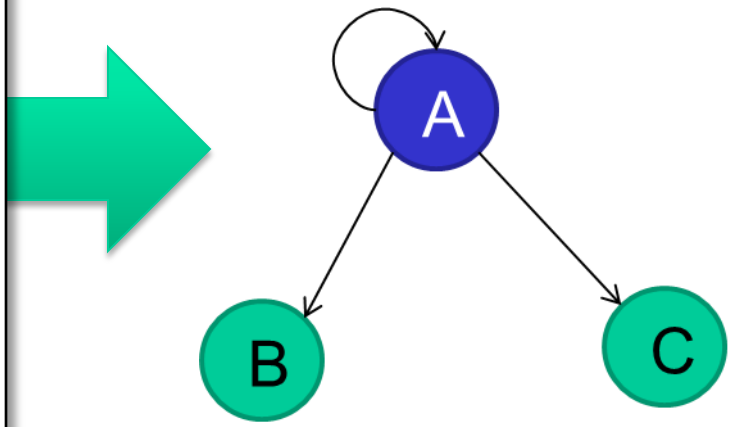
- 言語ベースのアプローチには限界がある
  同義語、同音異義語 など
- トピック自体が多様なメディアで特徴づけられている
  - Image, Video, URLs

⇒ リンク情報だけを用いて話題出現検知をしたい

SNSからの話題出現検知

ユーザのmentionリンクの崩れから話題の発生を検知





$$\mathbf{x} = (t, u, k, V)$$

#mentions

