This work was supported by the Independent Research Fund Denmark (No. 8022-00366B), the Australian Research Council Grant (Nos. FT180100140 and DP180103411), the NSF-China (No. 61672455), and the NSF of Zhejiang Province of China (No. LY18F020005)

AALBORG UNIVERSITY DENMARK

daisy
Center for Data-intensive Systems

36th IEEE International Conference on Data Engineering (ICDE 2020)

Indoor Mobility Semantics Annotation Using Coupled Conditional Markov Networks

Huan Li[1], Hua Lu[1], Muhammad Aamir Cheema[2], Lidan Shou[3], Gang Chen[3]

[1] Aalborg University [2] Monash University [3] Zhejiang University

AALBORG UNIVERSITY DENMARK

MONASH University

1897 ZHEJIANG UNIVERSITY

Motivation

- To understand **when-where-what** about indoor user movements
  - input: an object's indoor positioning records
  - output: a sequence of **mobility semantics (m-semantics)**, each including
    - a semantic region: e.g., a cashier or a shop
    - a time period
    - a mobility event: movement patterns like stay/pass

Positioning Records → **Positioning Records + Contexts** → **Mobility Semantics**

location, timestamp	
(120.21,62.94, 1F),	12:21:32
(120.61,63.04, 2F),	12:21:54
(120.54,63.01, 1F),	12:22:15
…	
(100.32,50.55, 1F),	12:42:19
(98.62,48.47, 1F),	12:44:26
…	
(53.77,11.01, 1F),	12:45:12
(52.89,10.82, 1F),	12:45:18
…	

John's Hotdog Deli,
12:21:32-12:22:15, stay

……

Food Market,
12:42:19-12:44:26, pass

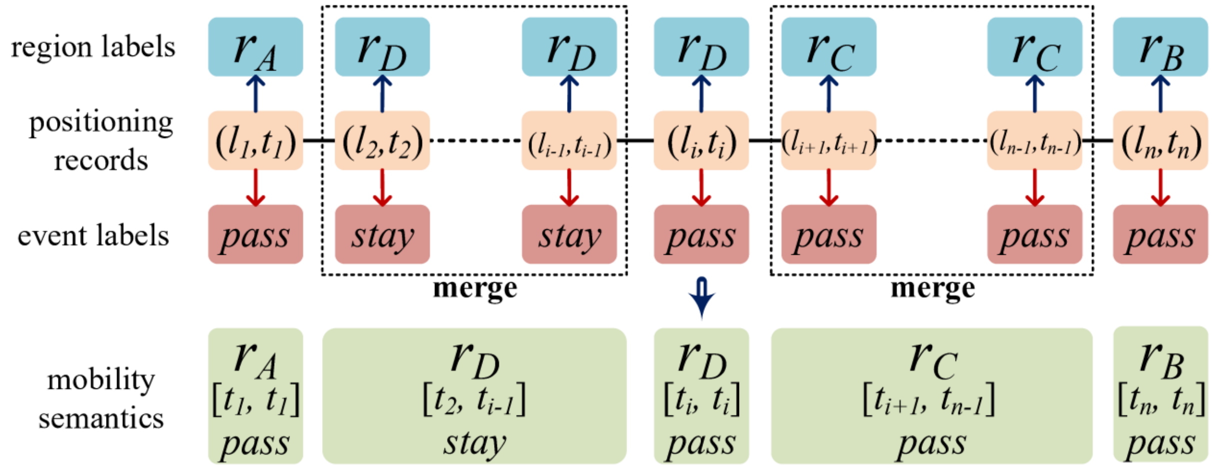
7-Eleven,
12:45:12-12:58:04, stay

……

- **M-semantics**: intuitive, concise, independent of positioning techniques
  - behavior inference: John's Hotdog Deil + stay → buying some food
  - estimate the conversion rate of people (distinction between stay and pass)

Idea and Challenges

- Label-and-Merge for constructing m-semantics
  - **label** each record with a region and an event
  - **merge** the consecutive records having the same region and event labels
  - merging can be performed at different region granularities



- Challenge 1: spatial and temporal uncertainties
  - hard to identify exact whereabouts and mobility states
- Challenge 2: indoor venue with small extent but complex topology
  - outdoor annotation inapplicable: POI category or human activity regularity
- Challenge 3: regions and events are always correlated
  - the correlations over sequence clearly increase the labeling complexity

