時系列サブシーケンスマイニングの再展開

安藤　晋
群馬大学工学研究科情報工学専攻
第8回IBISML研究会（12．3．12）

サブシーケンスマイニングのチャレンジ

- “Clustering Time Series Subsequence is Meaningless” [Keogh03]
- “… some anomalous subsequences are not found in sparse regions“ [Keogh04]
- Follow-ups and suggestions [Ide06,Goldin06]

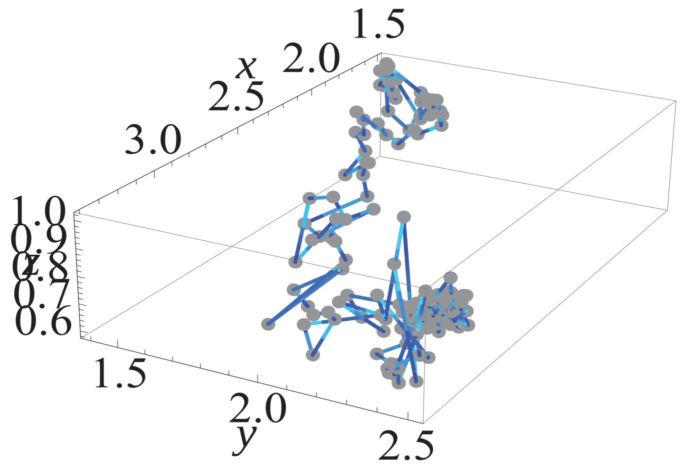
- 本発表:
  - サブシーケンスデータのトポロジー(位相)に関して考察
  - マイニング目的の再設定
  - 位相的手法・学習問題への応用

発表の内容

- 背景
  - 時系列サブシーケンスマイニングの目的とテクニック
  - サブシーケンスデータに関する実験報告
- サブシーケンスデータの位相的分析
  - 低次サブシーケンスの可視的解析
  - 高次サブシーケンス構造に関する考察
- サブシーケンスマイニングに関する指針
  - MDLアプローチ
  - 位相ベースアプローチ
  - 比較分析的アプローチ
- サブシーケンスの応用学習問題
  - 早期分類

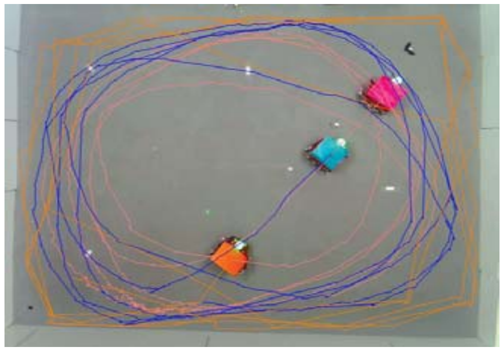
Temporal Knowledge Discovery

- Progressive Events/Dynamics
  - Time series, Trajectories, Data Streams
- Application Domains
  - Financial Time Series
  - Medical Log
  - Behavior Monitoring (Sensor/Surveillance)
  - Meteorological Records
  - Genomic Expressions
- Tasks:
  - Learning: Classification, Regression
  - Knowledge Discovery: Clustering, Association Rules



Subsequence Mining (1)

- Segmented:
  - Time Series=Sample
- Continuous:
  - Time Series= Sample Domain



[Piciarelli06]

[Ando11]

[Yang10]

- Selecting/disregarding relevant/irrelevant subsequence:
  - Difficult problem
- Ex. Movement Epenthesis in Sign Language
  - Word classification ACC: 90% (k-NN)
  - Significant drop-off for subseq. classification
  - Subsequences do not show good clustering tendencies
  - Epenthesis: move to/from initial position

Subsequence Mining (2)

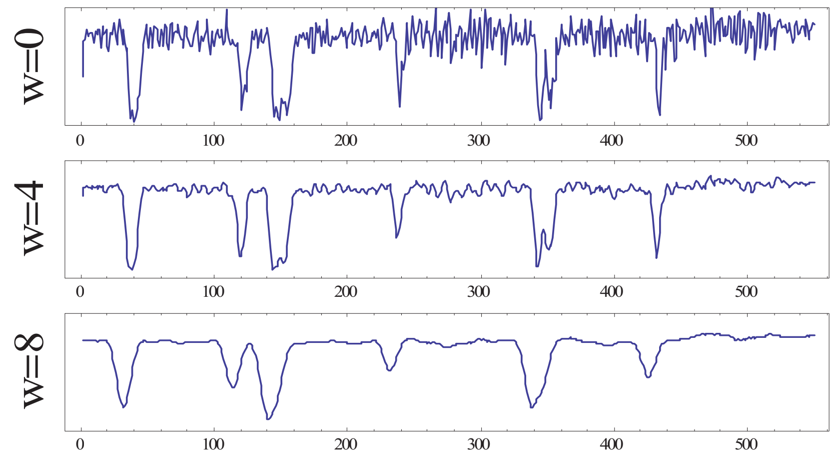
- Symbolic Discovery
  - Observed dynamics: *Continuous*

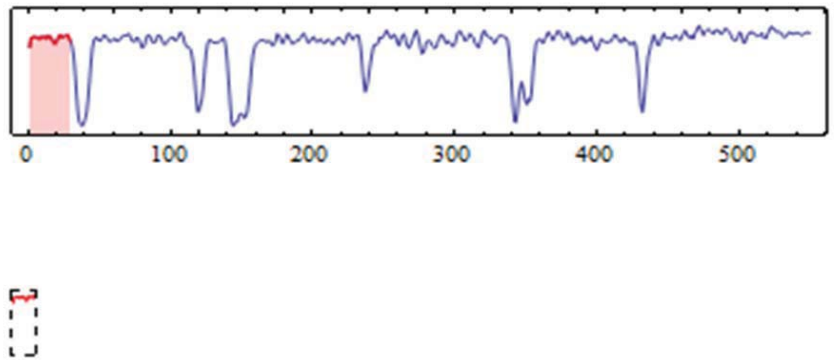
    ⇕ Filling the gap
  - Human Perception: *Symbolic*
- Goal: finding patterns related to relevant events
  - Practice: recurring/rare subsequence patterns within time series
  - Outliers, Clusters, Motifs

E.g. "A tight ***double bottom*** lead to the *move in the direction of longs*"

Preliminaries (Notations)

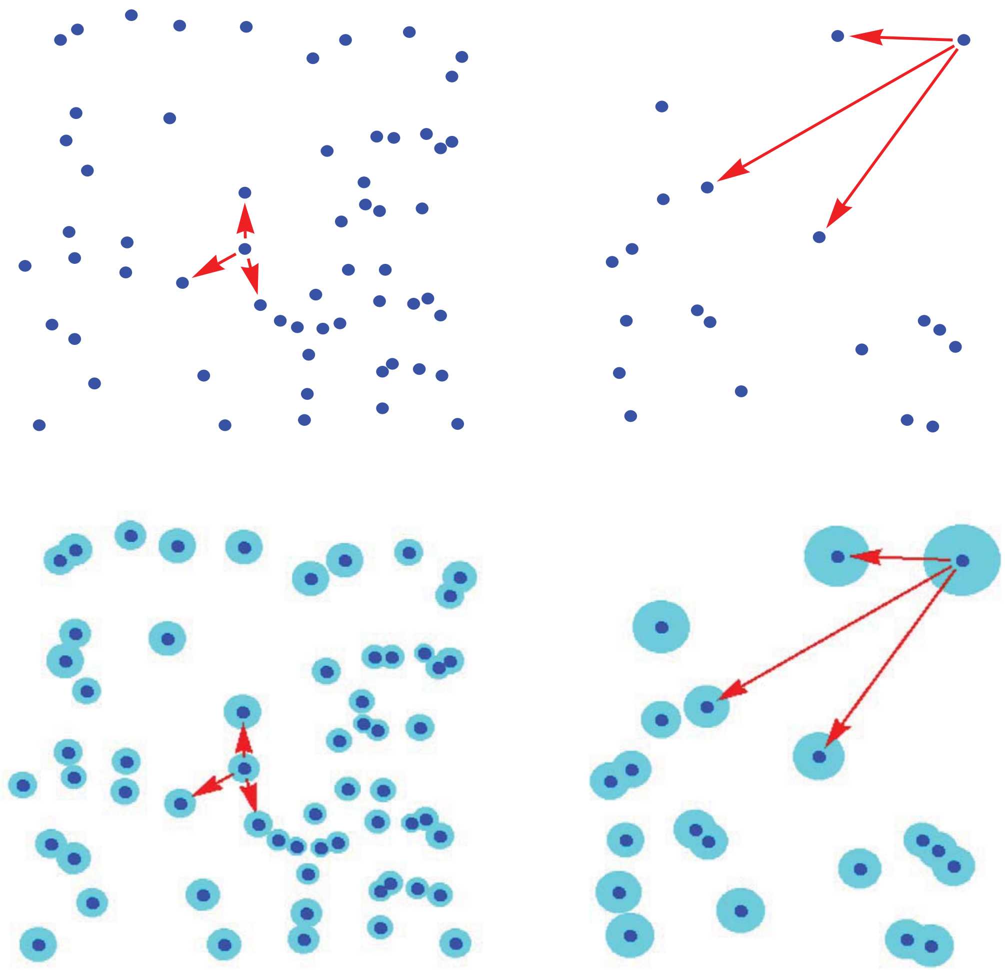
- Time Series: $X = \{x(t)\}_{t=1}^{T}$
- Moving Average
  - Window size: w
  - $f_{MA}: X \in R^T \rightarrow \bar{X} \in R^{T-w+1}$
  - $\bar{x}(t) = \frac{1}{w}\sum_{i=1}^{w} x(t-i+1)$
- Sliding Window
  - Window size: l
  - $f_{SLW}: X \in R^T \rightarrow V \in R^{n \times l}$
  - $V = \{(x(t), \dots, x(t+l-1))\}_{t=1}^{n}$
- Discretization
  - E.g. DWT, iSAX





Preliminaries (Density-based Methods)

- Outliers/anomalies:
  - Instances in the region of lowest-densities
- Density-based approach:
  - **Approximate point density by similarities to neighbors**
- LOF [Breunig00]
  - $N(\mathbf{v}) = \underset{S \subset V : \#S = n}{\operatorname{argmax}} \Sigma_{\mathbf{v}' \in S} D(\mathbf{v}, \mathbf{v}')$
  - $ND(\mathbf{v}) = \max\left(\{D(\mathbf{v}, \mathbf{v}')\}_{\mathbf{v}' \in N(\mathbf{v})}\right)$
  - $RD^{-1}(\mathbf{v}) = \langle \max\left(D(\mathbf{v}, \mathbf{v}'), ND(\mathbf{v}')\right) \rangle_{\mathbf{v}' \in N(\mathbf{v})}$
  - $LOF(\mathbf{v}) = RD^{-1}(\mathbf{v}) / \langle RD^{-1}(\mathbf{v}') \rangle_{\mathbf{v}' \in N(\mathbf{v})}$



Preliminaries: Clustering, Frequent Patterns

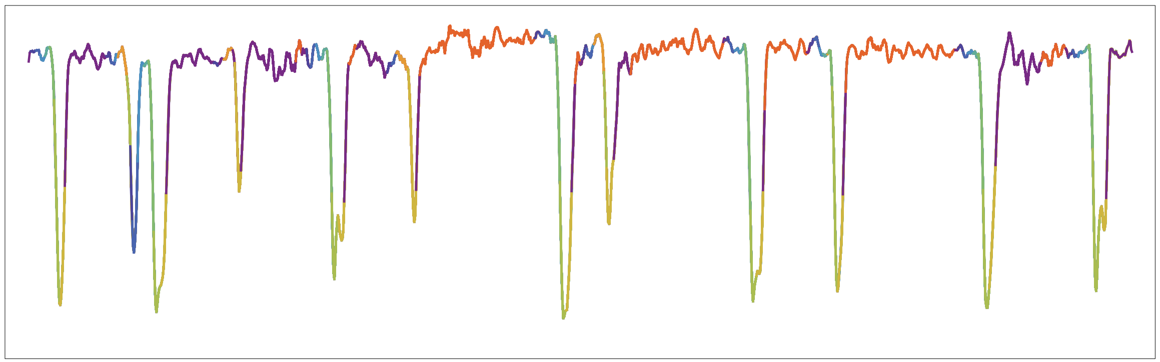
- Motifs
  - $MTF(\mathbf{v};\theta) = \#\{\mathbf{v}': D(\mathbf{v},\mathbf{v}') \le \theta\}$
- k–Means:
  - Minimize function
  - $F\left(V;\{\mu_i\}_{i=1}^{k}\right) = \Sigma_t\{\min_{\mu\in\{\mu_i\}} D(\mathbf{v}(t),\mu)\}$
  - $c_i = \{\mathbf{v}(t): \mu_i = \mathrm{argmin} D(\mathbf{v}(t),\mu)\}$
- Hierarchical Clustering
  - $H_0 = \{c_t: c_t = \{v(t)\}\}_{t=1}^{T'}$
  - $(c_j, c_k) = \mathrm{argmin}_{c_j,c_k \in C_i} D(c_i, c_j)$
  - $H_{i+1} = (H_i \setminus c_i \setminus c_j) \cup (c_j \cup c_k)$
  - $\#(H^*) = 1$
- Spectral Clustering (N–cut [Shi00])
  - Connection matrix: M
  - Graph Laplacian:
  - $L = I - M^{-0.5} D M^{-0.5}$
  - Eigenvectors $\{\mathbf{e}_i\}$
  - $\{\mathbf{v}(t)\} = [\mathbf{e}_1, \dots, \mathbf{e}_k]^\top$

Empirical Discovery (1)

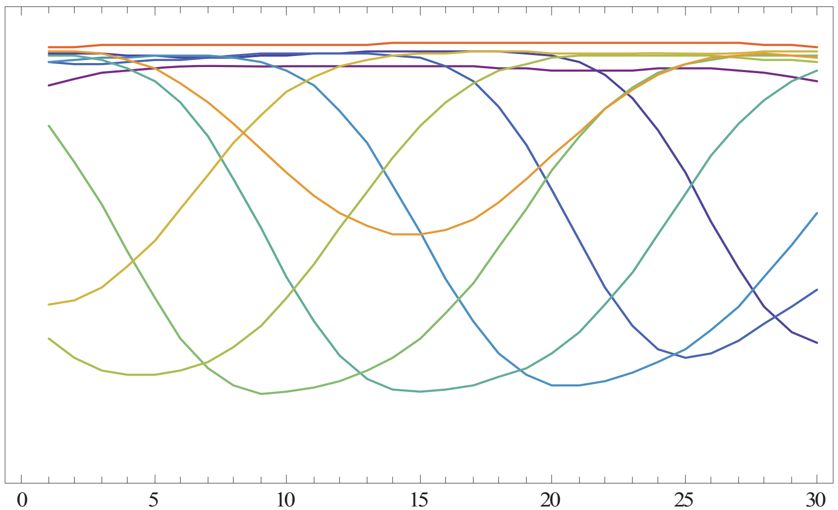
- "Cluster means form sine waves regardless of the original time series" [Keogh03]

- Hypothesis: Subsequence Clustering ≈ Eigen Problem
- Theoretical Analysis of SVD [Ide06]
  - Averaging transition patterns produce sine waves

Time Series

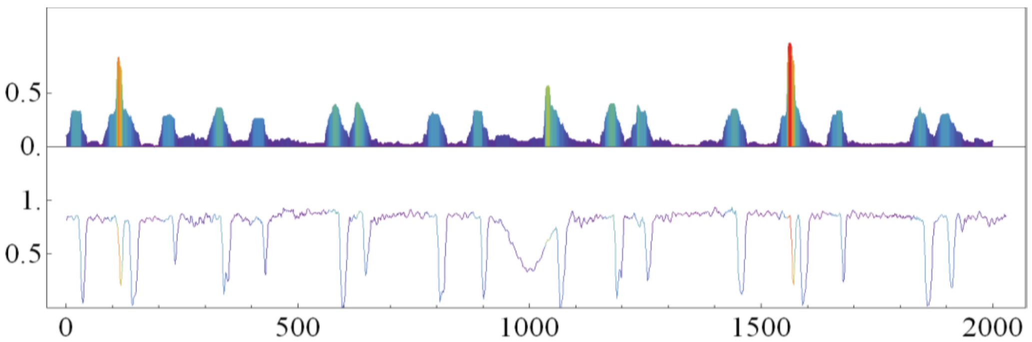


Subsequence Clusters



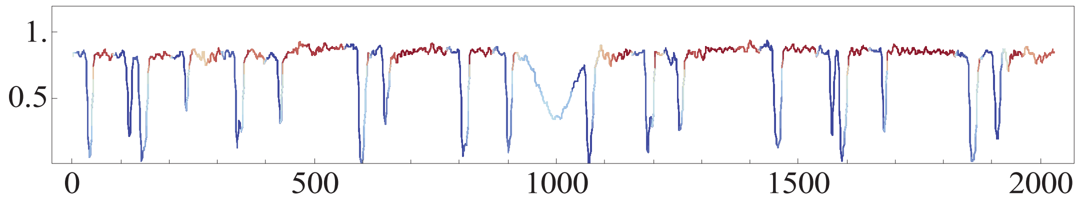
Empirical Discovery (2)

- "anomalous subsequence patterns are not found in the sparse regions" [Keogh03]
- Due to: transition patterns generated by sliding window



$ND(\mathbf{v})$

- Higher frequency → Sparser transition pattern
- Common patterns can be found in "*sparse*" region
- Difficulty finding high frequency Motifs