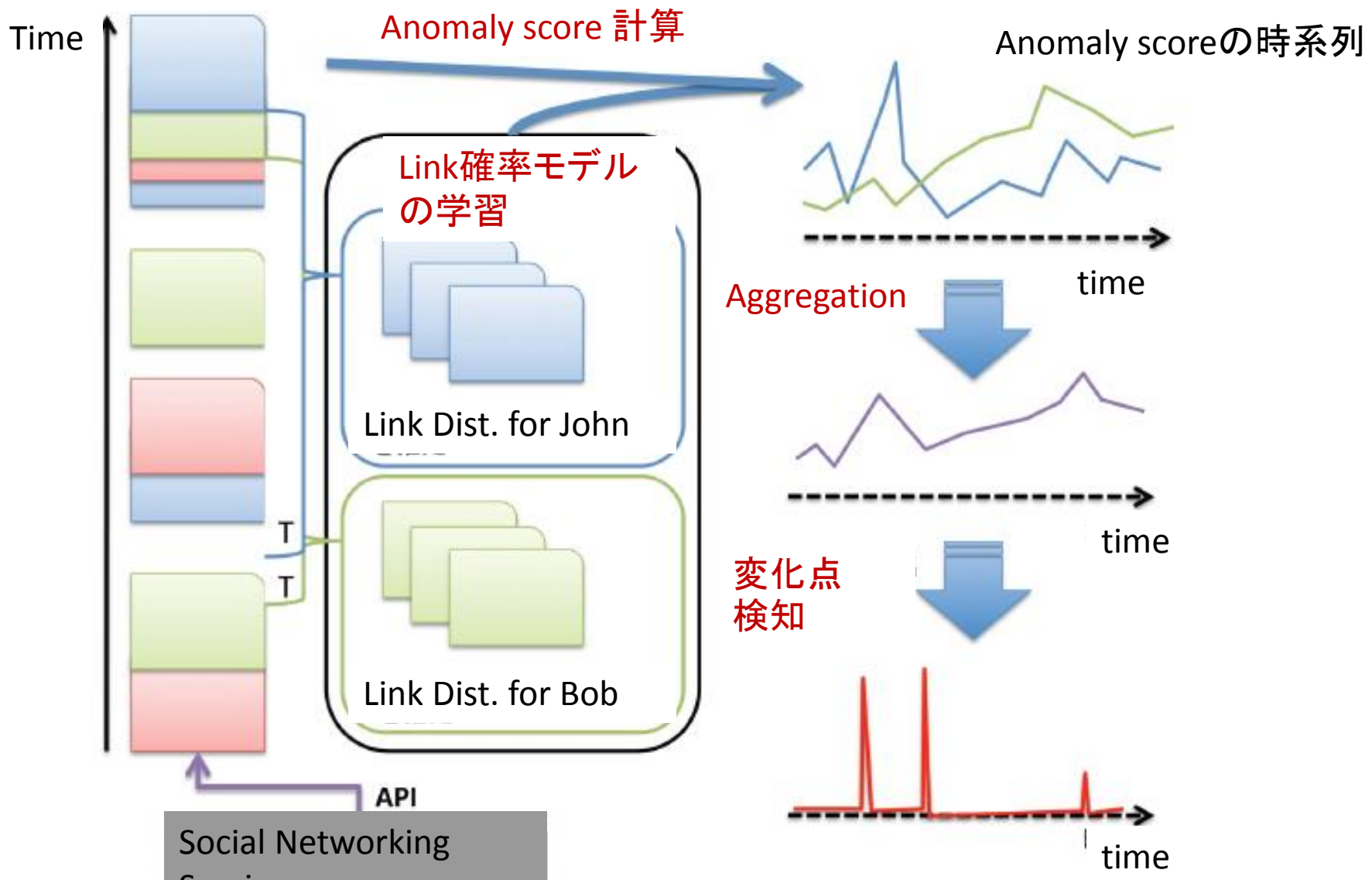
Linking Users

Submit Data (post)

Submit Time

Submit User

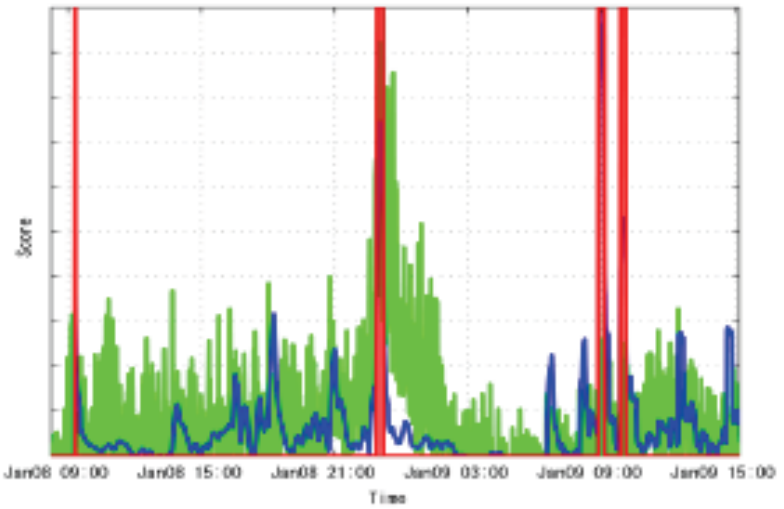
リンクベース話題出現検知の流れ



SNSからの話題出現検知例

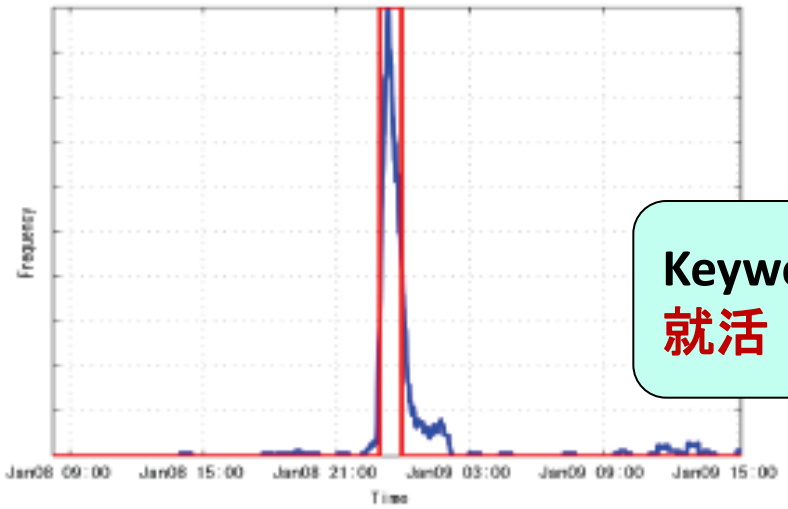
I氏の就職難発言の話題出現検知: 参加者数200

提案手法(変化点検出)