Technical Contributions

- Graphical model: coupled conditional Markov network
  - **joint relationship** as probabilistic dependencies among the positioning records, region labels, and event labels
  - probabilistic representation: overcome the spatiotemporal uncertainties (Challenge 1)

- Feature functions in the graphical model
  - incorporate **useful knowledge** about indoor topology and indoor mobility behaviors (Challenge 2)

- A novel alternate learning paradigm
  - progressively **estimate optimal parameters** for one label type with the other label type being fixed
  - mitigate the coupling of region labels and event labels (Challenge 3)

Problem Definition

- Positioning Sequence $P_o = \langle \theta_1, \theta_2, \dots, \theta_n \rangle$
  - positioning record $\theta_i = (l_i, t_i)$, location l_i as a 2D point on a floor

- M-Semantics $ms = (r, \tau, e)$
  - region r, time period τ, mobility event $e \in$ {stay, pass}

- M-Semantics Sequence $MS_o = \langle ms_1, ms_2, \dots, ms_m \rangle$

- Given $P_o = \langle \theta_1, \theta_2, \dots, \theta_n \rangle$, the goal of **m-semantics annotation** is to generate the **most-likely** $MS_o = \langle ms_1, ms_2, \dots, ms_m \rangle$

- Labeling problem: configure optimal target variable $\mathbf{Y}$ that maximizes the conditional distribution over observation $\mathbf{X}$

Conditional Markov Networks

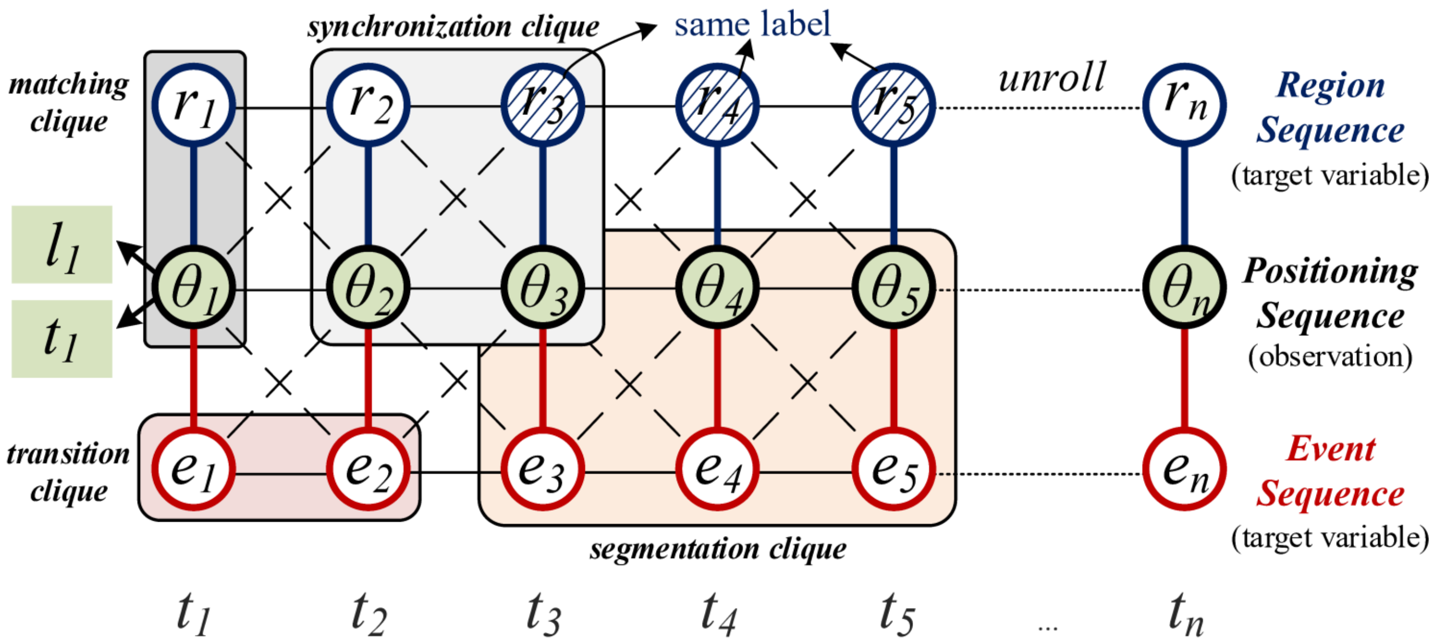
- Conditional Markov Networks (CMNs)
  - $P(\mathbf{y} \in \mathbf{Y} \mid \mathbf{x} \in \mathbf{X})$ factorized as a product of clique potentials $\phi_c(\mathbf{x}_c, \mathbf{y}_c)$
  - clique: a fully-connected sub-graph in the network
  - clique potentials: **compatibility** among clique nodes, i.e., the larger the potential value, the more likely the variable configuration for the clique nodes