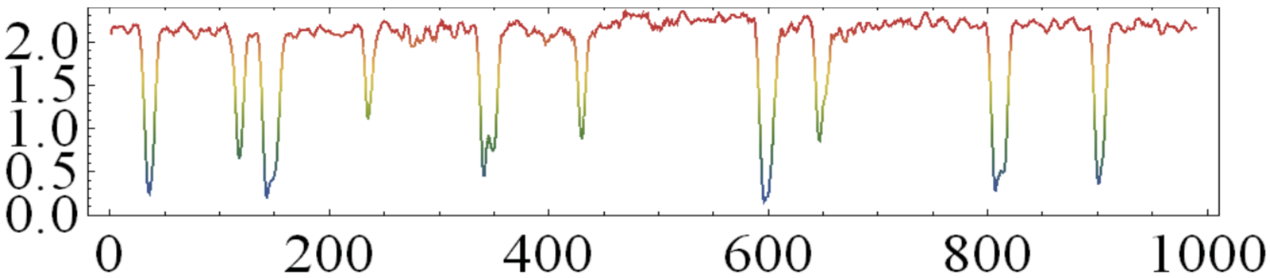


$MTF(\mathbf{v};\theta)$

サブシーケンスデータの
位相に関する考察

Exemplary Data

- Behavior Analysis Benchmark
  - Multi-agent Task experiment
- Velocity Norm Feature:
  - $\|\mathbf{v}(t)\| = \|\Delta\mathbf{x}/\Delta t\|$
  - $\Delta\mathbf{x} = \frac{1}{w+1}\{\sum_{i=0}^{w}\mathbf{x}(t+i) - \sum_{i=0}^{w}\mathbf{x}(t-i)\}$
  - $s_l(t) = (\|\mathbf{v}(t)\|, \dots, \|\mathbf{v}(t+l-1)\|)$
- Approximately 5min.
  - 4500 points

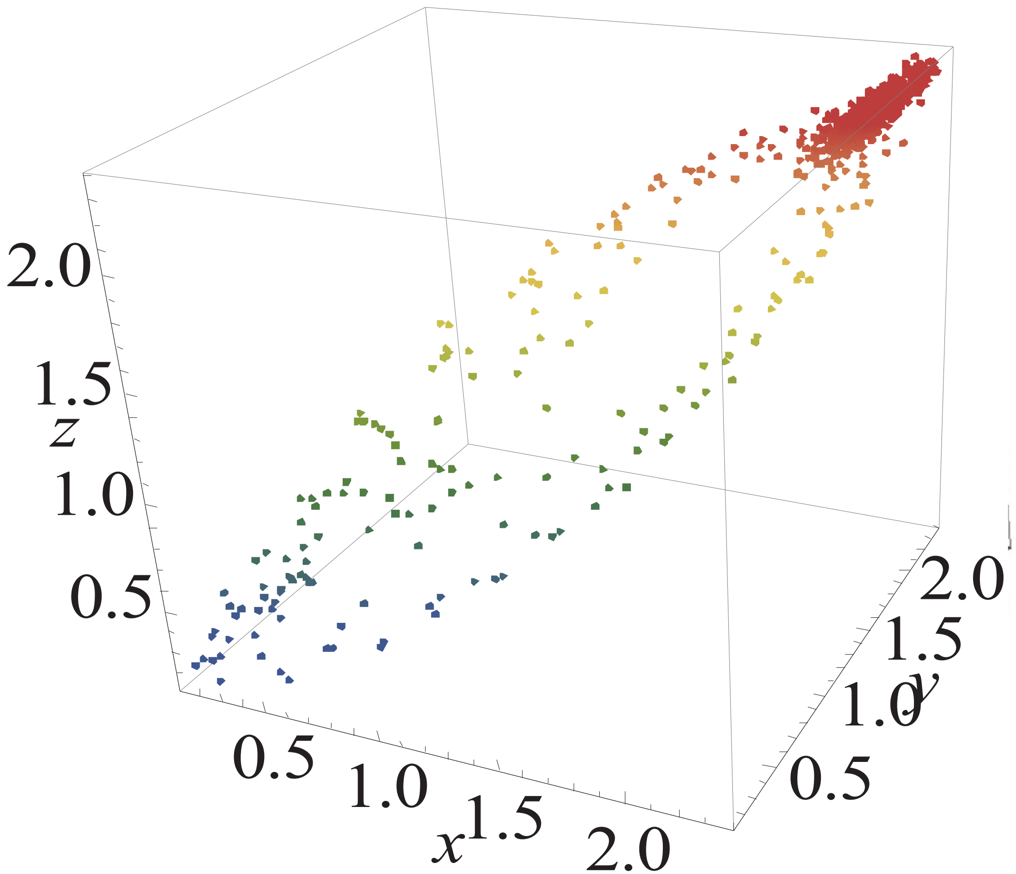


Velocity Norm Time Series (Orange)

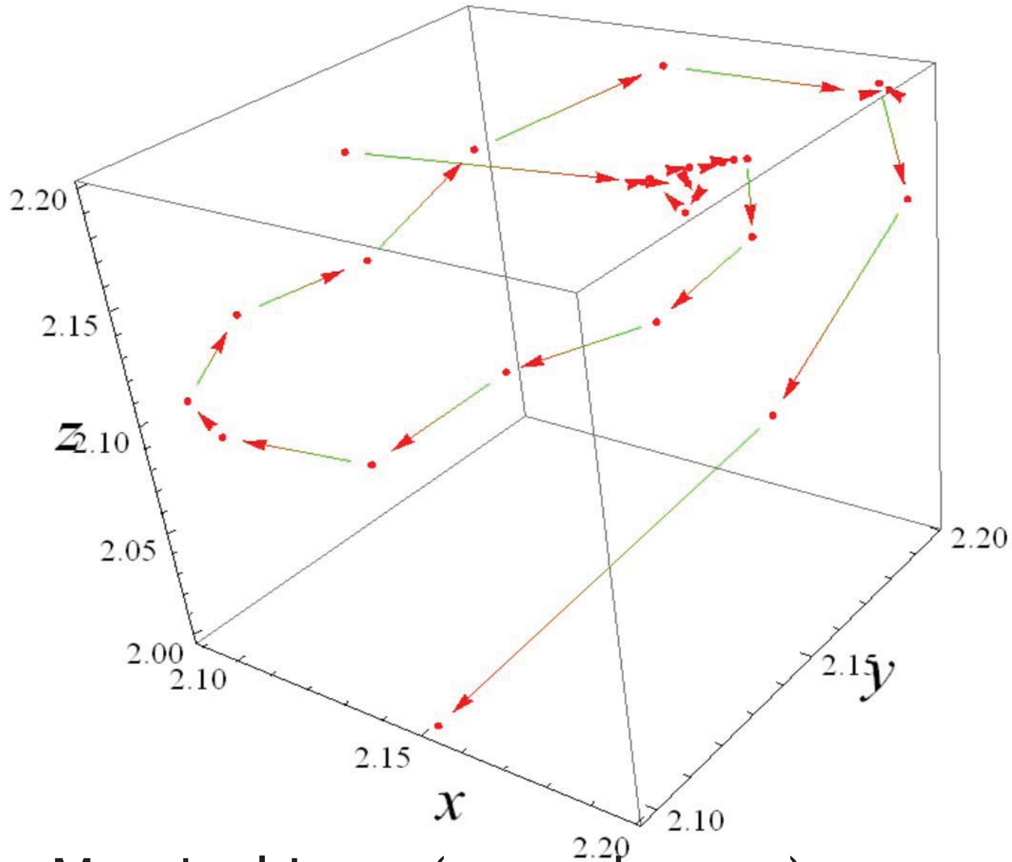
Pursuit Task

Visualization of Tertiary Structure

- MA/SLW: $l = 3$, $w=4$
- With temporal links



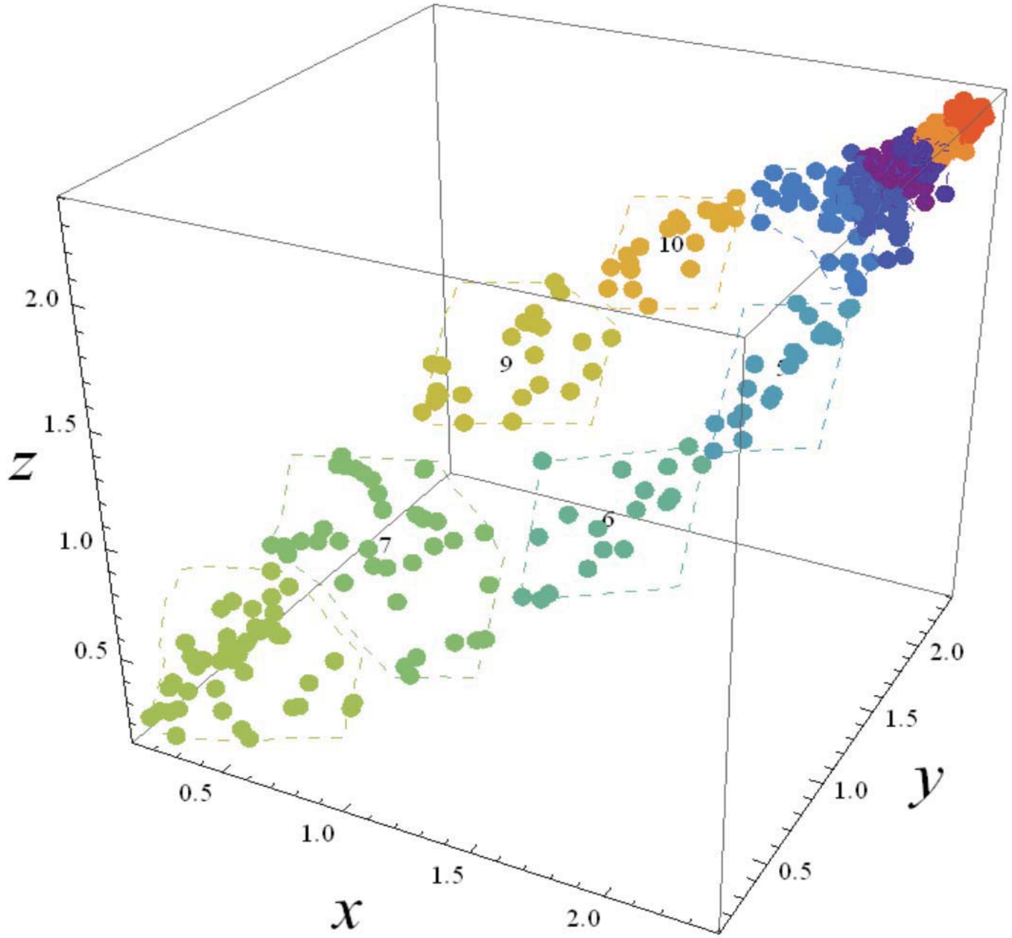
Primary Loop (along $x=y=z$)



Marginal Loop (around $x=y=z$)

Subsequence Clusters

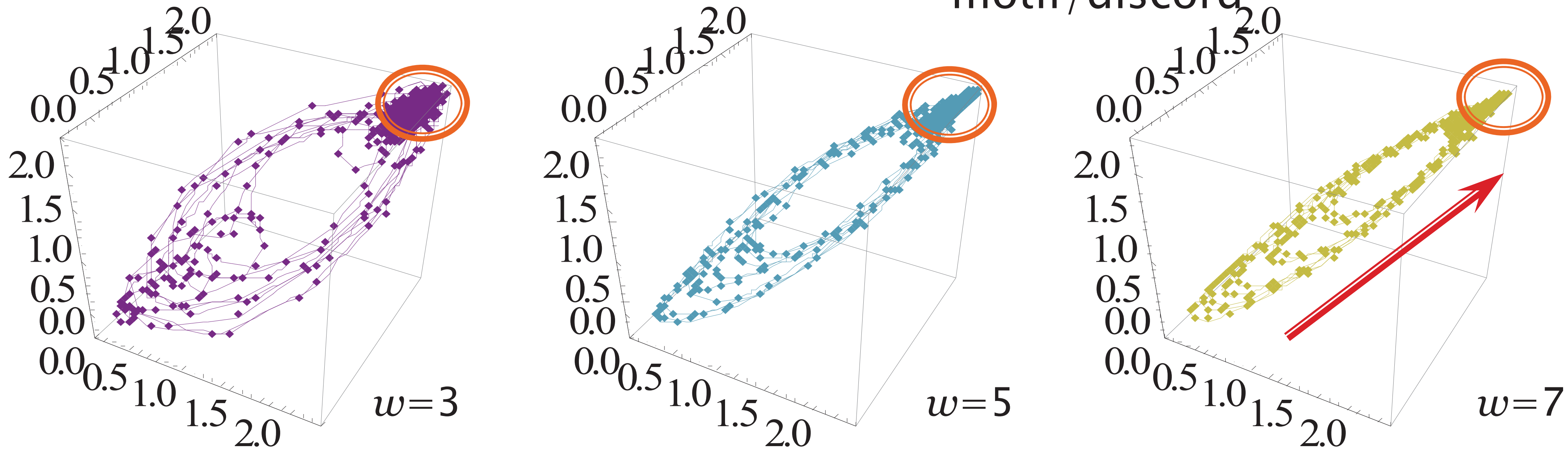
- k–Means (k=10)



- Cluster Means
  - Partial sine waves

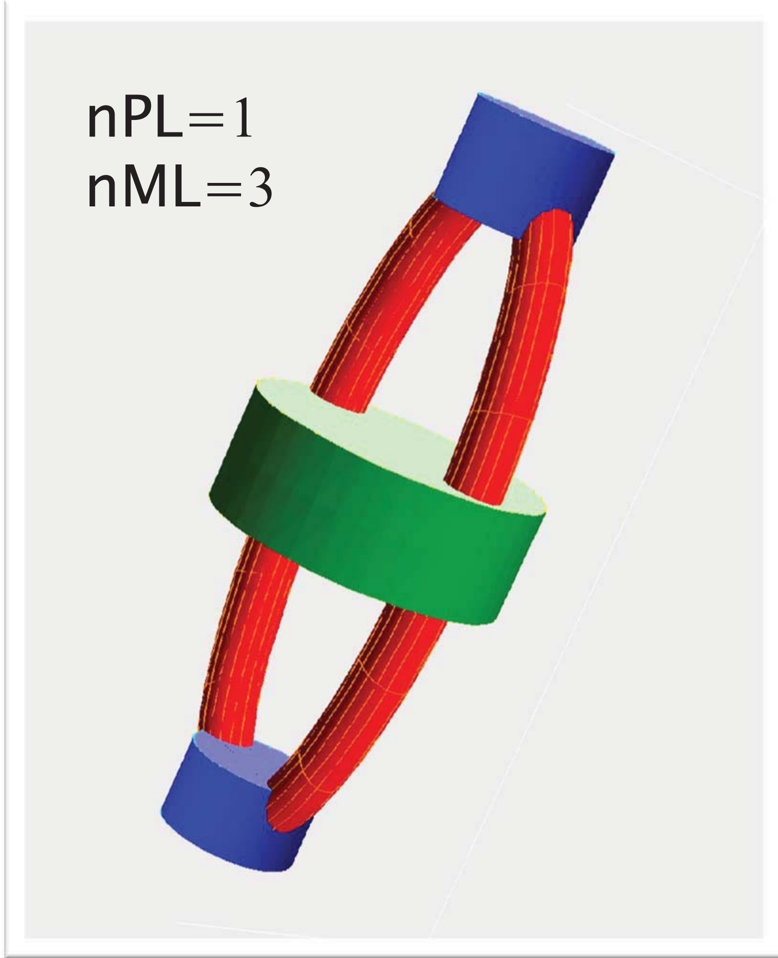
Effect of Parameter Choice

- Moving Average:
  - Reduce Marginal Loops
  - Scale Discrepancy
- Preferential criteria
  - w: remove high frequency noise
  - l : smaller than typical pattern/trend
    - “A pattern within motif/discord is also a motif/discord”



Extending Tertiary Analysis

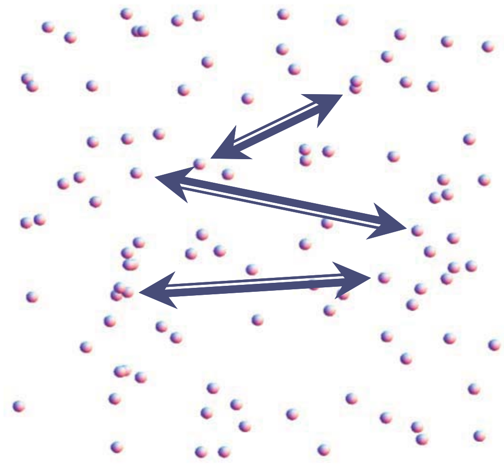
- Topology: low-rank components
  - Primary loop: large-scale trend/Dominant variance
- Density
  - Dependent on pattern frequency



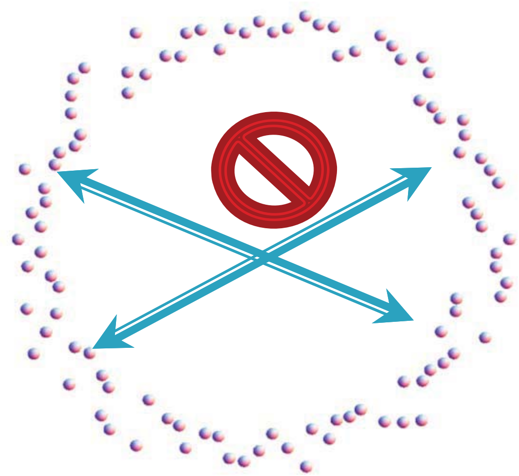
Conceptual Illustration of Subsequence Data Structure