緑:状態値, 青:変化点スコア
赤:アラームの状態
検出時刻・・・22:55
検出数・・・4

比較手法(キーワード頻度)



Keyword:
就活

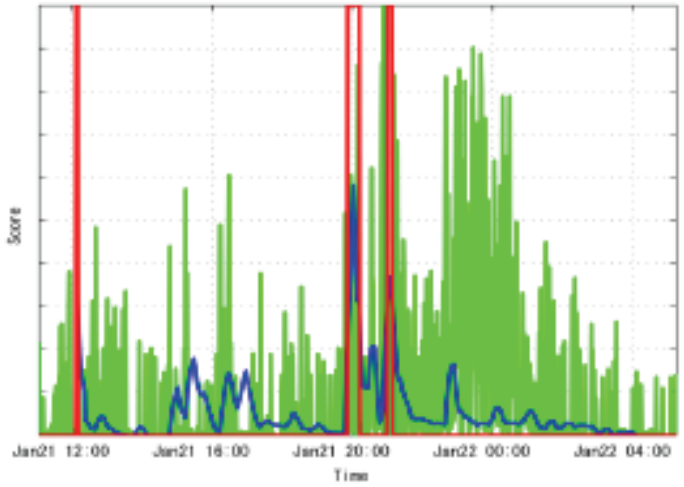
青:キーワードの登場頻度(/30分)
赤:アラームの状態
検出時刻・・・22:57
検出数・・・1

Linkベースがキーワードベースとほぼ同じ速さで話題出現を検知

SNSからの話題出現検知(9/9)

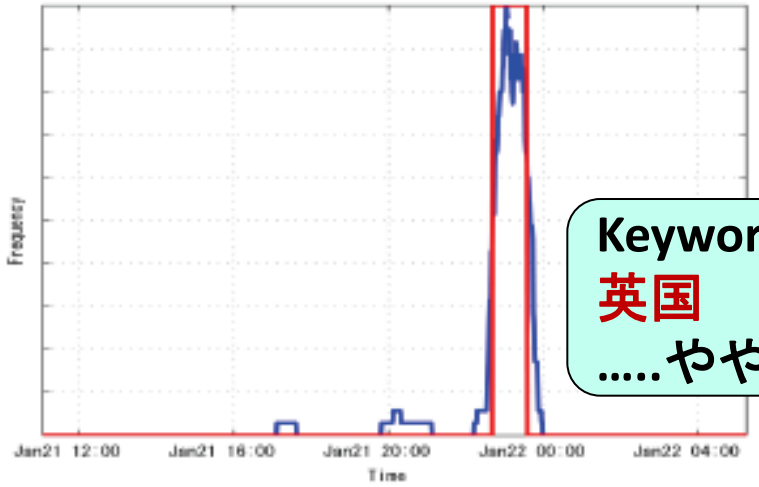
BBC問題発言の話題出現検知: 参加者数47

提案手法(変化点検出)



緑:状態値, 青:変化点スコア
赤:アラームの状態
検出時刻・・・19:52
検出数・・・3

比較手法(キーワード頻度)