$$
\begin{aligned}
\mathrm{P}(\mathbf{y} \mid \mathbf{x}) &= \frac{1}{Z(\mathbf{x})} \prod_{c \in \mathrm{C}} \phi_c(\mathbf{x}_c, \mathbf{y}_c) \\
&= \frac{1}{Z(\mathbf{x})} \prod_{c \in \mathrm{C}} \exp\{\mathbf{w}_c^{\mathsf{T}} \cdot \mathbf{f}_c(\mathbf{x}_c, \mathbf{y}_c)\} \\
&= \frac{1}{Z(\mathbf{x})} \exp\left\{ \sum_{c \in \mathrm{C}} \mathbf{w}_c^{\mathsf{T}} \cdot \mathbf{f}_c(\mathbf{x}_c, \mathbf{y}_c) \right\}
\end{aligned}
$$

- Unrolled CMNs and Parameter Sharing
  - unrolled net: **rather complex** as thousands of nodes involved over the time
  - parameter sharing: each clique **template** corresponds to one **weight vector**

$$
\mathrm{P}(\mathbf{y} \mid \mathbf{x}, \mathbf{w}) = \frac{1}{Z(\mathbf{x}, \mathbf{w})} \exp\left\{ \sum_{ct \in CT} \sum_{c \in \mathrm{C}(ct)} \mathbf{w}_{ct}^{\mathsf{T}} \cdot \mathbf{f}_c(\mathbf{x}_c, \mathbf{y}_c) \right\}
$$

Coupled CMN (C2MN)



- **Matching Cliques**: fitness of an observed node and a target node at a particular timestamp
- **Transition Cliques**: label smoothness of two consecutive target nodes
- **Synchronization Cliques**: transitional consistency for two pairs of an observed node and a target node
- **Segmentation Cliques**: comparability of multiple consecutive pairs of an observed node and a target node, in which the other type of target nodes have the same label

Feature Functions in C2MN

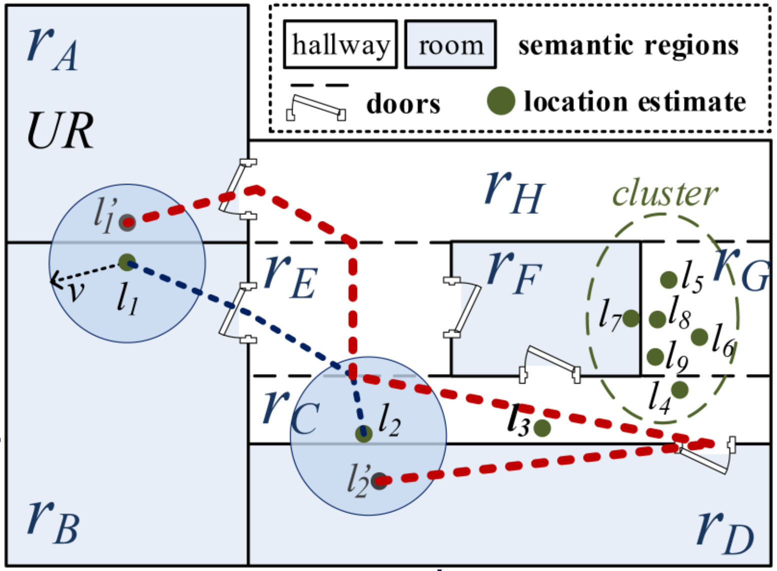
- A feature function: evaluates the **labeling plausibility** of the nodes in a particular type of cliques

- C2MN framework chooses the label configuration with the **highest** overall evaluation as the final result

- The **importance** of each function in the framework is determined by the weights **learned from training data**.