Manifold Learning

- Manifold Learning
  - Manifold: neighborhood of each point can be represented as a Euclidean subspace
- Goal: find/parametrize $\{M : M \subset R^l\}$
- Techniques
  - $\times$Topological Theories
  - Principal Component Analysis
  - Spectral (sparse graph) representation

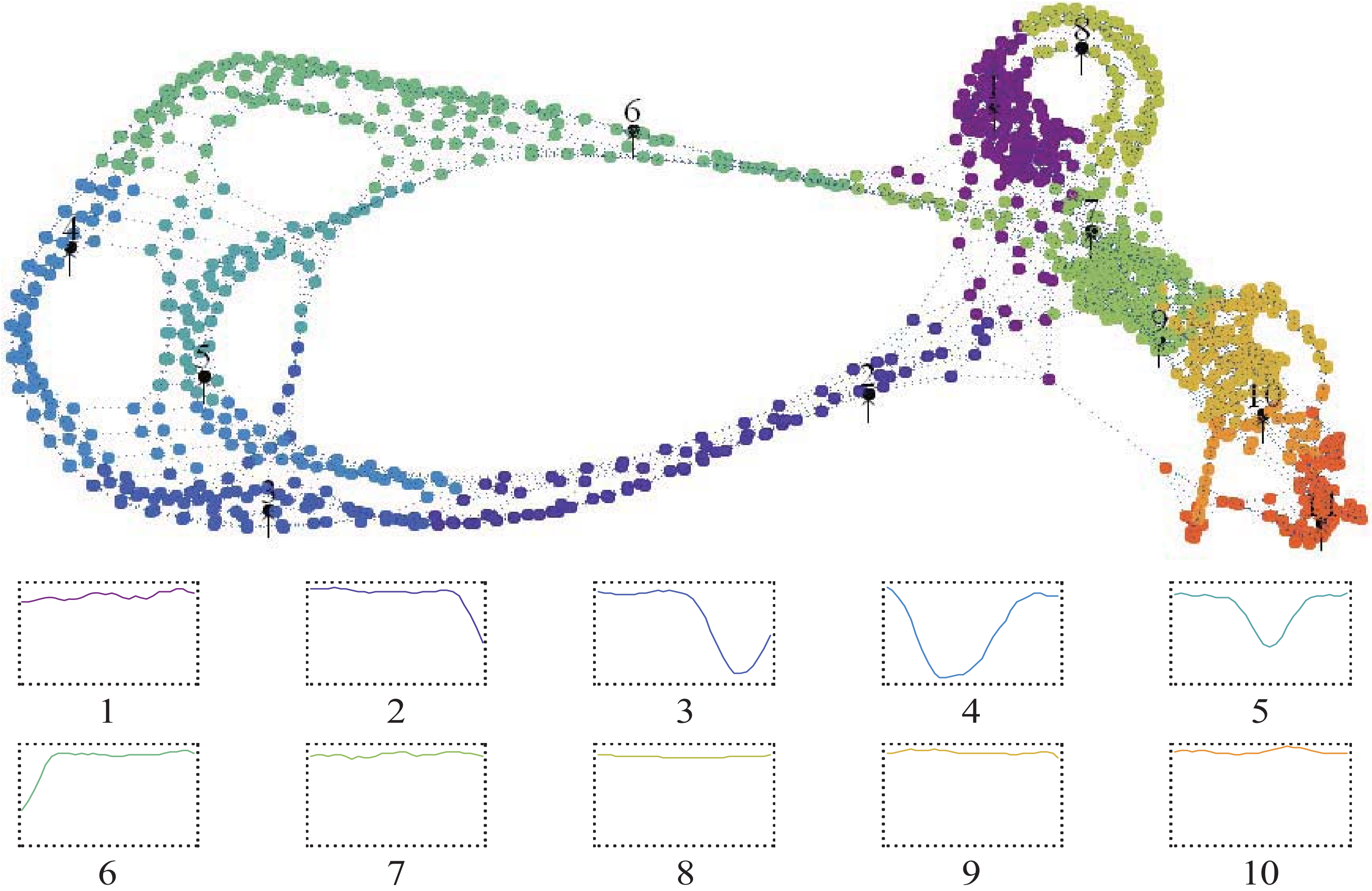


Metric space



Manifold

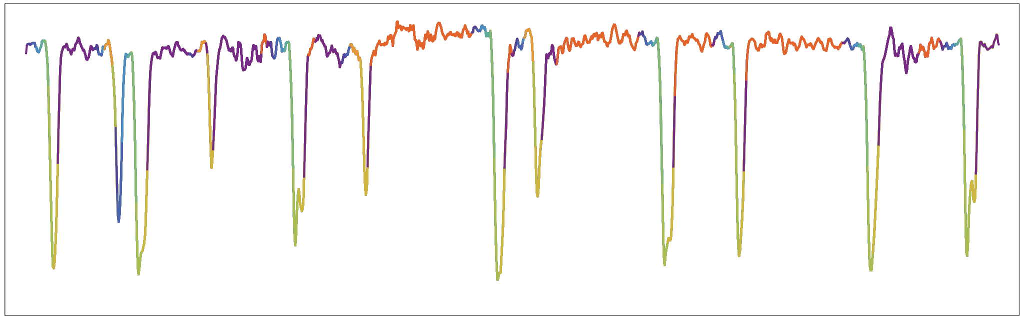
Nearest-neighbor Graph



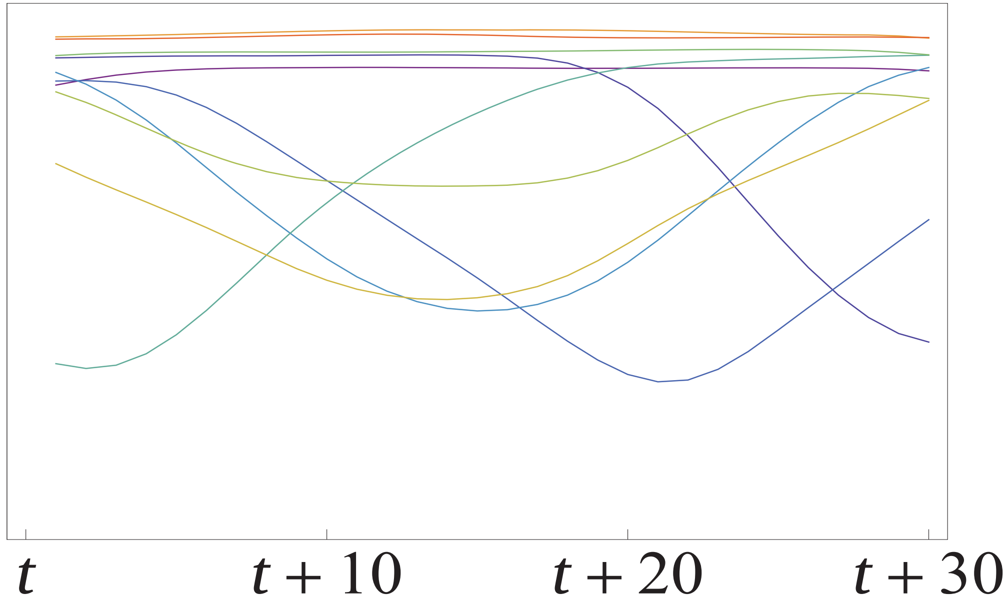
- w=4, l = 20, k=10
- n=4

- Graph rendering: spring-electric model

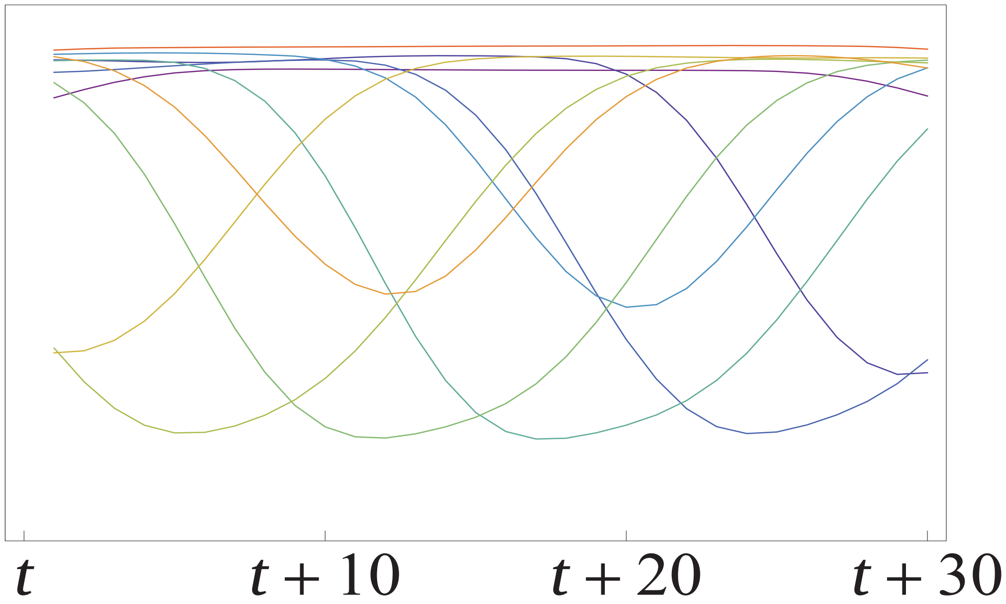
Cluster Means



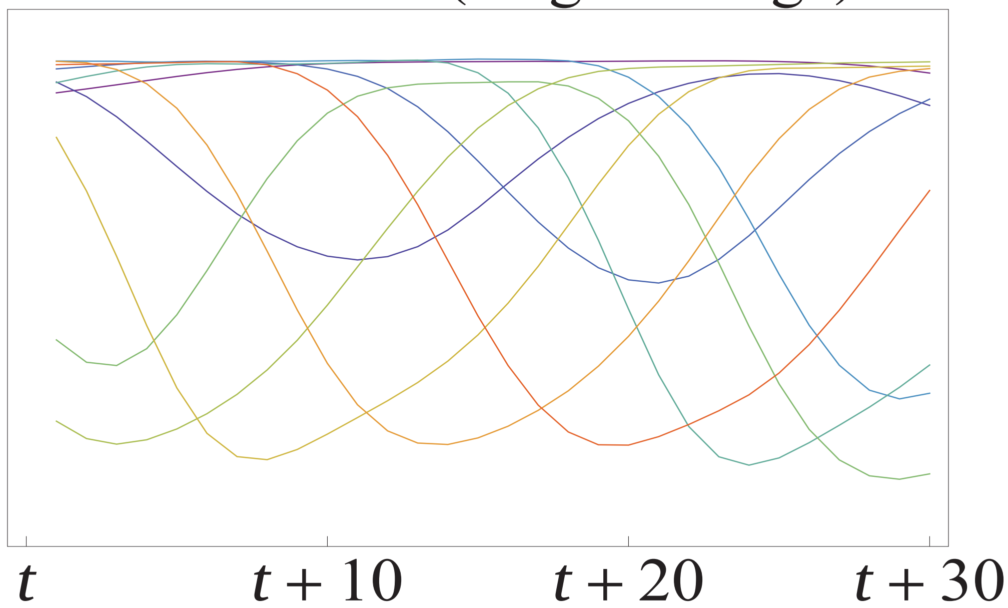
Normalized cut



k−means

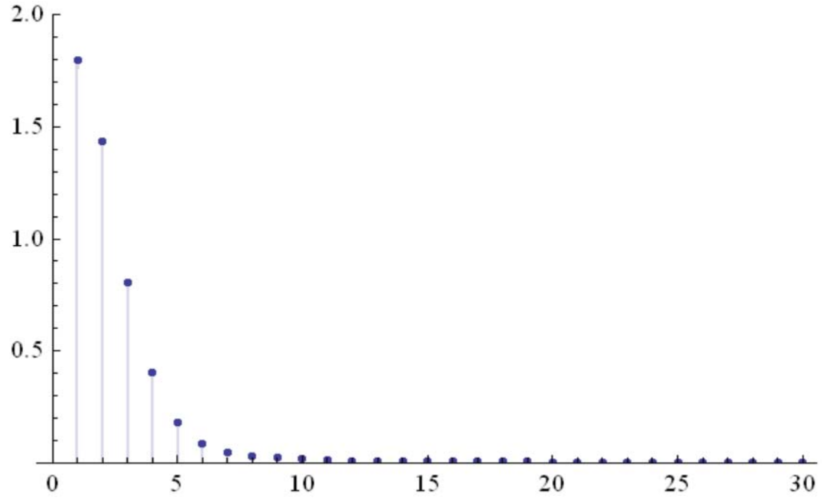


Hierarchical (Avg. Linkage)

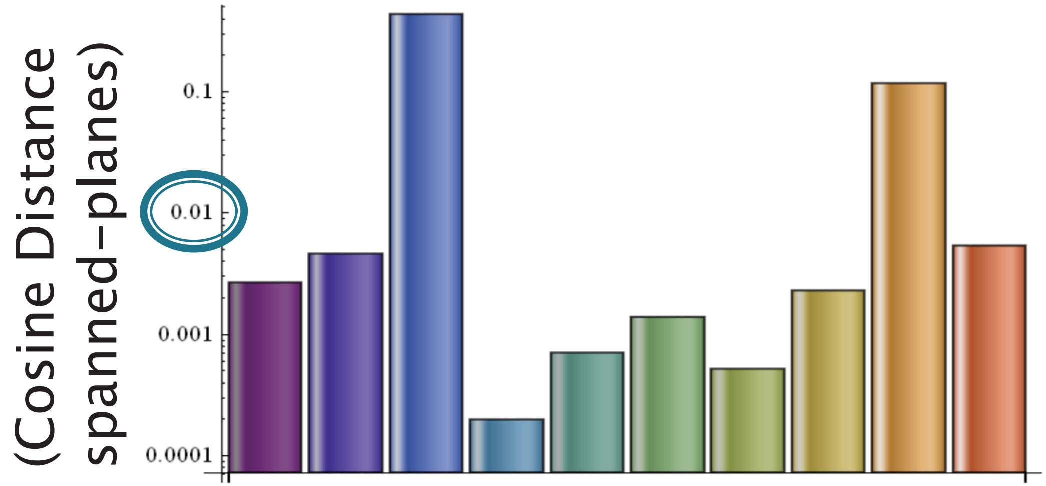


Principal Component Analysis

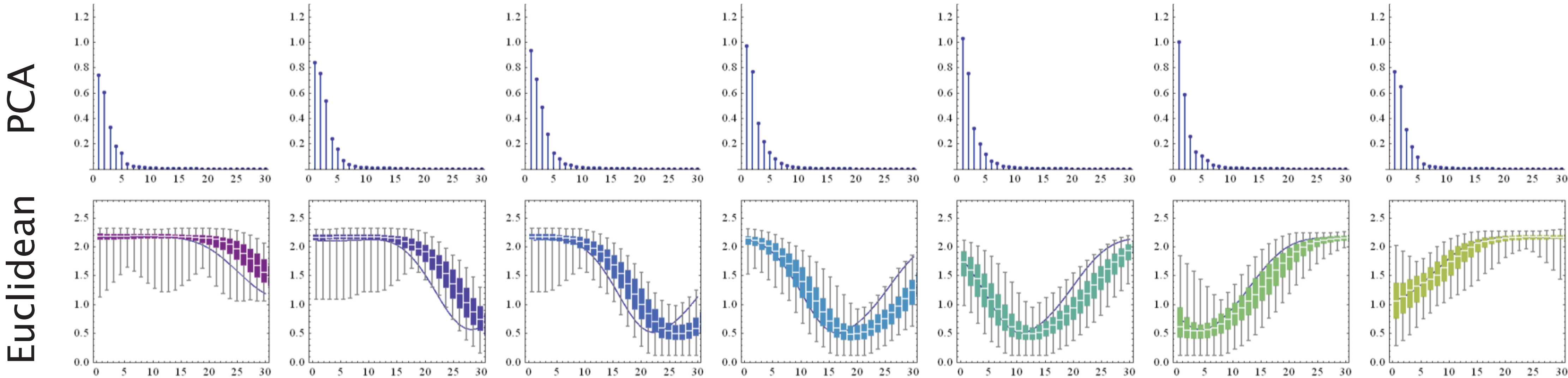
Global Principal Component Deviations



Component-wise similarity of Primary and Secondary Components



Cluster-wise Component Deviations



サブシーケンスマイニングの指針

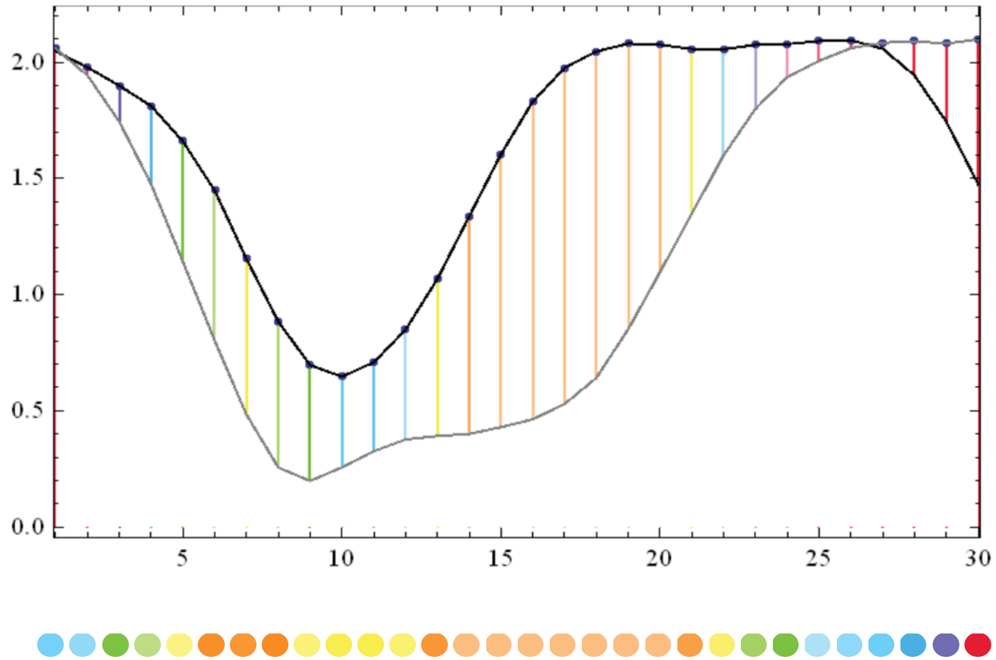
Deconstructing Subsequence Mining

- Reconstider motivation:
  - What can we find ?
  - What are the relevant patterns ?
- Methods and algorithms:
  - Designed for the topology of subsequence
  - Scale discrepancy
  - Skewed density
- Directions
  - Ignore some data (MDL)
  - Referential Approach
  - Topological technique