Keyword:
英国
.....やや曖昧

青:キーワードの登場頻度(/30分)
赤:アラームの状態
検出時刻・・・22:41
検出数・・・1

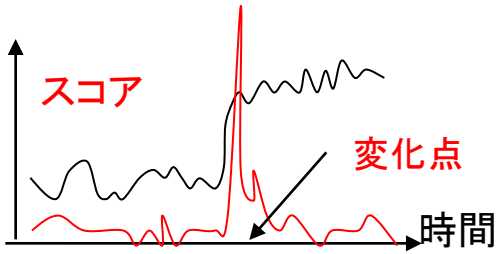
リンクベースはキーワードベースよりも早期に話題検知している

３．潜在的構造変化検知

変化検知の分類

顕在的変化検知

データの表層的性質が変化する



潜在的構造変化検知

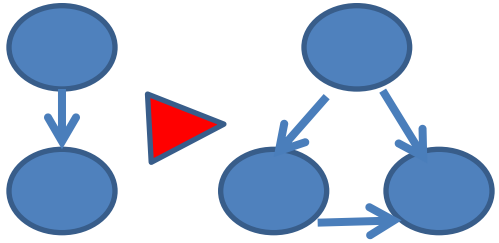
データの潜在的構造が変化する

第一の潜在的ダイナミクス

潜在変数の状態が変化

第二の潜在的ダイナミクス

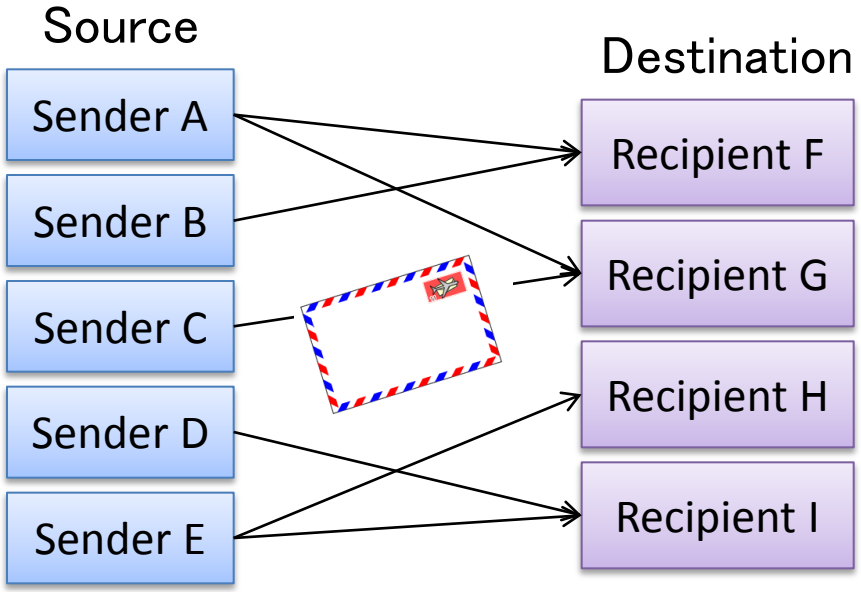
潜在空間の構造が変化



グラフ分割構造変化検知

2部グラフデータ

例
- 都道府県間の人口移動（何県から何県へ何人）
- メールの送受信件数（誰が誰に何通）
- SNS上のやり取り etc…



グラフ分割

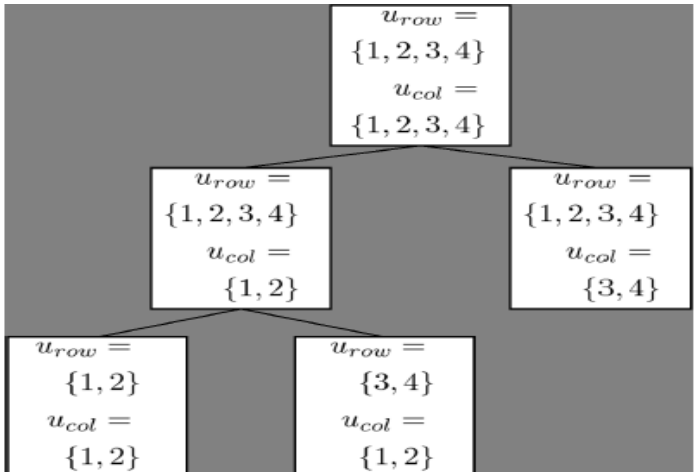
$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{cccc} F & G & H & I \end{array} \\ \left(\begin{array}{cccc} 1 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 3 & 1 \end{array} \right) \end{array}$$