Clique Catagory	Region Relevant Dependencies	Event Relevant Dependencies
Matching Cliques	(1) *Spatial Matching Function* $\mathbf{f}_{sm}(\theta_i, r_i)$	(2) *Event Matching Function* $\mathbf{f}_{em}(\theta_i, e_i)$
Transition Cliques	(3) *Space Transition Function* $\mathbf{f}_{st}(r_i, r_{i+1})$	(4) *Event Transition Function* $\mathbf{f}_{et}(e_i, e_{i+1})$
Synchronization Cliques	(5) *Spatial Consistency Function* $\mathbf{f}_{sc}(\theta_i, \theta_{i+1}, r_i, r_{i+1})$	(6) *Event Consistency Function* $\mathbf{f}_{ec}(\theta_i, \theta_{i+1}, e_i, e_{i+1})$
Segmentation Cliques	(7) *Event-based Segmentation Function* $\mathbf{f}_{es}(c_{es}^{i:j})$	(8) *Space-based Segmentation Function* $\mathbf{f}_{ss}(c_{ss}^{i:j})$

Feature Function Design (I)



(1) Spatial Matching Function

$$\mathbf{f}_{sm}(\theta_i, r_i) = \frac{UR(\theta_i.l, v) \cap Area(r_i)}{UR(\theta_i.l, v)}$$

intersection ratio

(2) Event Matching Function

$$\mathbf{f}_{em}(\theta_i, e_i) = \begin{cases} 1, & \text{if } (e_i, \theta_i.\mathrm{D}) = (stay, \text{core}) \text{ or } (pass, \text{noise}); \\ \alpha, & \text{if } (e_i, \theta_i.\mathrm{D}) = (stay, \text{border}); \\ \beta, & \text{if } (e_i, \theta_i.\mathrm{D}) = (pass, \text{border}); \\ 0, & \text{otherwise.} \end{cases}$$

spatiotemporal density

(3) Space Transition Function

$$\mathbf{f}_{st}(r_i, r_{i+1}) = \exp\left\{ -\gamma_{st} \cdot \mathrm{E}_{p \in r_i, q \in r_{i+1}} [\mathrm{d}_I(p, q)] \right\}$$

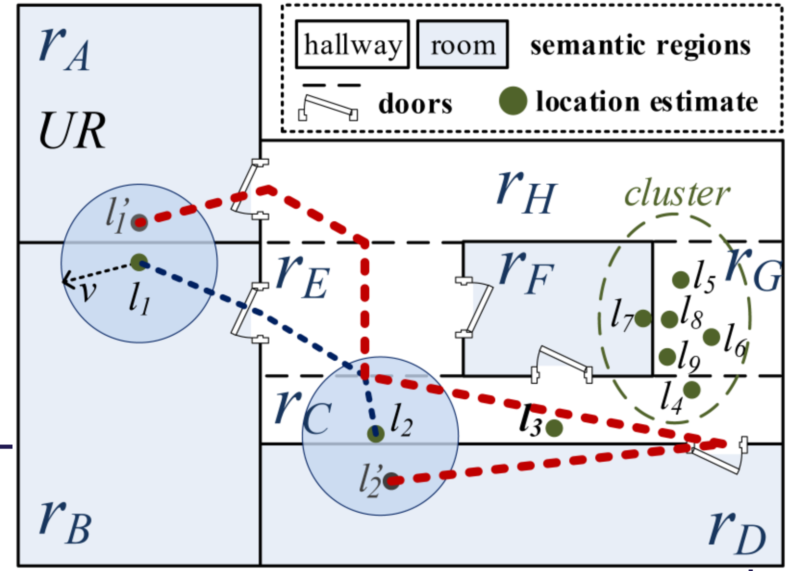
distance between regions

Feature Function Design (II)

(4) Event Transition Function

$$\mathbf{f}_{et}(e_i, e_{i+1}) = \begin{cases} 1, & \text{if } e_i = e_{i+1}; \\ 0, & \text{otherwise.} \end{cases}$$

event change



(5) Spatial Consistency Function

$$\mathbf{f}_{sc}(\theta_i, \theta_{i+1}, r_i, r_{i+1}) = \exp\left\{ -\left|\mathsf{E}_{p\in r_i, q\in r_{i+1}}[\mathsf{d}_I(p,q)] - \mathsf{d}_E(\theta_i.l, \theta_{i+1}.l)\right| \right\}$$