MDL-principled Approach [Rakthanmanon11]

- "Clustering Time Series Streams Requires Ignoring Some Data"
- Minimize Description Length L of Data D given Model M
  - $\text{argmin}_M L(M) + L(D|M)$
- Quantization:
  - $f_Q: X \in R^T \rightarrow Q \in \{q_1, \dots, q_k\}^T$
- Entropy: lower-bound of expected # of bits
  - $H(Q) = \Sigma_{i=1}^{k} p(q_i) \log p(q_i)$
- Mean Model: $L(\mathbf{v}|M;\mu)$
  - $H\big(f_Q(\mu)\big) + H\big(f_Q(\mathbf{v}-\mu)\big)$

- Example



$f_Q(\mathbf{v}-\mu) = (\textit{Blue}, \textit{Cyan}, \dots, \textit{Purple}, \textit{Violet})$

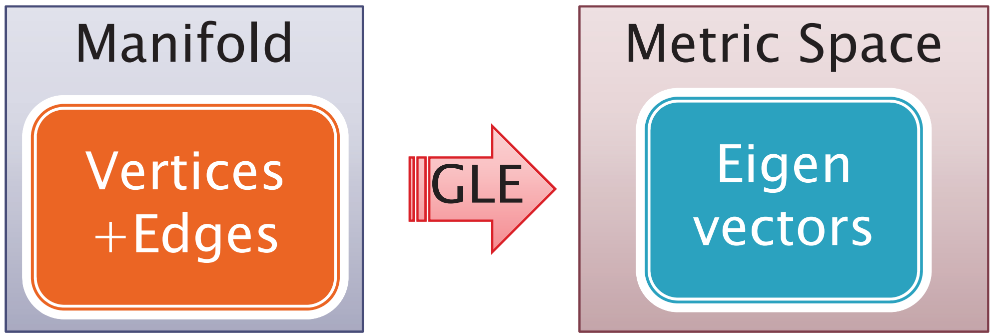
$$H(Q) = p(B|Q)\log p(B|Q) +, \dots, + p(V|Q)\log p(V|Q)$$

BitSave Algorithm

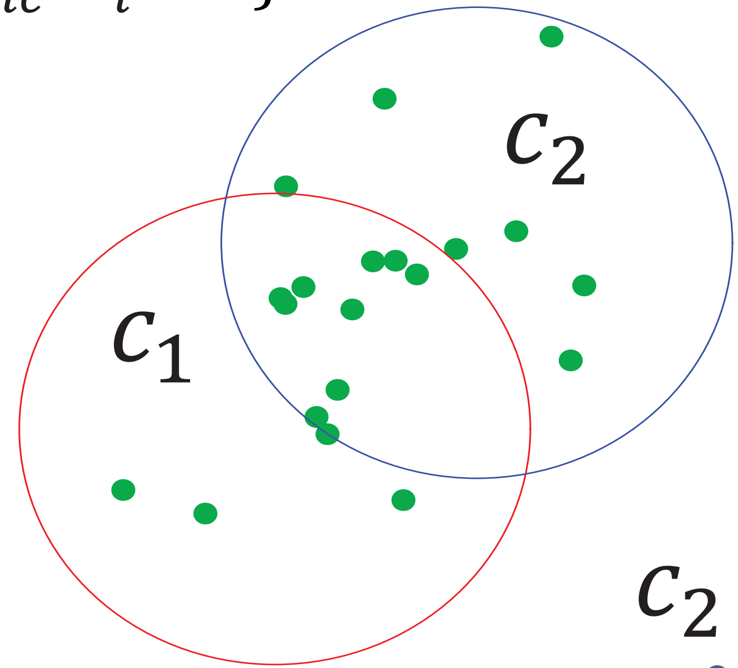
- Clustering Model DL:
  - $L(\{\mu_i\}) + L\left(\{\mathbf{v}_j\} | \{\mu_i\}, \{c_i\}\right)$
- Entropic Measure:
  - $\Sigma_i H\left(f_Q(\mu_i)\right) + \Sigma_i \left\{ H\left(f_Q(\mu_i)\right) + \Sigma_c H\left(f_Q\left(\mathbf{v}_j - \mu_i\right)\right) \right\}$
- Hierarchical Clustering – iterate joining clusters/instances
  - *BitSave*: (reduced entropy) after each iteration
  - Terminate when *BitSave* <0

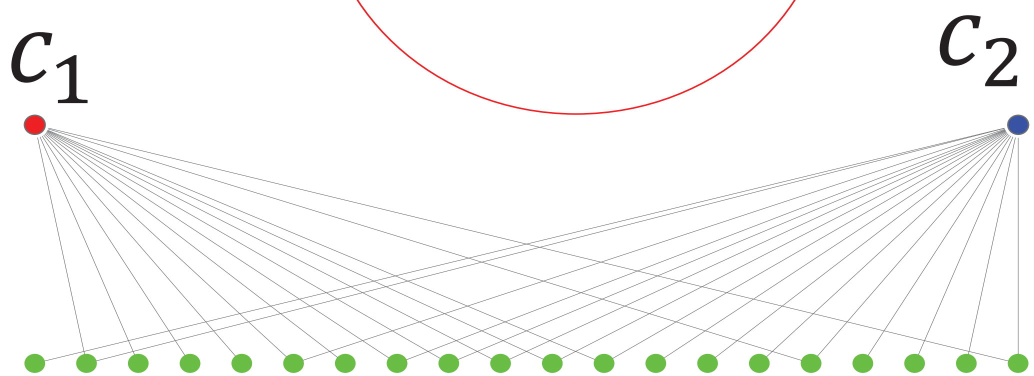
Topology-based Techniques

- Neighbor Graph: $G(V,E)$
  - $E = \{e_{ij} : \mathbf{v}_j \in N(\mathbf{v}_i)\}$
- Graph Partitioning:
  - $(V,E) \rightarrow \{V_k : V_k \cup V\}$
- Normalized-cut [Shi00]
  - GP=GLEP (Partitioning Graph Laplacian Eigenvectors)
  - GLE: $(V,E) \rightarrow \{\mathbf{e}_i\}_{i \in V}$



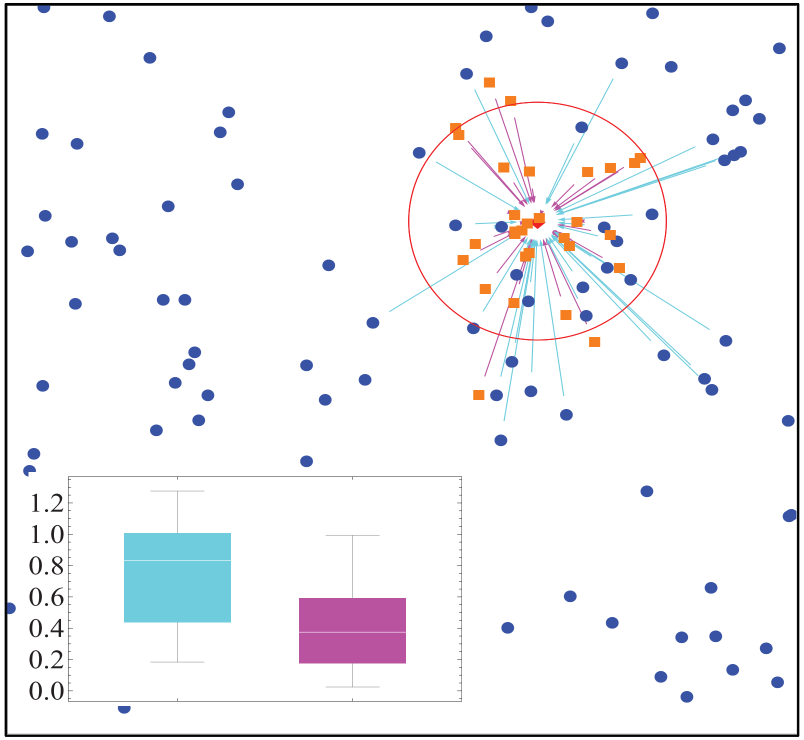
- Hybrid Graph Formulation [Fern04]
  - $\cup_c \{e_{ic} : \mathbf{v}_i \in c\}$





Referential Approach

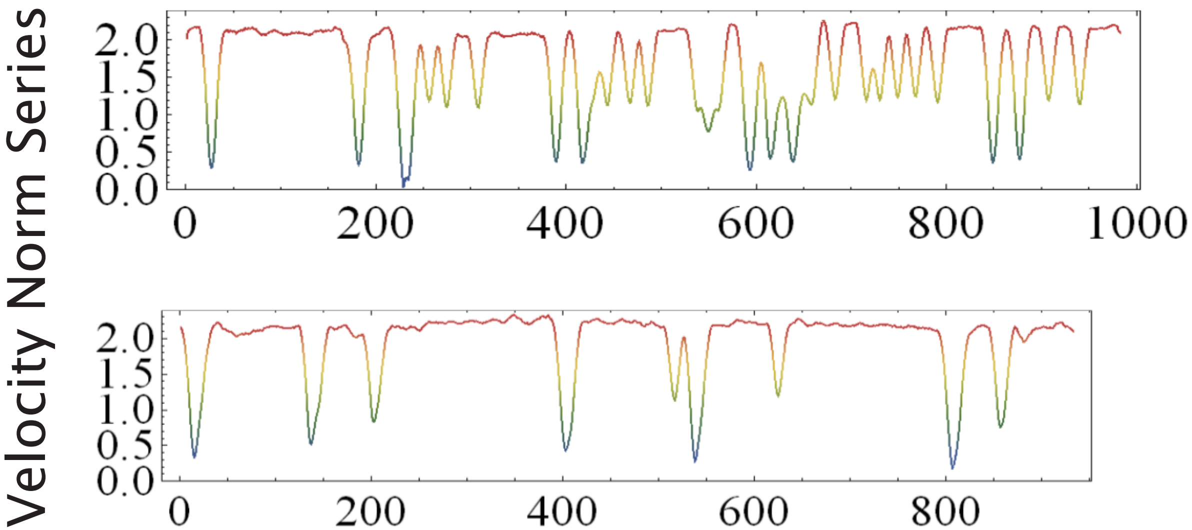
- Inlier-based Outlier Detection [Hido10]:
  - Inlier(Normal)/Test Density Ratio Estimation
- Anomalous Cluster Discovery [Ando11]: Target/Auxiliary (Normal) sets



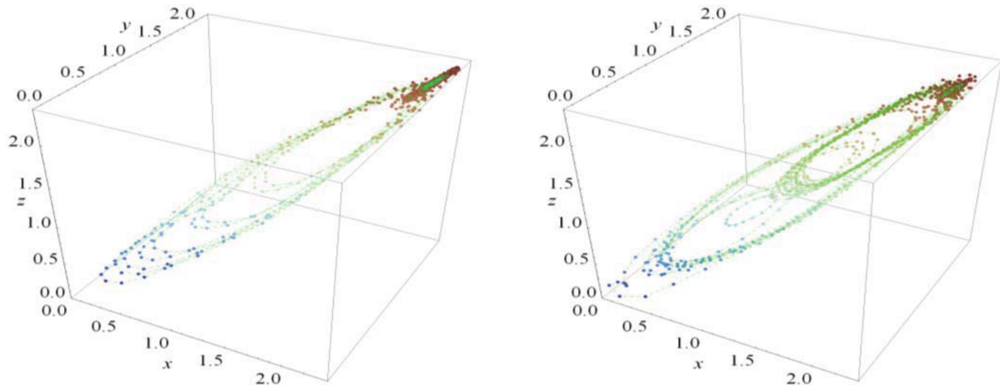
- Target data
- Auxiliary(Normal) data

Referential Dataset

- Reference Data
  - Orange: Exploring
- Target Data
  - Cyan: Pursuing



Tertiary Subsequence Structure

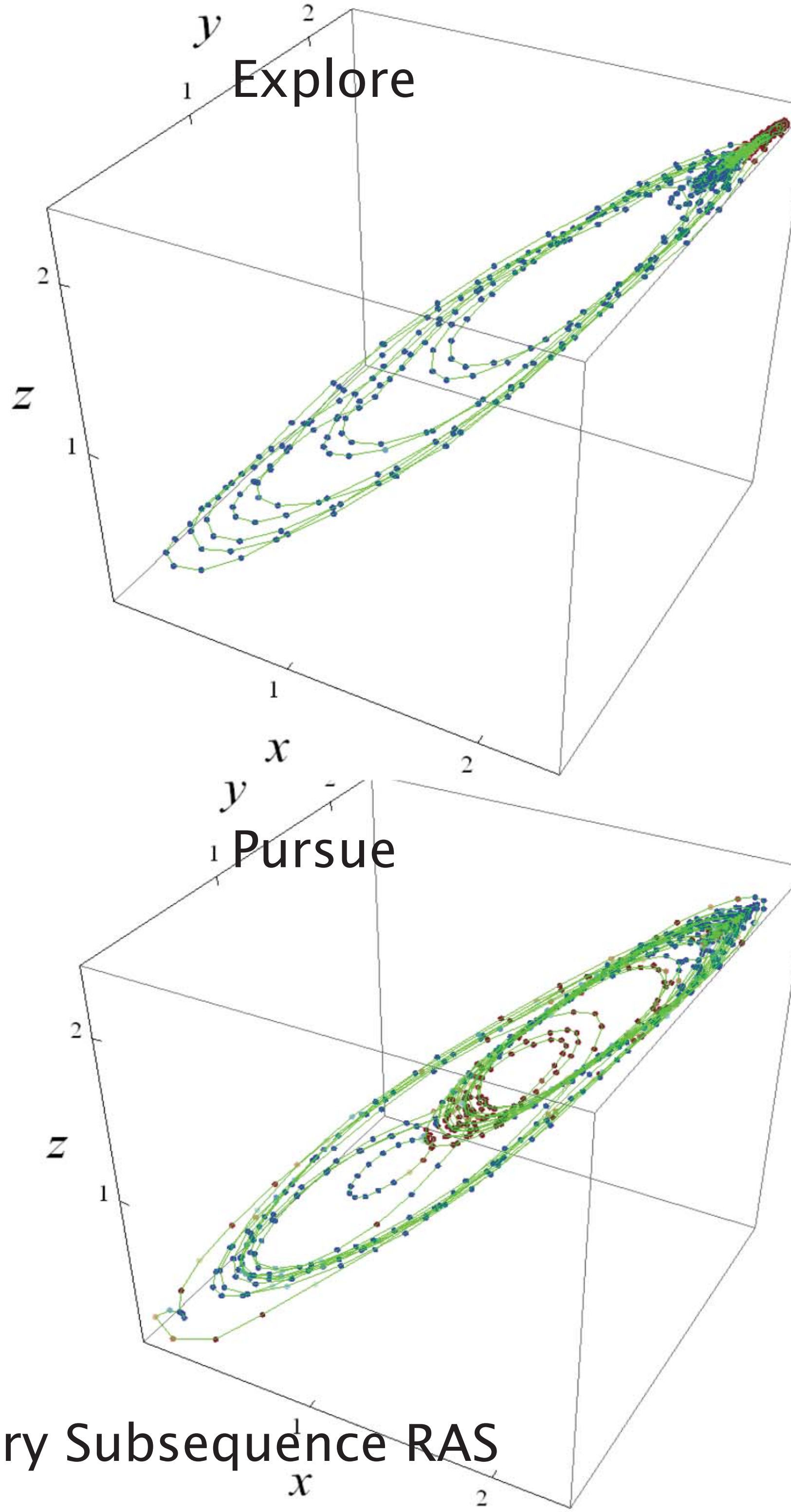


Reference

Target

Referential Density-based Approach

- Target dataset: $V^T = \{\mathbf{v}_i^T\}$
- Reference set: $V^R = \{\mathbf{v}_j^R\}$
- Referential neighbors:
  - $RN(\mathbf{v}, V^R) = \underset{S \subset V^R : \#S = k}{\operatorname{argmax}} \sum_{j \in S} D(\mathbf{v}, \mathbf{v}_j^R)$
- Referential Anomaly Score
  - $RDD(\mathbf{v}, V^R) = \max\left(\{D(\mathbf{v}, \mathbf{v}')\}_{\mathbf{v}' \in RN(\mathbf{v})}\right)$
  - $RRD^{-1}(\mathbf{v}, V^R) = H\langle \max\left(D(\mathbf{v}, \mathbf{v}'), RDD(\mathbf{v})\right)\rangle_{\mathbf{v}' \in RN(\mathbf{v})}$
  - $RAS(\mathbf{v}, V^R) = RRD^{-1}(\mathbf{v}) / \langle RRD^{-1}(\mathbf{v}')\rangle_{\mathbf{v}' \in RN(\mathbf{v}')}$