木構造を用いたグラフ分割構造変化検知

Key Idea:　グラフ分割構造変化検知＝木構造変化検知

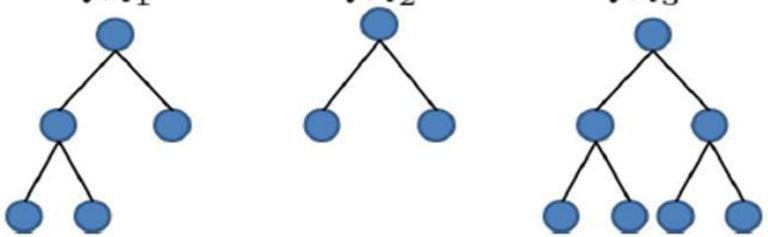
[Sato Yamanishi ICDM 2013]

$$G \overset{1}{\underset{2}{\underset{3}{\underset{4}{}}}} \begin{pmatrix} 1 & 2 & 5 & 3 \\ 2 & 1 & 4 & 7 \\ 9 & 8 & 6 & 3 \\ 12 & 9 & 7 & 4 \end{pmatrix} \quad \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \quad =$$



$$\overbrace{G_1, G_2, G_3}^{\mathcal{G}^1}, \overbrace{G_4, G_5}^{\mathcal{G}^2}, \overbrace{G_6, G_7, G_8}^{\mathcal{G}^3}, \ldots$$

$$\underbrace{M_1, M_2, M_3}_{\mathcal{M}_1}, \underbrace{M_4, M_5}_{\mathcal{M}_2}, \underbrace{M_6, M_7, M_8}_{\mathcal{M}_3}, \ldots$$



グラフ分割構造変化検知

動的モデル選択

For given graph sequence: $\mathcal{G} = G_1, G_2, \ldots, G_T$ (T: data size)
Select a tree sequence: $M^t = M_1 M_2 \ldots M_t$
so that the ***DMS (dynamic model selection)*** is minimum:

$$\mathcal{L}(\mathcal{G}; \mathcal{M}) \stackrel{\text{def}}{=} \sum_{t=1}^{T} (-\log P(G_t|M_t)) + \sum_{t=1}^{T} (-\log P(M_t|M^{t-1}))$$

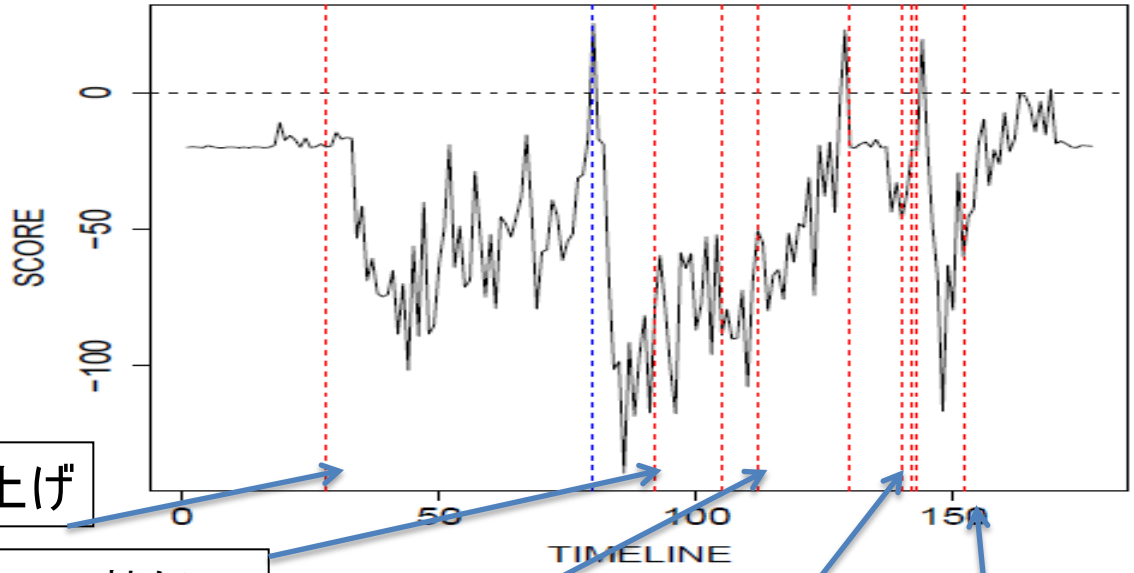
木構造系列に対するグラフ系列の予測符号長

木構造系列の予測符号長.