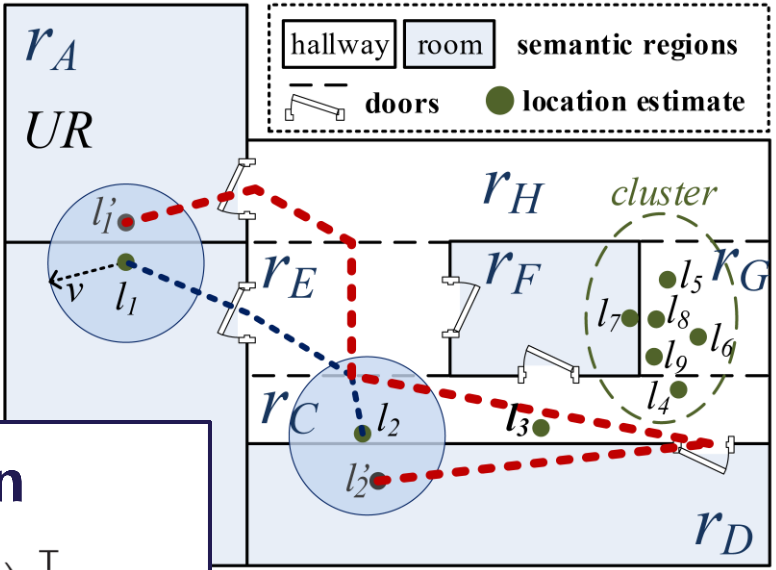
spatial change under indoor topology

(6) Event Consistency Function

$$\mathbf{f}_{ec}(\theta_i, \theta_{i+1}, e_i, e_{i+1}) = \exp\left\{ -\left| \min\left(1, \gamma_{ec} \cdot \frac{\theta_{i+1}.l - \theta_i.l}{\theta_{i+1}.t - \theta_i.t}\right) - \frac{\mathsf{I}_{\triangleright}(e_i) + \mathsf{I}_{\triangleright}(e_{i+1})}{2} \right| \right\}$$

moving speed

Feature Function Design (III)



Segmentation cliques can only be identified when **one target variable** has been **configured**

(7) Event-based Segmentation Function

$$\mathbf{f}_{es}(c_{es}^{i:j}) = \left(2 \cdot \mathsf{I}_{\triangleright}(c_{es}^{i:j}.e) - 1\right) \cdot \begin{pmatrix} DISTNUM(r_i, \ldots, r_j) \\ \sum_{x=i}^{j-1} \mathsf{d}_E(\theta_x.l, \theta_{x+1}.l)/(\theta_j.t - \theta_i.t) \\ -TURNNUM(\theta_i.l, \ldots, \theta_j.l) \end{pmatrix}^{\mathsf{T}}$$

region number, moving distance, turn number

(8) Space-based Segmentation Function

$$\mathbf{f}_{ss}(c_{ss}^{i:j}) = \begin{pmatrix} -NUM(e_i, \ldots, e_j)/(\theta_j.t - \theta_i.t) \\ -(\sum_{x=i}^{j-1} \mathbf{f}_{et}(e_x, e_{x+1})/(\theta_j.t - \theta_i.t) \\ \mathsf{I}_{\triangleright}(e_i) + \mathsf{I}_{\triangleright}(e_j) \end{pmatrix}^{\mathsf{T}}$$

event number, event change number, stay number in head/tail

Supervised Learning of C2MN

- Given full labeled data as (P, R, E), to find all $w_{\mathrm{ct}} \in \boldsymbol{w}$ that maximize $\mathrm{P}(R, E \mid P, \boldsymbol{w})$
- Objective function for C2MN (**parameter sharing**)

$$
\begin{aligned}
\mathsf{L}(\mathbf{w}) &= -\log \mathsf{P}(R, E \mid P, \mathbf{w}) + \frac{\mathbf{w}^{\mathsf{T}}\mathbf{w}}{2\sigma^2} \\
&= \sum_{ct \in CT} \sum_{c \in \mathsf{C}(ct)} \left(-\mathbf{w}_{ct}^{\mathsf{T}} \cdot \mathbf{f}_c(P_c, R_c, E_c) \right) + \log Z(P, \mathbf{w}) + \frac{\mathbf{w}^{\mathsf{T}}\mathbf{w}}{2\sigma^2} \\
&= \sum_{ct \in CT} \left(-\mathbf{w}_{ct}^{\mathsf{T}} \cdot \sum_{c \in \mathsf{C}(ct)} \mathbf{f}_c(P_c, R_c, E_c) \right) + \log Z(P, \mathbf{w}) + \frac{\mathbf{w}^{\mathsf{T}}\mathbf{w}}{2\sigma^2} \\
&= \sum_{ct \in CT} -\mathbf{w}_{ct}^{\mathsf{T}} \cdot \mathbf{f}_{ct}(P_{ct}, R_{ct}, E_{ct}) + \log Z(P, \mathbf{w}) + \frac{\mathbf{w}^{\mathsf{T}}\mathbf{w}}{2\sigma^2} \\
&= -\mathbf{w}^{\mathsf{T}} \cdot \mathbf{f}(P, R, E) + \log Z(P, \mathbf{w}) + \frac{\mathbf{w}^{\mathsf{T}}\mathbf{w}}{2\sigma^2}
\end{aligned}
$$