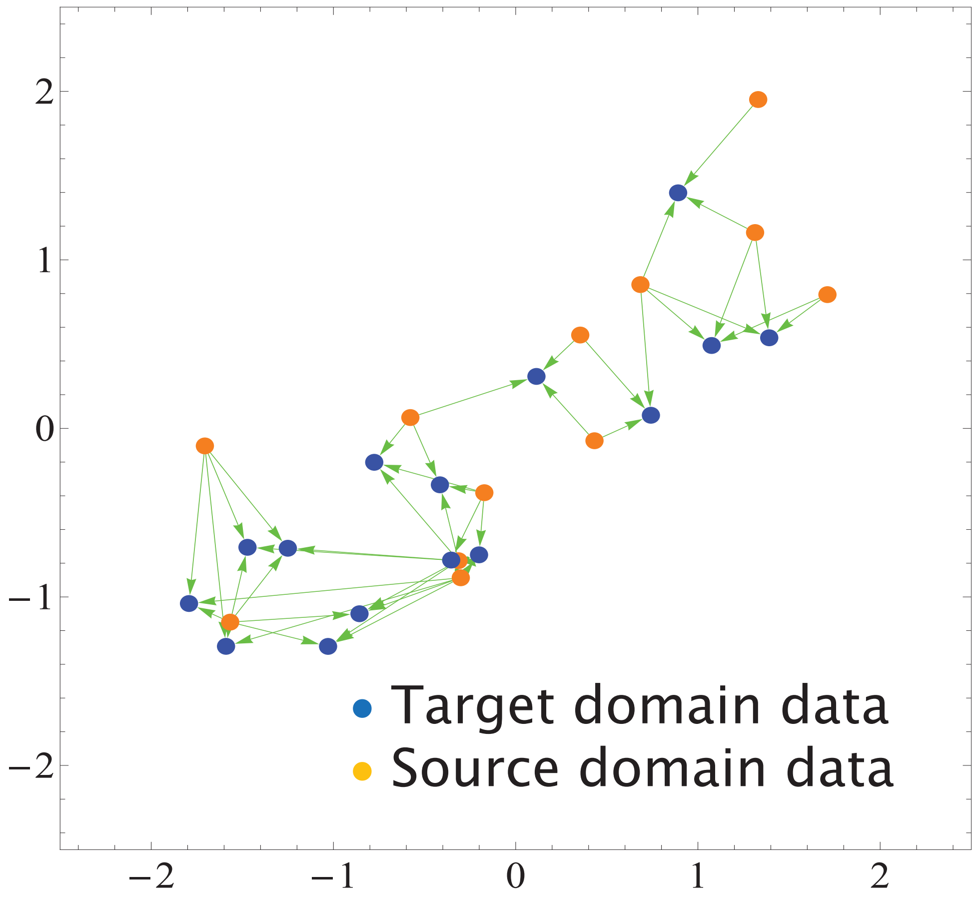


Tertiary Subsequence RAS

Referential Graph Formulation

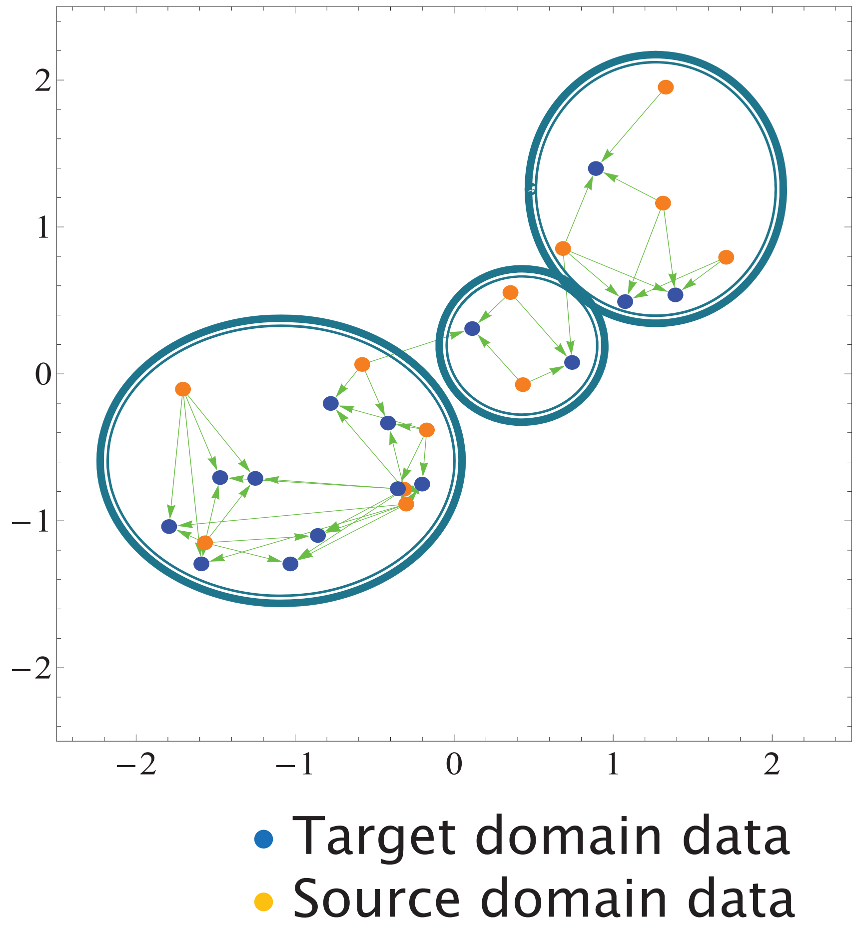
- Role-specific Behavior Analysis [Ando11]
  - Referential-neighbor links: $E_i = \{e_{ij} : \mathbf{v}_j^R \in N(\mathbf{v}_i^T)\}$
  - Edges: $E = \cup_i E_i$
  - Vertices: $V = V^T \cup V^R$
  - Graph: $G = (V, E)$
- G is a bipartite graph

- No links to transition subsequence



Reference Graph Partitioning

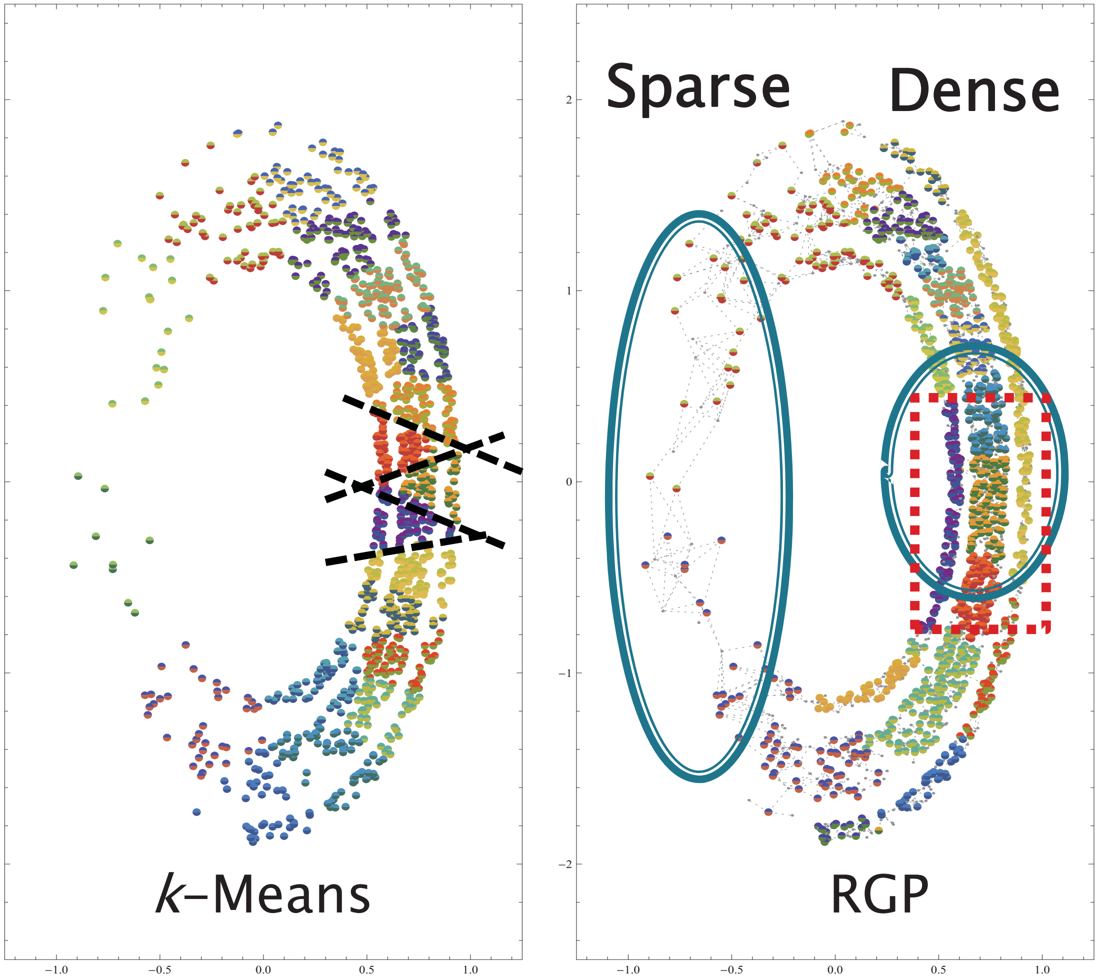
- Normalized-cut
  - $GLE(V,E) \rightarrow \{\mathbf{e}_h\}_{\mathbf{v}_h \in V} = \{\mathbf{e}_i\}_{\mathbf{v}_i \in V^T} \cup \{\mathbf{e}_j\}_{\mathbf{v}_j \in V^R}$
- Cluster-pairing
  - Subgraphs: $\{V_k : V = \cup_k V_k\}$
  - $V_k = c^T \cup c^R$



RGP: Property

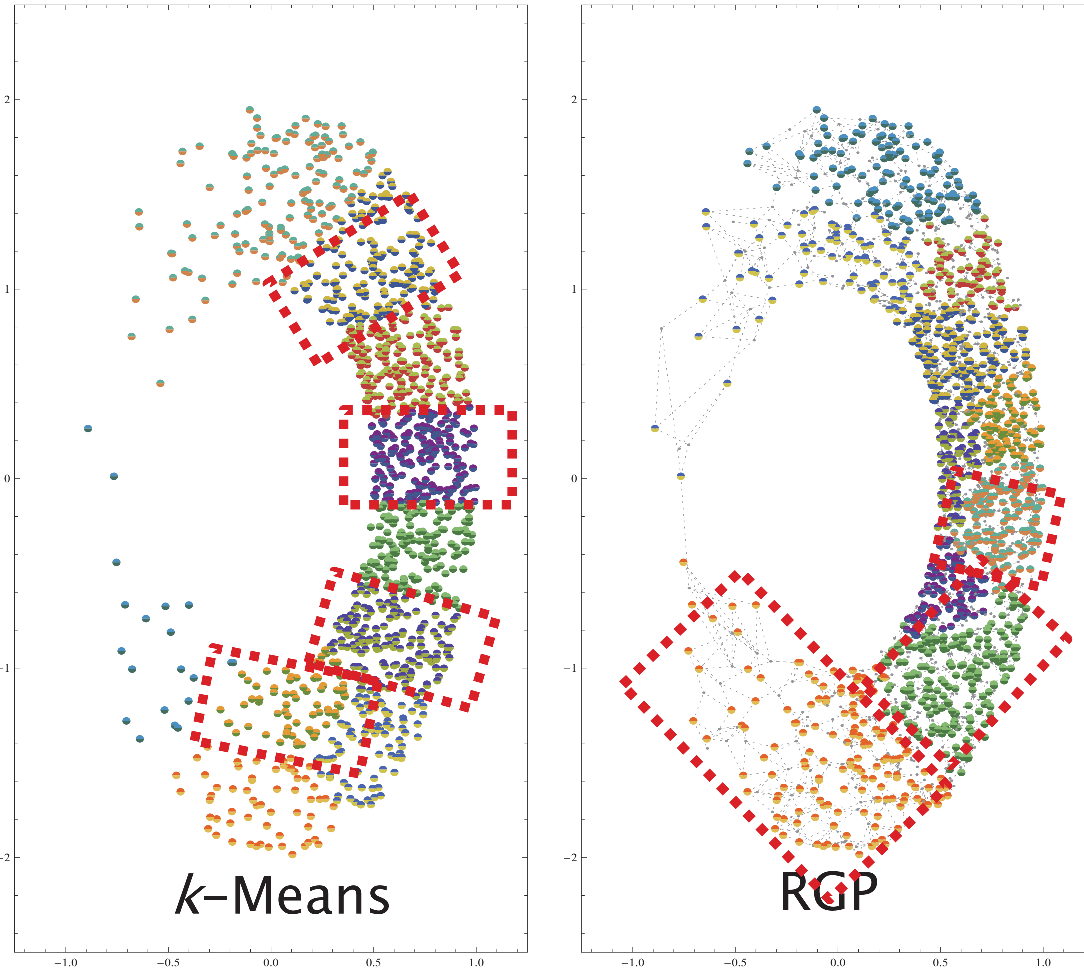
- Against discrepant scale

Skewed Saturn Ring (Cantor Set)



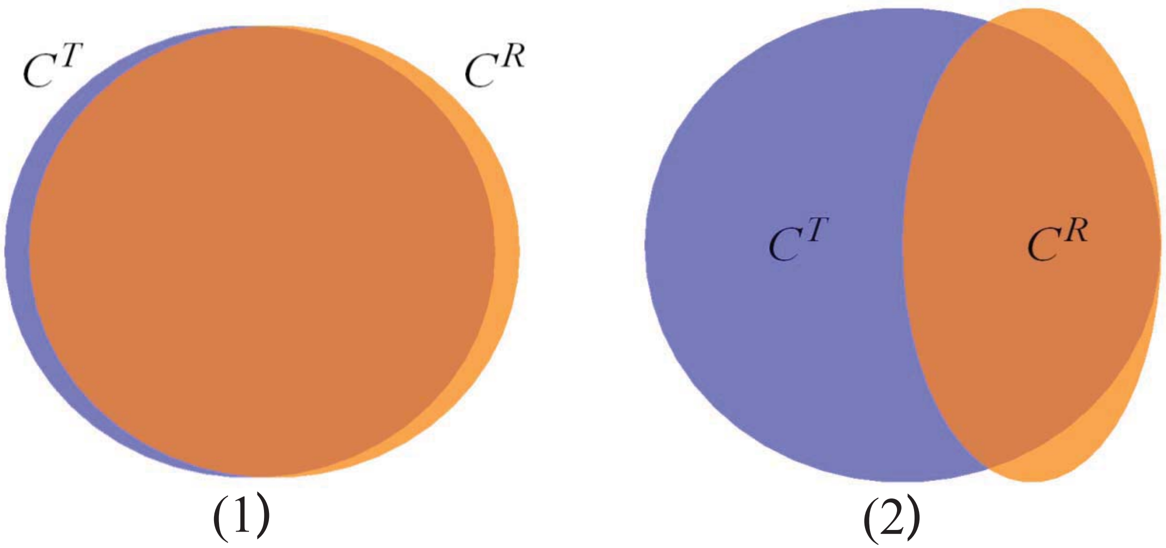
- Against skewed density

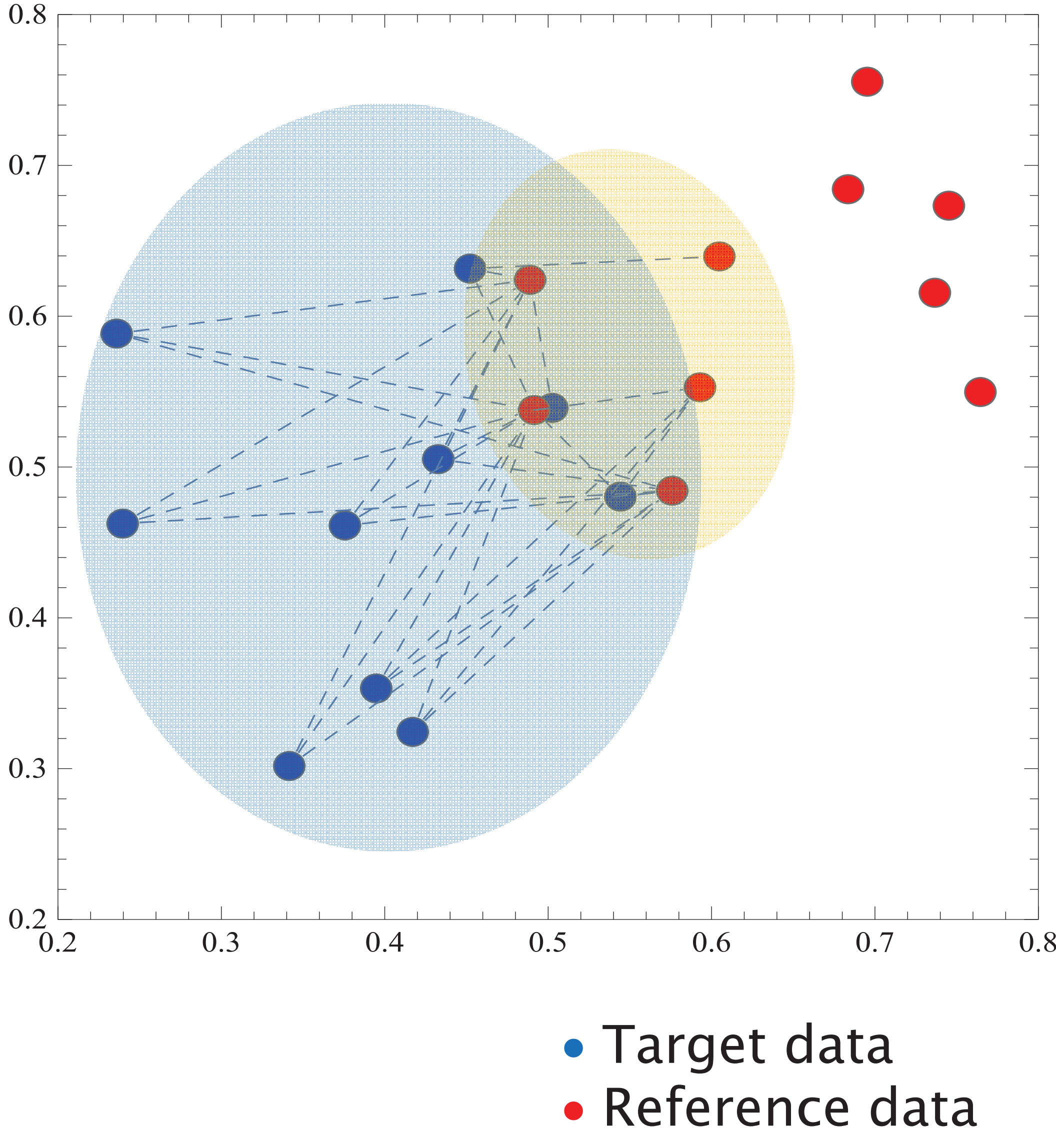
Skewed Ring



RGP: Property (Cluster Pair)