Minimum Description Length (MDL) Principle [Rissanen 78]
DMS…..MDL のモデルが変化する場合への拡張
[Yamanishi and Maruyama KDD2005, IEEE IT 2007]

Enron社のEvent検知

ENRON社内でのemailのやり取りの有無を週単位で記録
(1999年1月～2002年7月） #sender=#recipient=150



ENRON社立ち上げ

J.Skilling CEO就任

Jエネルギー部門売却

FBI捜査開始

会計改革法案提出

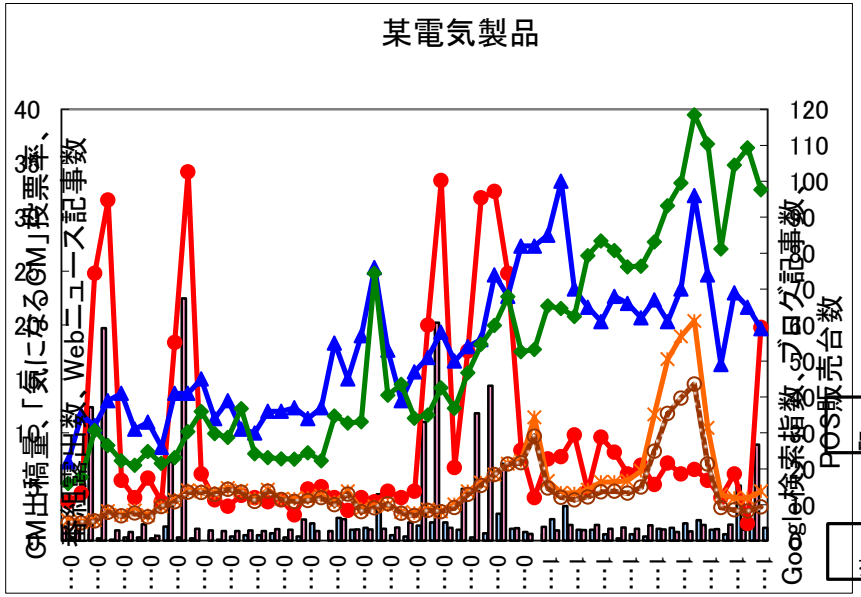
スコア＝構造変化がない場合のDMS規準
－構造変化がある場合のDMS規準

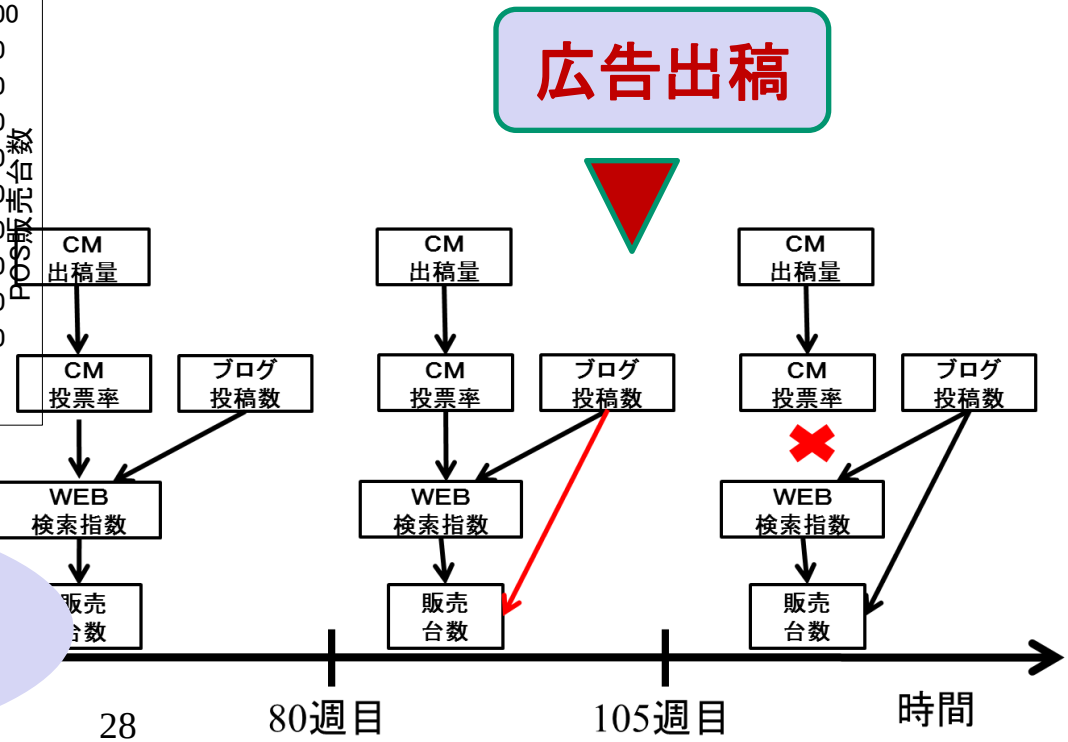
ベイジアンネット構造変化検知