- Computing $\mathrm{Z}(P, \boldsymbol{w})$ needs to consider all possible configurations for R and E: to employ **pseudo-likelihood** based on Markov blanket (immediate neighbors)

Alternate Learning Algorithm

- **Optimization**: Quasi-Newton method L-BFGS
  - requires both pseudo-likelihood and its gradient

$$\mathsf{PL}(\mathbf{w}) = -\log \sum_{y_i \in R \cup E} \mathsf{P}(y_i \mid \mathrm{MB}(y_i), \mathbf{w}) + \frac{\mathbf{w}^\mathsf{T}\mathbf{w}}{2\sigma^2}$$

$$= -\sum_{y_i \in R \cup E} \left(\mathbf{w}^\mathsf{T} \cdot \mathbf{f}(y_i, \mathrm{MB}(y_i)) + \log Z(\mathrm{MB}(y_i), \mathbf{w})\right) + \frac{\mathbf{w}^\mathsf{T}\mathbf{w}}{2\sigma^2}$$

$$\nabla\mathsf{PL}(\mathbf{w}) = \sum_{y_i \in R \cup E} \left(-\mathbf{f}(y_i, \mathrm{MB}(y_i)) + \mathsf{E}_{\mathsf{P}(y_i' \mid \mathrm{MB}(y_i), \mathbf{w})}[\mathbf{f}(y_i', \mathrm{MB}(y_i), \mathbf{w})]\right) + \frac{\mathbf{w}}{\sigma^2}$$

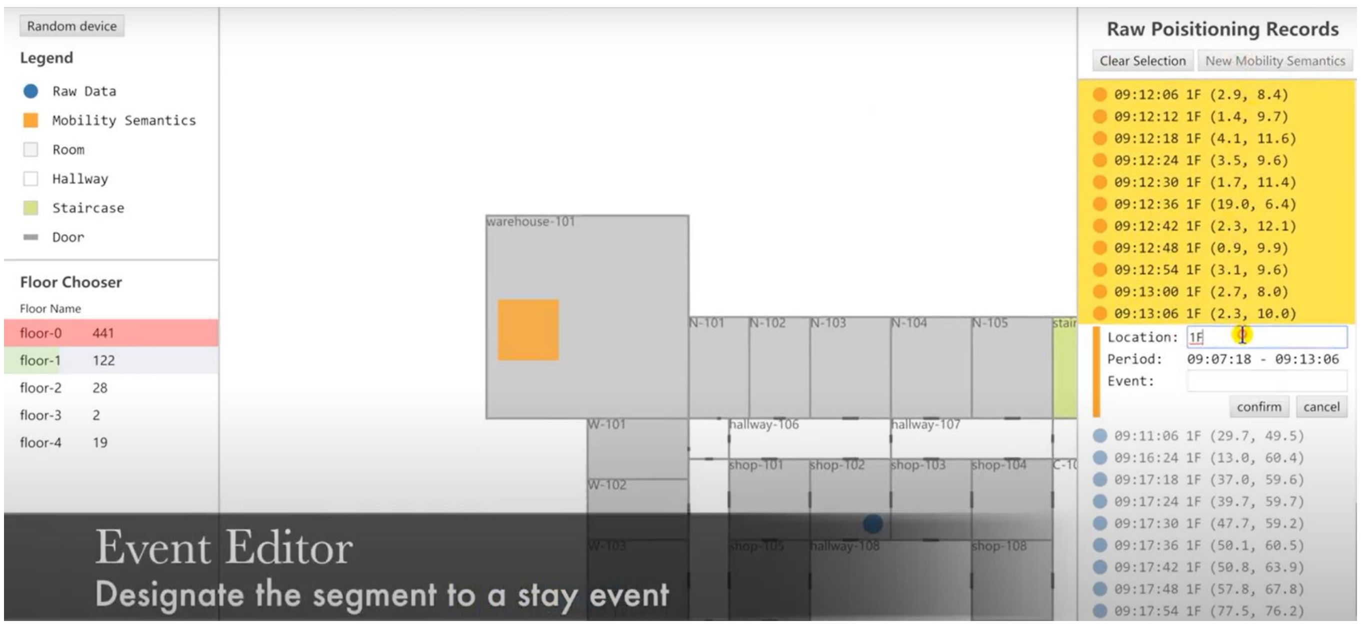
- **Computation**: alternate learning with MCMC inference
  - problem: R and E are **correlated**
  - in each iteration, fix R (or E), and **learn parameters for another**
  - use the current parameters to conduct **MCMC sampling**
  - compute pseudo-likelihood and its gradient
  - **update** parameters and **exchange R and E**
  - **back to the loop** until convergence