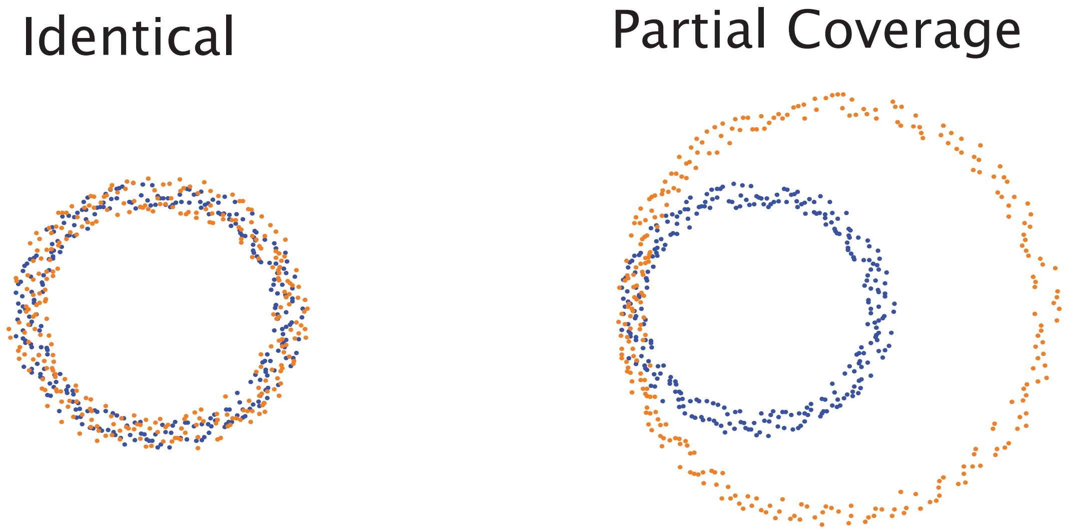
- c^{T} vs. c^{R}
  1. Overlap
  2. Partial Coverage

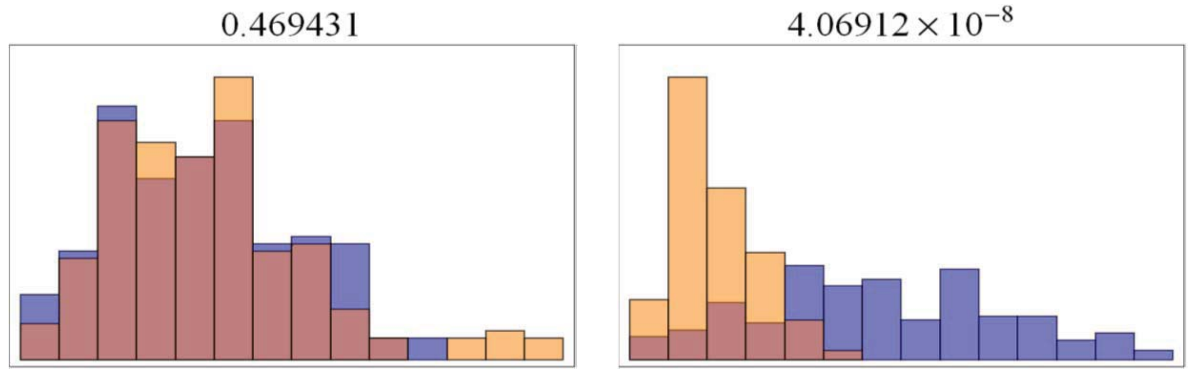




Detecting Partial Coverage

- Compare w/i cluster approximate densities
  - Target: $\bar{r}^T = \langle \{RRD(\mathbf{v}^T)\}_{c_k^T} \rangle$
  - Referential: $\bar{r}^R = \langle \{RRD(\mathbf{v}^R)\}_{c_k^R} \rangle$
- Test of Means
  - 1) Null Hypothesis: $\bar{r}^T = \bar{r}^R$
  - 2) Alternate: $\bar{r}^T \neq \bar{r}^R$
- Mann–Whitney test

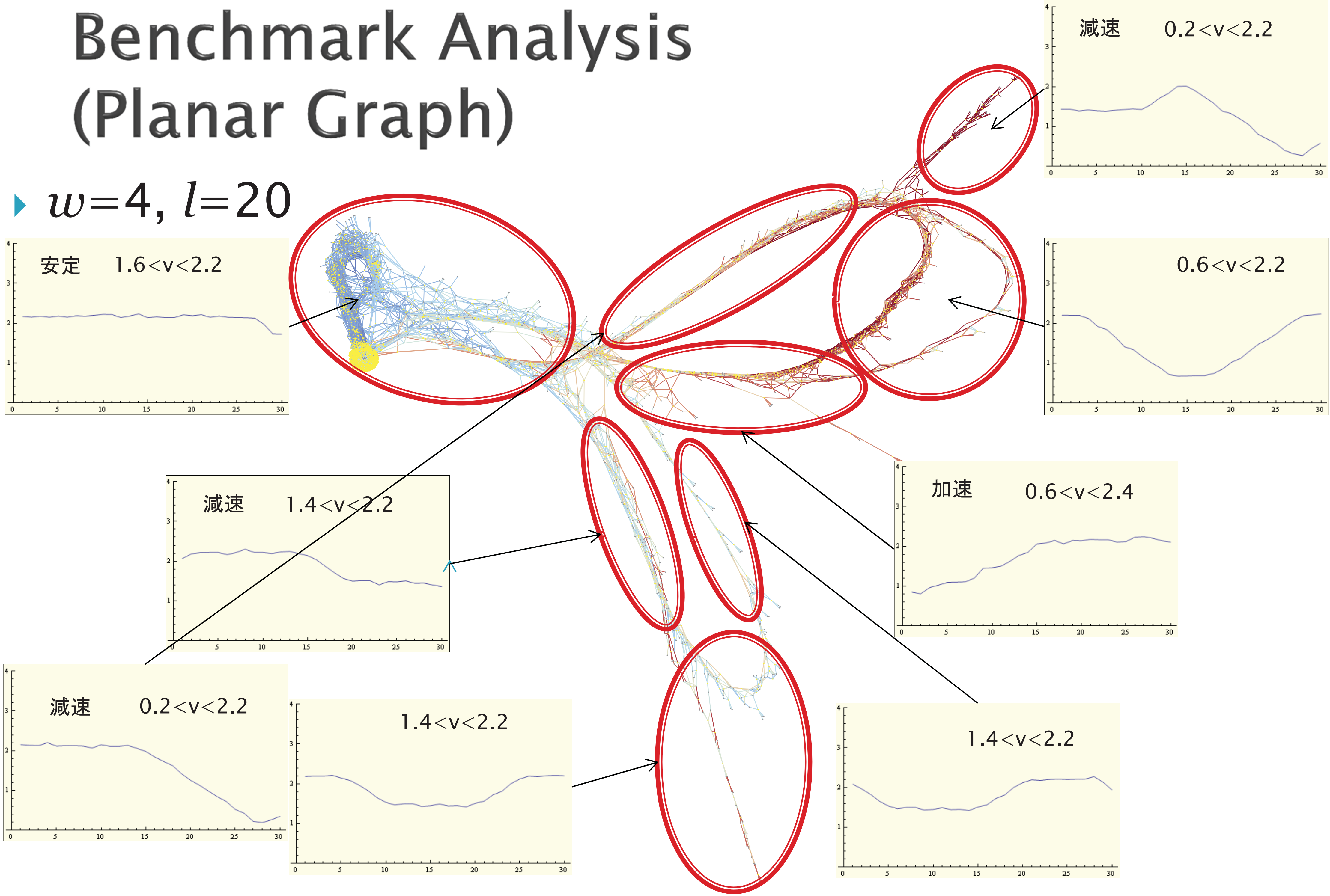




Additional distance evaluation required

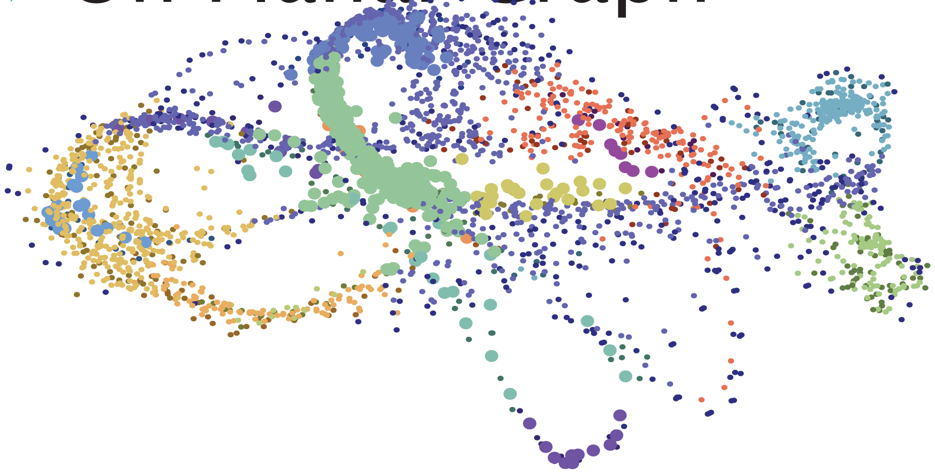
Benchmark Analysis (Planar Graph)