多次元時系列データからベイジアンネットワークの構造変化検知により広告効果を測定

[Hayashi Yamanishi ICDM2012, DMKD2015]

データ提供： 博報堂





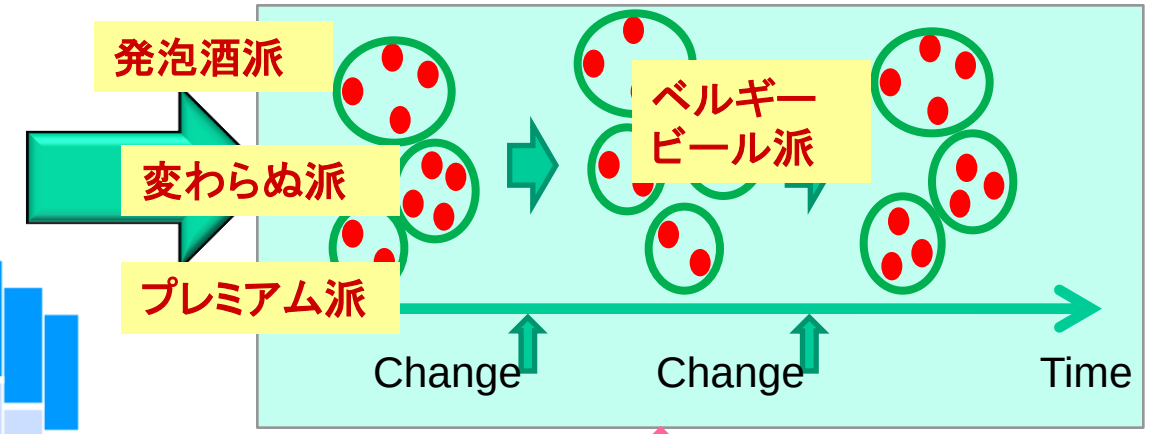
広告出稿により
ブログでの評判が高まり
売り上げに結び付いた

クラスタリング構造変化検知

クラスタリング構造変化検知によるマーケットの構造変化検知

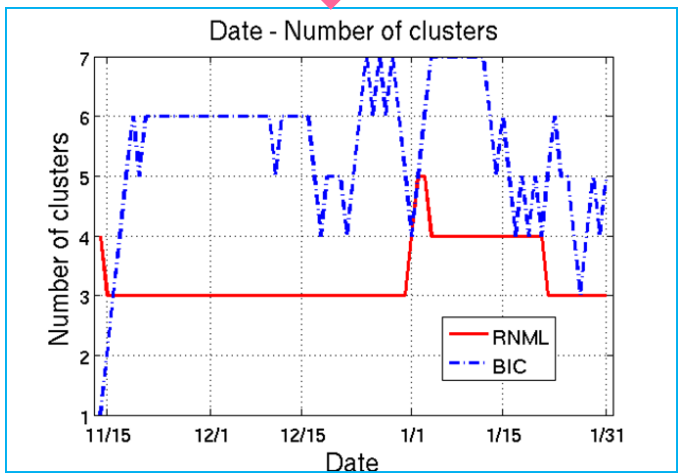
[Hirai Yamanishi KDD2012, IEEE IT2013]

購買層クラスター系列



多変数時系列

	Beer 1	Beer 2	. . .
User 1	350	700	. . .
User 2	1050	350	. . .
. . .	. . .	. . .	. . .