Experimental Settings

- Compared Methods
  - **SMoT**: speed threshold (for event) + nearest-neighbor regions
  - **HMM+DC**: HMM for region labeling + Density-based Clustering (i.e., ST-DBSCAN) for event labeling
  - **SAP**: first dynamic-velocity (**SAPDV**) or density-area (**SAPDA**) based event segmentation, then segmentation-level HMMs for regions
  - **CMN**: decouples event and region labeling with two parallel CMNs
  - C2MN variants
    - **C2MN/Tran** without transition cliques
    - **C2MN/Syn** without synchronization cliques
    - **C2MN/ES** without event-based segmentation cliques
    - **C2MN/SS** without space-based segmentation cliques
- Performance Metrics
  - region accuracy (RA); event accuracy (EA)
  - combined accuracy CA = 0.7 * RA + 0.3 * EA
  - perfect accuracy (PA): fraction of records having both labels correct

Real Data Results

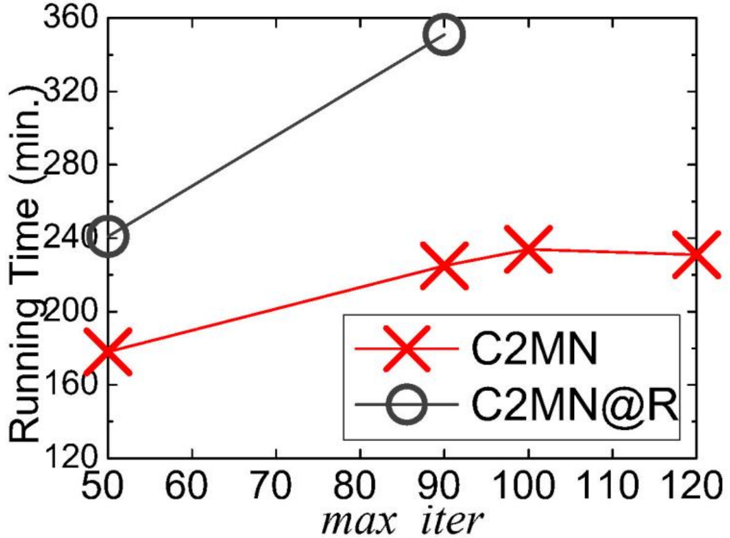
- Wi-Fi positioning data, 7-floor shopping mall, 1 month
  - 5,218,361 positioning records for 44,863 p-sequences
- Training data processing
  - **TRIPS*** backend for coarsely annotation + Event Editor for confirmation (pairs of volunteers)
  - 10-fold cross-validation with a **70/30** train/test split



* TRIPS: A system for translating raw indoor positioning data into visual mobility semantics. Huan Li, Hua Lu, Feichao Shi, Gang Chen, Ke Chen, and Lidan Shou. PVLDB, pp. 1918-1921, 2018.

Real Data Results

Methods	*RA*	*EA*	*CA*	*PA*
SMoT	0.7254	0.8125	0.7515	0.6687
HMM+DC	0.7443	0.8769	0.7841	0.6780
SAPDV	0.7028	0.8296	0.7408	0.6485
SAPDA	0.7394	0.8781	0.7810	0.6943
CMN	0.8860	0.8983	0.8897	0.6684
C2MN/Tran	0.8994	0.9109	0.9026	0.7474
C2MN/Syn	0.9332	0.9073	0.9254	0.8268
C2MN/ES	0.9222	0.9495	0.9304	0.7387
C2MN/SS	0.9014	0.9525	0.9167	0.7616
C2MN	**0.9492**	**0.9691**	**0.9552**	**0.8866**