- $w=4$, $l=20$

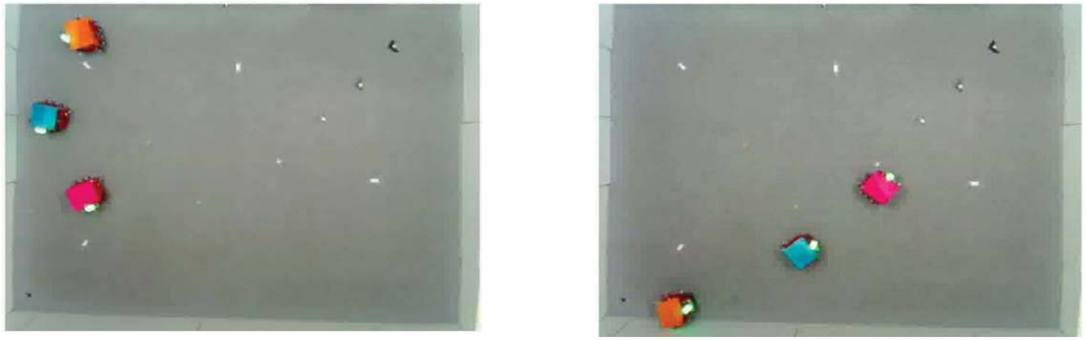


Context-specific Clusters

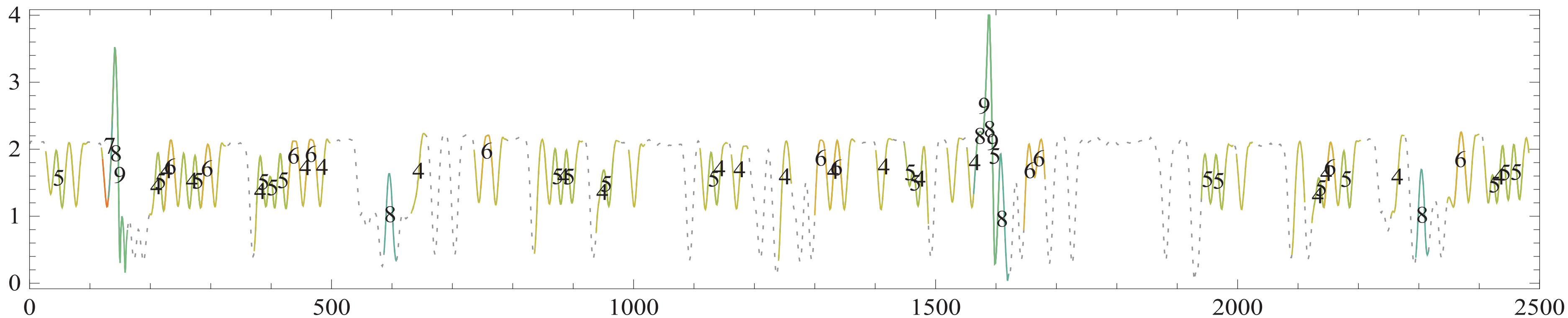
- On Planar Graph



- High RAS Instances



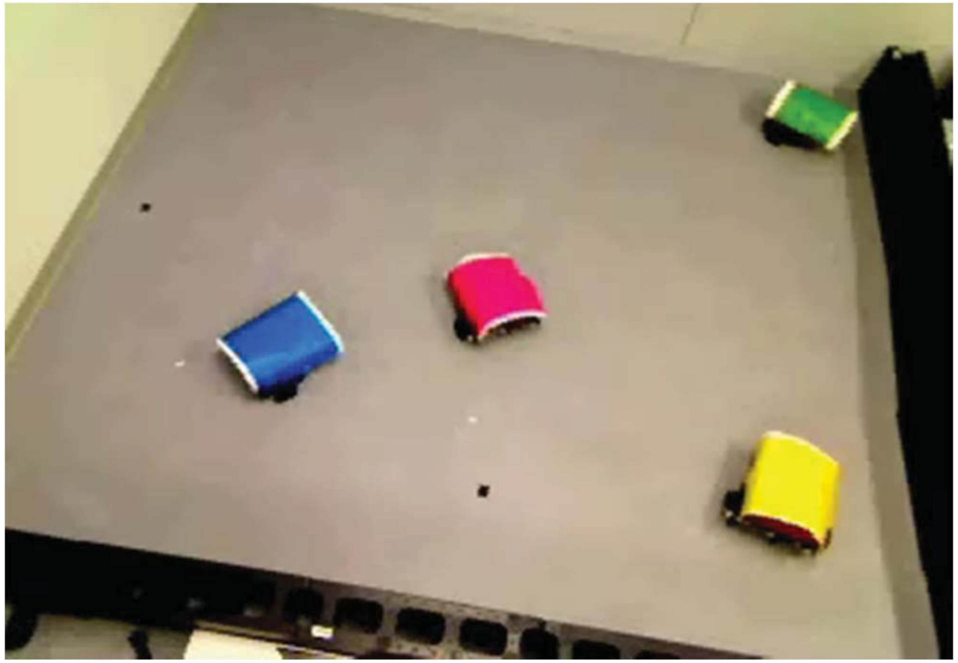
- On Time-series



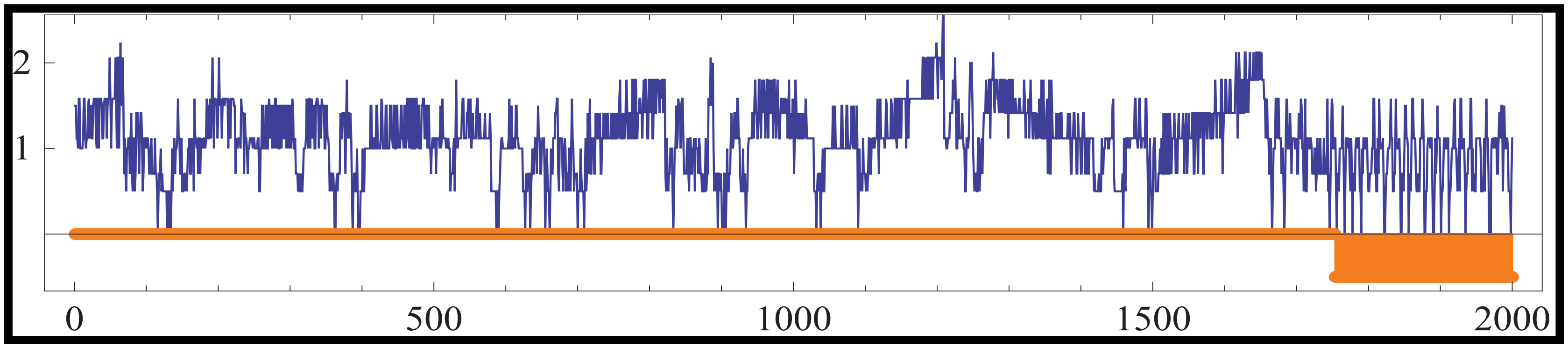
Anomaly Detection Task

- Exploring Agents

Experiment Video Log

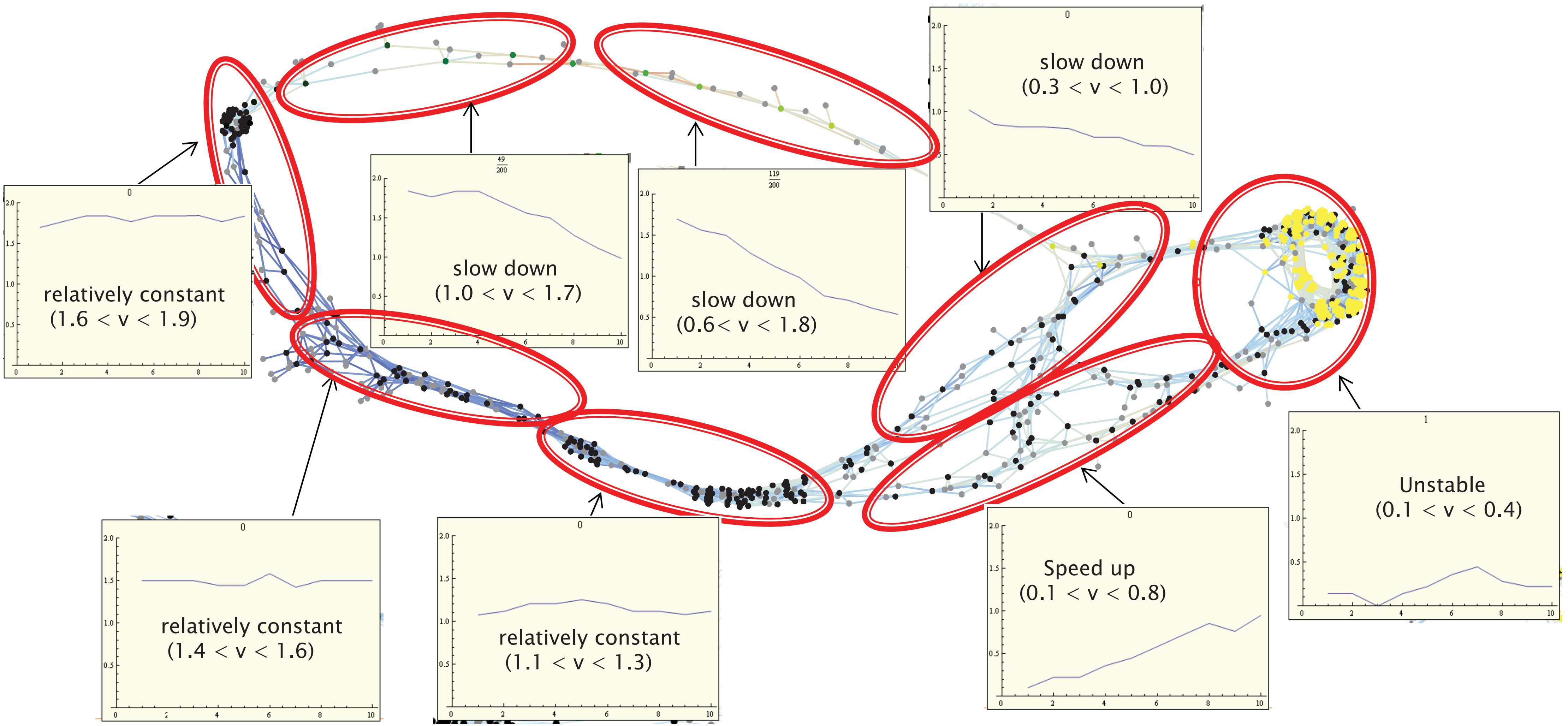


Velocity Norm
Time Series



- Setup
- w=3, l=7
- 3 trials: T=4122,4677,5380
- Loss of purpose: pivoting against field corner

Planar Graph Visualization

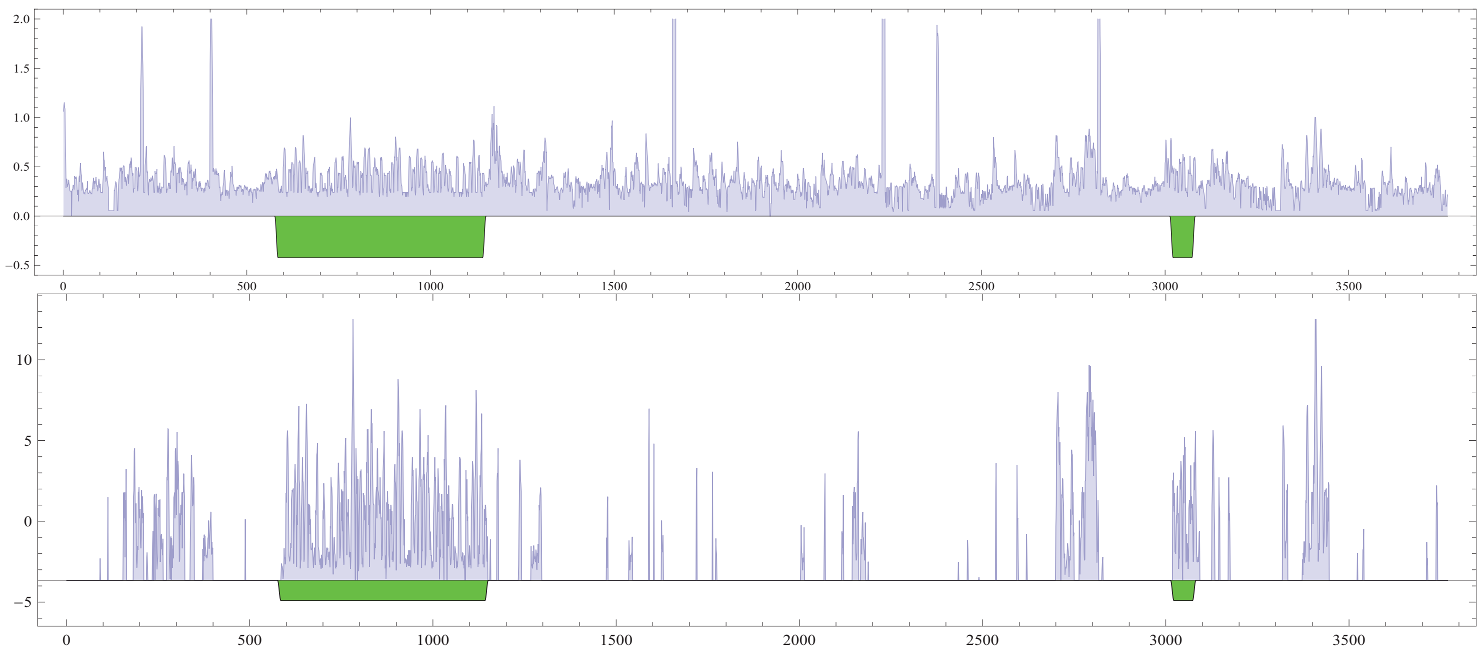


Performance Evaluation

- Baseline Methods
  - Inlier-based DB
  - LOF[Breunig]
  - Discord Discovery [Yankov]

Summary of AUC

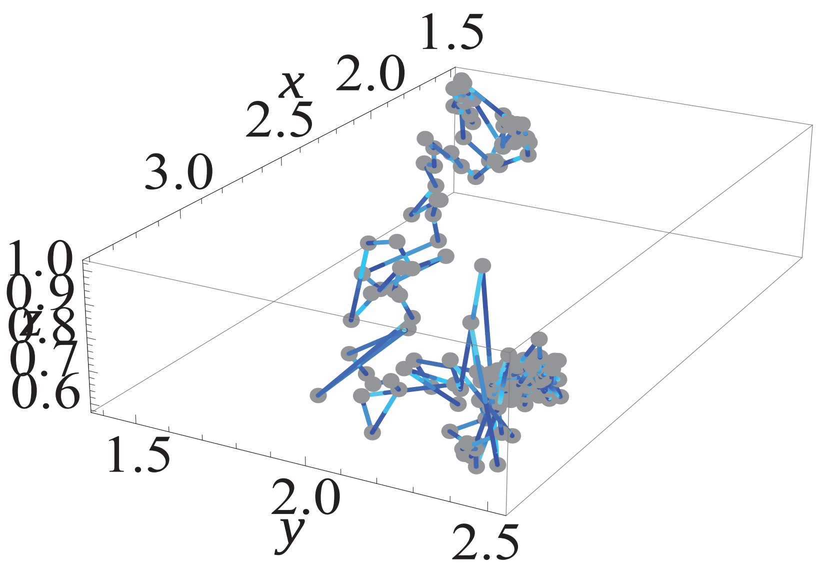
	T1	T2	T3
IbDB	0.79	0.62	0.45
LOF	0.44	0.38	0.31
DD	0.47	0.44	0.41
C/RAS	**0.80**	**0.90**	**0.79**



Anomaly Score Time Series

Semi-supervised Data Cleaning

- Person Activity Data
  - 3D coord. from sensors on ankles, belt chest
  - Perform a sequence of activities: walking, sitting, falling, etc.



Trajectory (Belt) during Walking→Falling

- k-NN Classification
  - Training data: velocity norm time series segmented by activities
  - Test data: subsequences of entire series (different trial)
- Cleaning:
  - RGF: Training set-Reference
  - Reject anomalous cluster

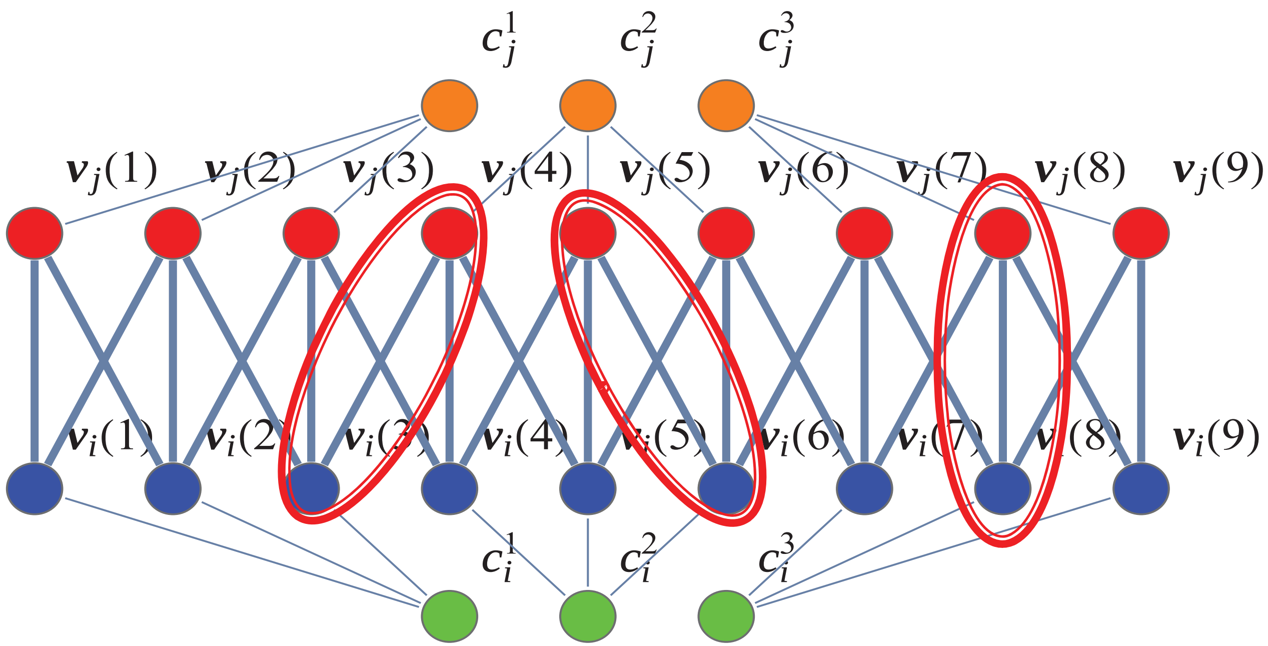
k=3	k-NN	k-NN + Cleaning
ACC	81.1	79.6
ERR	18.9	4.6
RJT	0.0	15.8

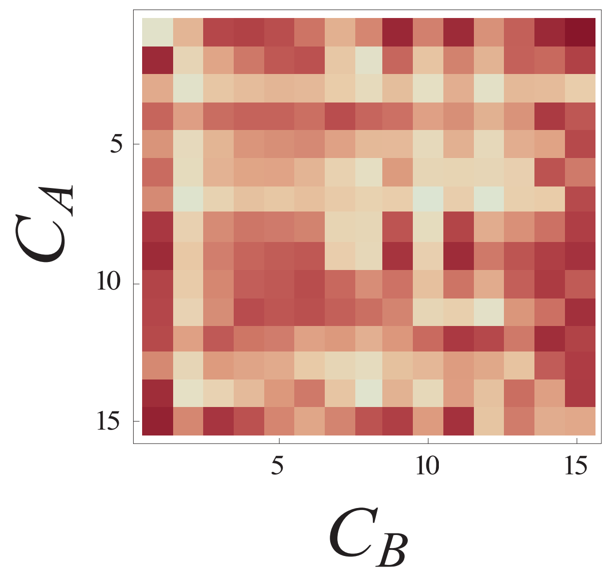
Multivariate Subsequence Analysis

- Multivariate time series:
  - $\mathbf{X} = \{X_1, \dots, X_p\}$,
  - $X_i = \{x_i(t)\}_{t=1}^{T}$
  - $\mathbf{v}_i(t) = (x_i(t), \dots, x_i(t + l - 1))$
- Symbolic Discovery
  - Temporal correlation of patterns (clusters) between time series
- Variable-wise Referential Clusters
  - $\mathbf{C} = \{C^i\}_{i=1}^{p}, C^i = \{c_k^i\}$
- Related problem:
  - Classification/anomaly detection based on correlation among multiple time series [Papadimitriou06,Ide07]
- Caveats
  - Irrelevant patterns in each time series
  - Relation to context-specific patterns

Multivariate Cluster Graph

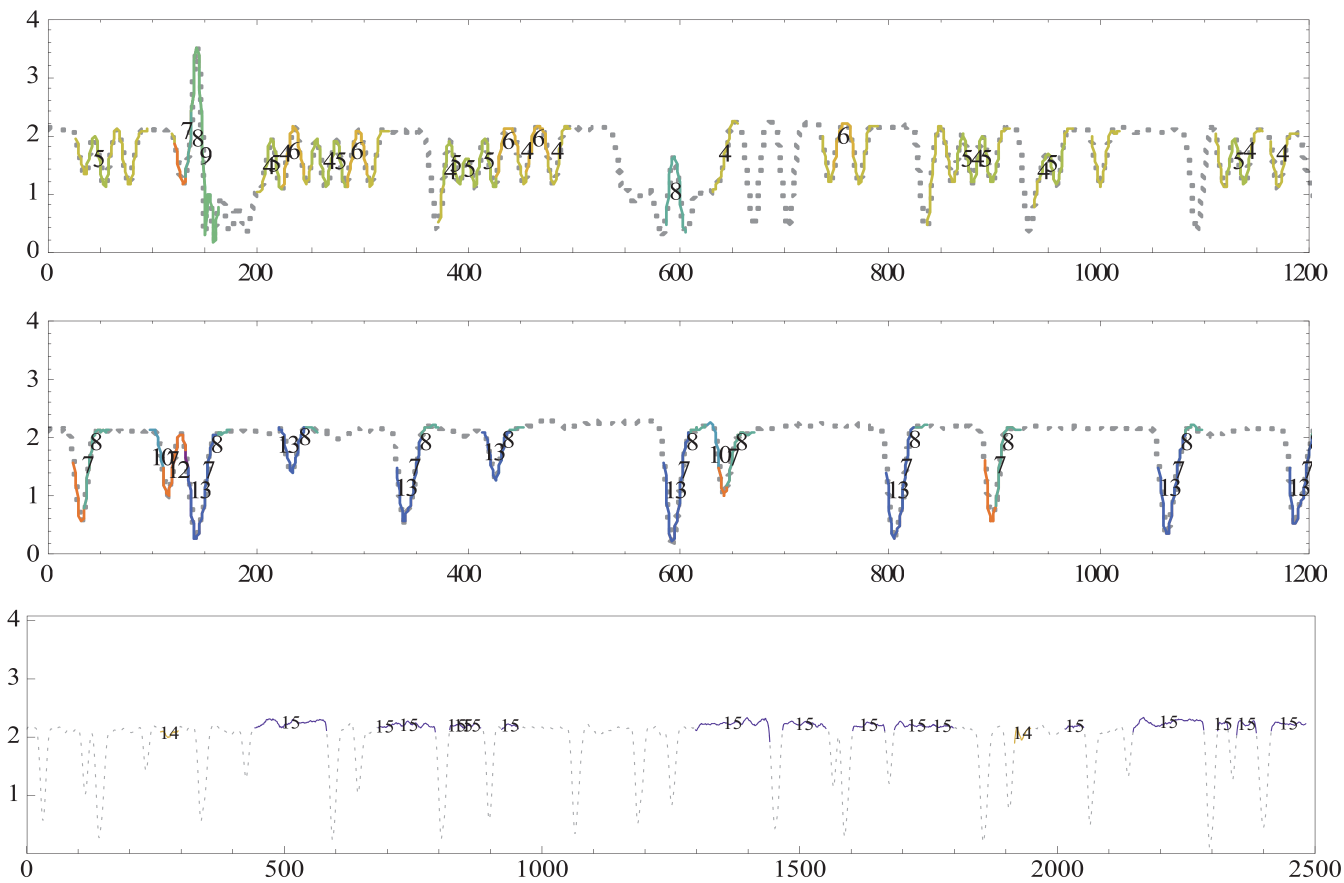
- Graph Formulation
  - $G(V, E)$
  - $E_{ik} = \{e_{i,k}: \mathbf{v}_i(t) \in c_k^i\}$
  - $E = \cup_{i,k} E_{ik}$
  - $V = \{(\mathbf{v}_i(t), \mathbf{v}_j(t'))\}_{\{(t,t'):|t-t'| \leq L\}}$
- Cluster spectral features
- $\{\mathbf{e}_c\}_{c \in C} = GLE(V, E)$
- Cross-distance matrix of $\{\mathbf{e}_c\}$