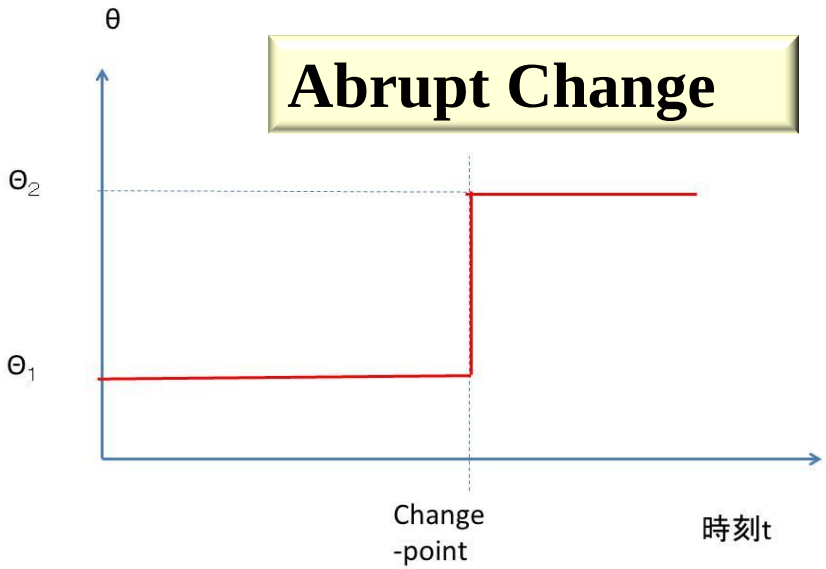
データ提供： マクロミル、博報堂

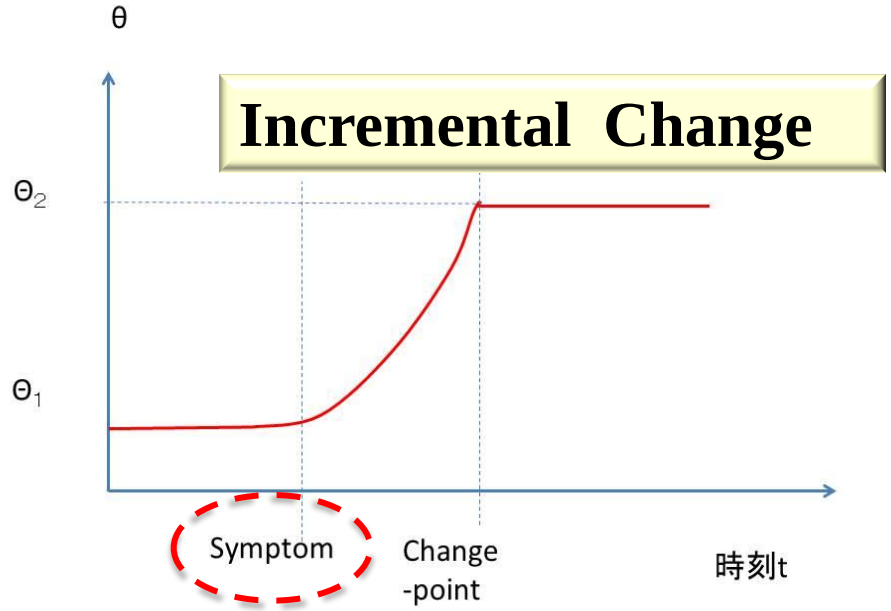
Inc.

4. 変化兆候検知

変化兆候検知

変化は徐々に起こる⇒変化の開始点を検知





Change Detection
パラメータ値が
不連続に変化

Change Symptom Detection
パラメータ値が
徐々にに変化

おわりに

・変化にこそ知識発見あり　（Change = Novelty）

　…..異常が異常でなくなるとき

・潜在的構造変化検出　（Latent Dynamics）

　…..潜在世界にこそ変化の本質がある

・変化兆候検知　(Symptom Detection)

　…,,変化の開始を見極める　⇒　　Challenge!