- **Two-way methods** CMN and HMM+DC, **two-step methods** SMoT and SAPDV/SAPDA, **cannot** learn the interaction between regions and events
- C2MN with a **complete structure** outperforms all its variants
- **Joint labeling** on regions and events significantly improves overall accuracy
- Initially configuring events labels (using ST-DBSCAN) is more efficient

Synthetic Data Results

- Indoor simulator VITA*: generate trajectories with different levels of **sampling frequency** (T) and **positioning error** (u)
- The **quality of m-semantics** cannot be directly measured by labeling accuracy, we turn to the precision of queries
  - Top-k Popular Region Query (TkPRQ)
  - Top-k Frequent Region Pair Query (TkFRPQ)

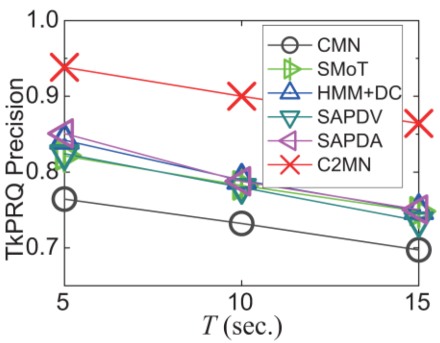


Fig. 15: T*k*PRQ Precision vs. *T*

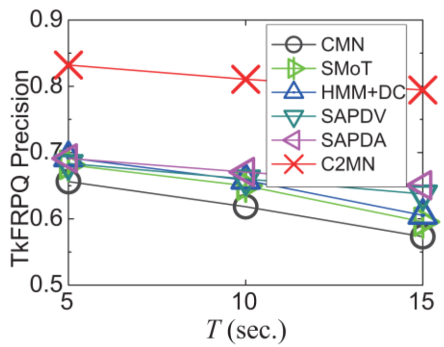


Fig. 16: T*k*FRPQ Precision vs. *T*

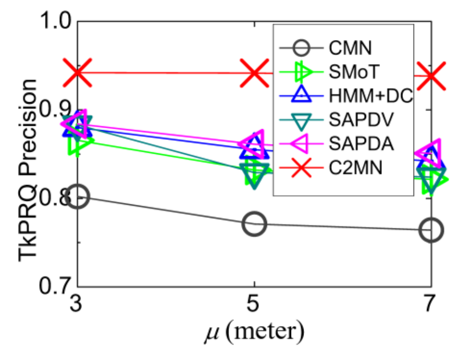


Fig. 18: T*k*PRQ Precision vs. μ

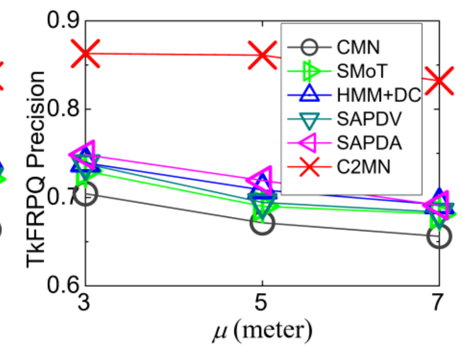


Fig. 19: T*k*FRPQ Precision vs. μ

- C2MN-based approach works very **effectively** even when the **mobility data quality is low**.

* Vita: A versatile toolkit for generating indoor mobility data for real-world buildings. Huan Li, Hua Lu, Xin Chen, Gang Chen, Ke Chen, and Lidan Shou. PVLDB, pp. 1453-1456, 2016.

Summary

- The annotation of indoor mobility data with a semantic region, a time period, and a mobility event
- The **C2MN** is proposed to capture probabilistic dependencies among positioning records, regions, and events
- A set of **feature functions** are designed to incorporate indoor topology and mobility behaviors
- An **alternate learning paradigm** is proposed to enable parameter estimation over coupling of regions and events
- Future directions
  - more diverse mobility events (ontology extension)
  - outdoor scenarios, sparse mobility data
  - NER techniques (e.g., BiLSTM-CRF) for mobility annotation

AALBORG UNIVERSITY
DENMARK

daisy
Center for Data-intensive Systems

Indoor Mobility Semantics Annotation Using Coupled Conditional Markov Networks

Thank You

AALBORG UNIVERSITY
DENMARK

MONASH University

浙江大学 1897 ZHEJIANG UNIVERSITY