Spectral–similar Behavior Clusters

- Similar behavior clusters of agents in pursuit
  - Turn by pursued agent
- Association Rule
  - Frequent but not necessarily related pattern



学習問題への応用

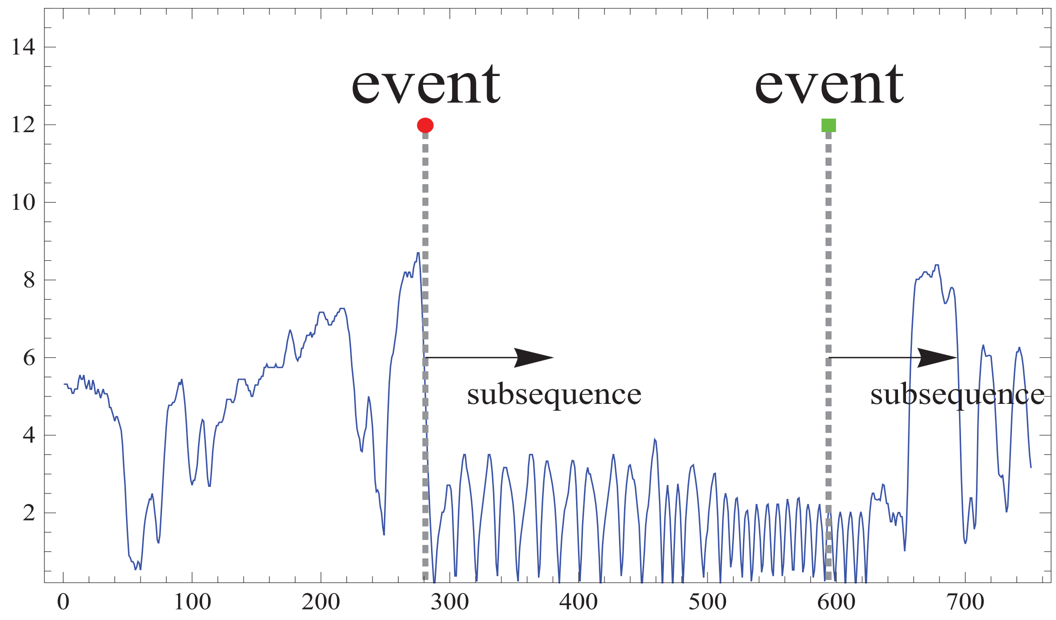
Early Classification of Time Series

- Practical incentives to make prediction early [Xing09]
  - Ex. Medical diagnosis
- Trade-off: response time vs. accuracy
- Goal: make prediction early and maintain accuracy
- Related:
  - Anytime Learning [Esmeir11]

- Time Series (Segmented):
  - $\mathbf{X} = \{X_1, \dots, X_n\}$
- Label: $Y = \{y_i\}_{i=1}^n$
- Prefix:
  - $P_\lambda(X) = \{x(1), \dots, x(\lambda)\}$
- Early Predictor: f_λ
  - $y' = f_\lambda(P_\lambda(X))$
- When to predict
  - Minimum Prediction Length (MPL)

Early Subsequence Classification

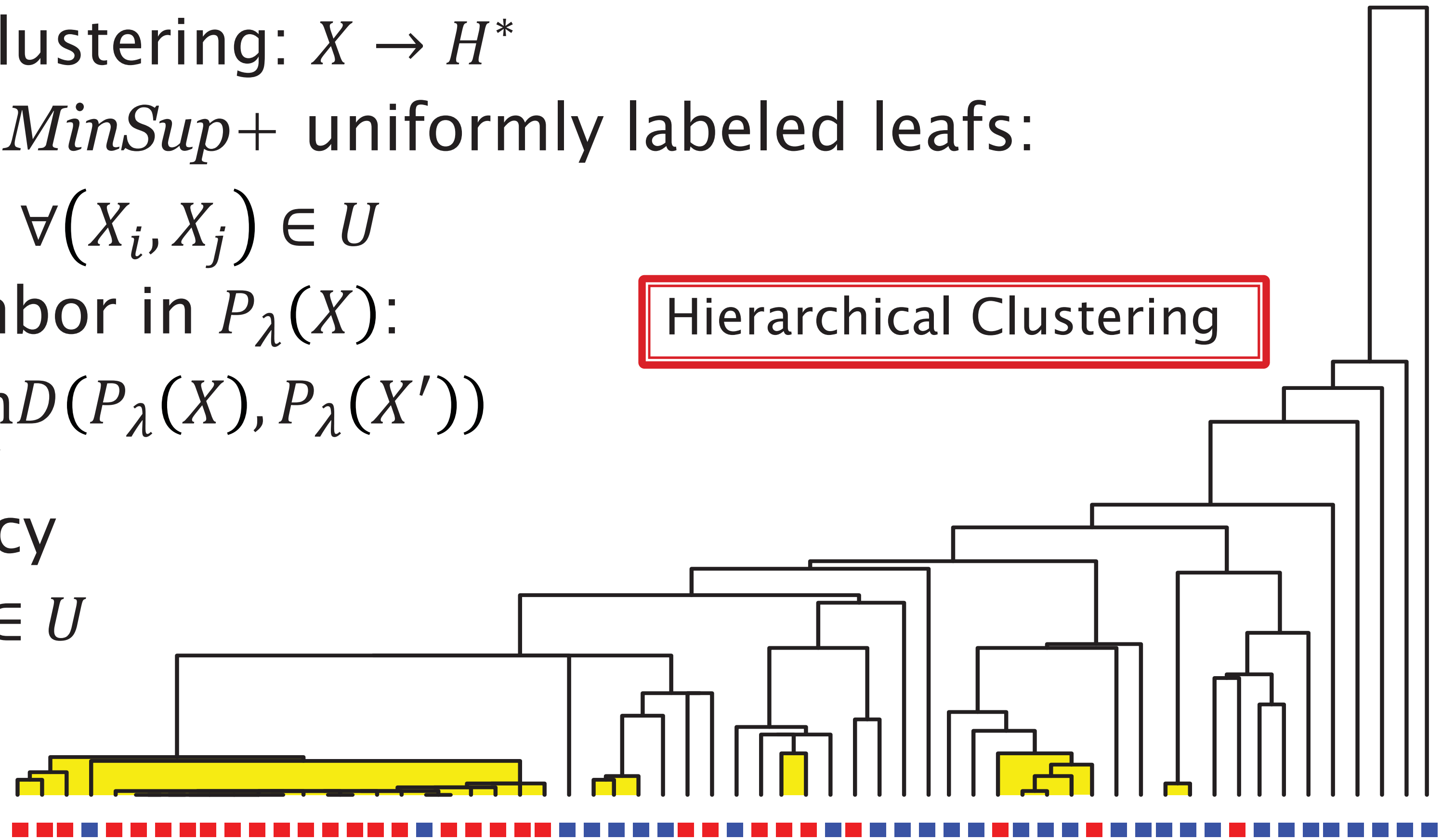
- Predict event from subsequent observations



- Event Label:
  - $Y = \{y(t)\}_{t=1}^{T}$
- Subsequences:
  - $V = \{\mathbf{v}(t)\}_{t=1}^{n}$
- Prefixes:
  - $V_\lambda = \{\mathbf{v}_\lambda(t)\}_{t=1}^{n}$
  - $\mathbf{v}_\lambda(t) = \big(x(t+1), \dots, x(t+\lambda)\big)$
- Early Prediction:
  - $y'_\lambda(t) = f(\mathbf{v}_\lambda(t))$

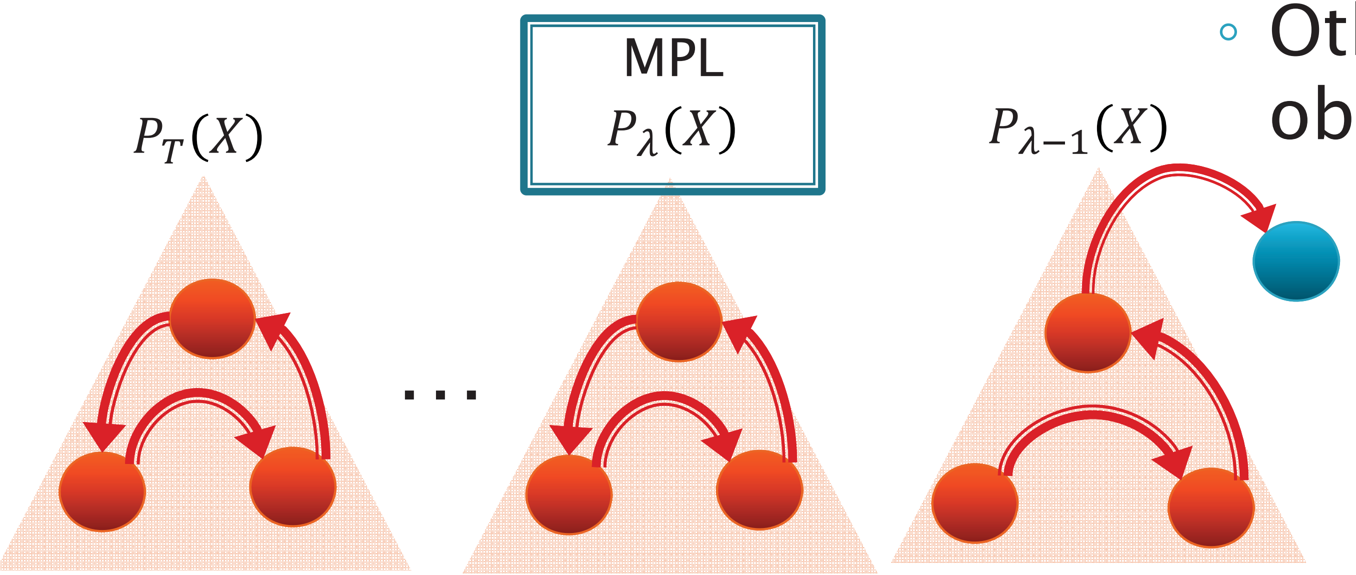
Nearest-neighbor Approach (1)

- 1-NN Classification
  - Predict with label of the nearest training sample
- Procedure (train)
  - Hierarchical Clustering: $X \rightarrow H^*$
  - Subtrees with $MinSup+$ uniformly labeled leafs:
  - $U \in \{H\}: y_i = y_j \ \forall \left(X_i, X_j\right) \in U$
  - Nearest-neighbor in $P_\lambda(X)$:
  - $N_\lambda^*(X) = \underset{X' \in \mathbf{X} \setminus X}{\operatorname{argmin}} D(P_\lambda(X), P_\lambda(X'))$
  - NN-consistency
  - $N_\lambda^*(X) \in U: \forall X \in U$



Nearest-neighbor Approach (2)

- Procedure (contd.)
  - For $X \in U$, $MPL(X)$ is the smallest λ s.t. $N_\lambda^*(X)$ is consistent in U
- Testing: Prefix of test instance $P_\lambda(X')$



- Procedure (test)
  - For $\lambda = \min\{MPL(X)\}, \ldots, T$
  - Find nearest neighbor
  - $N_\lambda^*(X') = \underset{P_\lambda \in \{P_\lambda(X)\}}{\operatorname{argmin}} D(P_\lambda, P_\lambda(X'))$
  - If $MPL(N_\lambda^*) \le \lambda$, make the prediction
  - Otherwise wait for more observations

Guarantees:
Zero error rate with Hold-one-out CV

Early Subsequence Classification

- 1-NN is not as effective with subsequences
- Preliminary Exp.
  - Personal Activity
- Extension to k-NN
  - Make the prediction if $MPL(X) = \lambda : \forall X \in N_{\lambda}^{*}(X')$
- Not improved

	ECTS (1-NN)	ECTS (3-NN)	1-NN
Accuracy	68.2	69.1	80.2
Avg. len.	11.5	13.4	18

Topological consistency-based EC

- Topology-based approach:
  - Goal: maintain neighborhood structure

Consistency of prediction
$\Rightarrow$ Preserved neighborhood structure

- Procedure
  - $\{\mathbf{e}_\lambda(t)\} = GLE\left(RGF_k(V_\lambda)\right)$
  - Supervised k-means: $\{c_i\}_{i=1}^{k}$
  - $f = SVM\left(\left\{\left(\mathbf{e}(t), y(t)\right)\right\}\right)$
  - For $\mathbf{v}(t) \in c_i$, $MPL\left(\mathbf{v}(t)\right)$ is the smallest λ s.t. $y(t) = f\left(\mathbf{e}_\lambda(t)\right), \forall t \in c_i, \lambda \in \{\lambda, \ldots, l\}$

Performances

- Multi-agent task trajectories

	ECTS (1-NN)	Proposed	1-NN
Accuracy	68.2	80.2	80.2
Avg. len.	11.5	14.5	18

- Personal Activities

	ECTS (1-NN)	Proposed	1-NN
Accuracy	71.2	81.2	84.5
Avg. len.	24.3	27.1	32

まとめ

- 時系列サブシーケンスデータの位相は低次部分構造から構成される
  - 局所的には距離関数空間
- 有効なテクニック
  - 近傍グラフのスペクトル分解で距離空間に射影
  - 参照グラフ表現
- 応用例
  - 固有行動の発見，異常検出
  - 時系列に対する早期予測問題

Appendix

References

- [Breunig00] LOF: Identifying Density-based Local Outliers
- [Keogh02] "Probabilistic Discovery of Time Series Motifs"
- [Fern04] "Solving cluster ensemble problems by bipartite graph partitioning" ICML
- [Ide06] "Why does subsequence time-series clustering produce sine waves?," ECML/PKDD
- [Piciarelli06] "On-line trajectory clustering for anomalous events detection", Pattern Recog. Letters
- [Ide07] "Computing Correlation Anomaly Scores Using Stochastic Nearest Neighbors," ICDM
- [Xing09] "Early Prediction on Time Series," IJCAI
- [Yang10] "Handling Movement Epenthesis and Hand Segmentation Ambiguities in Continuous Sign Language Recognition Using Nested Dynamic Programming," IEEE TPAMI
- [Rakthanmanon11] "Time Series Epenthesis: Clustering Time Series Streams Requires Ignoring Some Data," ICDM

References (2)

- [Ando11a] "ACE: Anomaly Clustering Ensemble for Multi-perspective Anomaly Detection in Robot Behaviors," SDM
- [Ando11b] "Role-Behavior Analysis from Trajectory Data by Cross-Domain Learning," ICDM
- [Roddick02l "A Survey of Temporal Knowledge Discovery Paradigms and Methods," IEEE TKDE
- [Keogh02] "On the need for time series data mining benchmarks: a survey and empirical demonstration," SIGKDD
- [Cotofrei02] "Classification Rules + Time = Temporal Rules"
- [Liao05] "Clustering of time series data-a survey," Pattern Recogn.
- [Xing10] "A brief survey on sequence classification," SIGKDD Explorer

Public Benchmarks

- Trajectory Data (UCI Repository)
  - Australian Sign Language
  - Localization Data for Person Activity
  - Wall-Following Robot Navigation Data
  - Vicon Physical Action Data Set
  - EMG Physical Action Data Set
- Surveillance Image Sequence (CAVIAR)
  - http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/
- American Sign Language Lexicon Video Dataset
  - http://vlm1.uta.edu/~athitsos/asl_lexicon
- Multi-task Agent Trajectory (LEMIR)
  - http://www.i.kyushu-u.ac.jp/~suzuki/lemir2011.html
- ECG etc. (Physio.net)
  - http://www.physionet.org
- FX Daily Return Time Series (ACM GECCO 2011 Industrial Challenge)
  - http://gociop.de/gecco-2011-industrial-challenge