

Weekly Meeting

2020.3.18



Why We Go Where We Go: Profiling User Decisions on Choosing POIs

**Renjun Hu^{1*}, Xinjiang Lu^{2*}, Chuanren Liu³, Yanyan Li², Hao Liu², Jingjing Gu⁴,
Shuai Ma¹ and Hui Xiong⁵**

¹SKLSDE Lab, Beihang University & Beijing Adv. Inno. Center for Big Data and Brain Computing

²Business Intelligence Lab, Baidu Research

³Business Analytics and Statistics, University of Tennessee, Knoxville

⁴Nanjing University of Aeronautics and Astronautics

⁵Rutgers University

{hurenjun, mashuai}@buaa.edu.cn, {luxinjiang, liyanyanliyanyan, liuhao30}@baidu.com,
cliu89@utk.edu, gujingjing@nuaa.edu.cn, hxiong@rutgers.edu

Motivation

- little progress has been made for understanding why and how people make their decisions for the selection of POIs;
- decision profiling needs to unify **heterogeneous** factors, e.g., the basic spatiotemporal influence and the hidden preference and functionality impacts for choosing POIs;
- the contributions of factors can differ greatly from one decision to another, which is hard to pre-define, it is more desired to determine the various factor contributions automatically;
- the complex decision structures need to be preserved at the same time.



Find key factors!

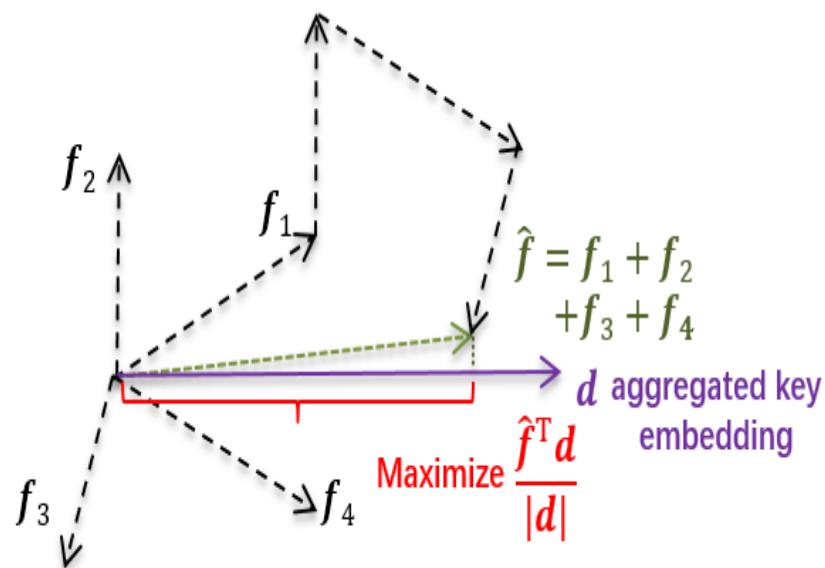
Contributions

- study user decision profiling to provide explanations for people's decisions;
- propose a novel **scalar projection maximization** objective for the problem;
- devise a framework **PROUD** which is able to directly estimate the likelihood of each factor to be a key factor;
- demonstrate the effectiveness of PROUD quantitatively and qualitatively through extensive experiments.

Factors Choosing

- three aspects: user, POI, and context
- User-related factors are user identifier and frequently-visited areas/POIs;
- POI-related factors contain POI identifier, category, brand, and POI popularity;
- context-related ones are decision time (*i.e.*, hour) and the distance to home, work, and POI at the decision time;
- discretize the continuous popularity into six levels based on the standard scores z of log-scaled popularity;
- discretize distance into five levels.

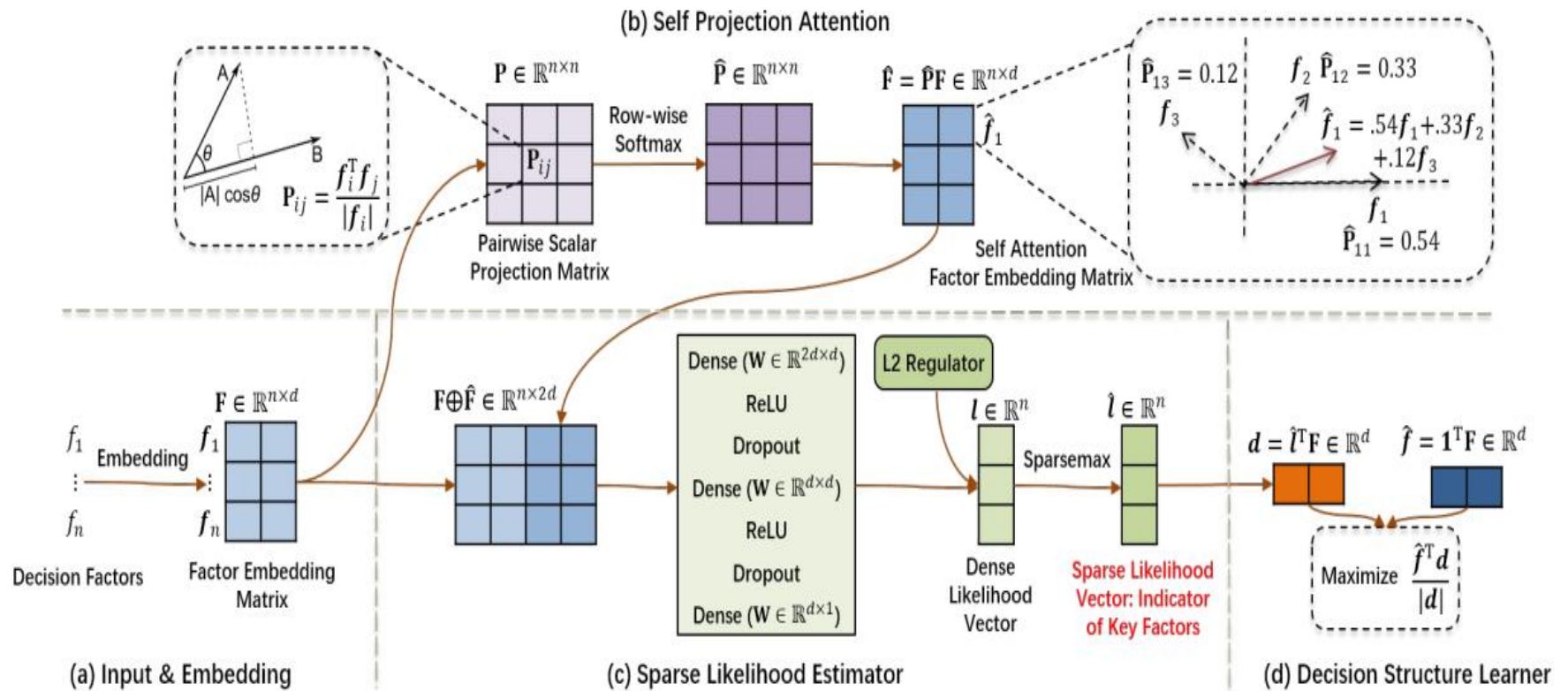
Scalar Projection



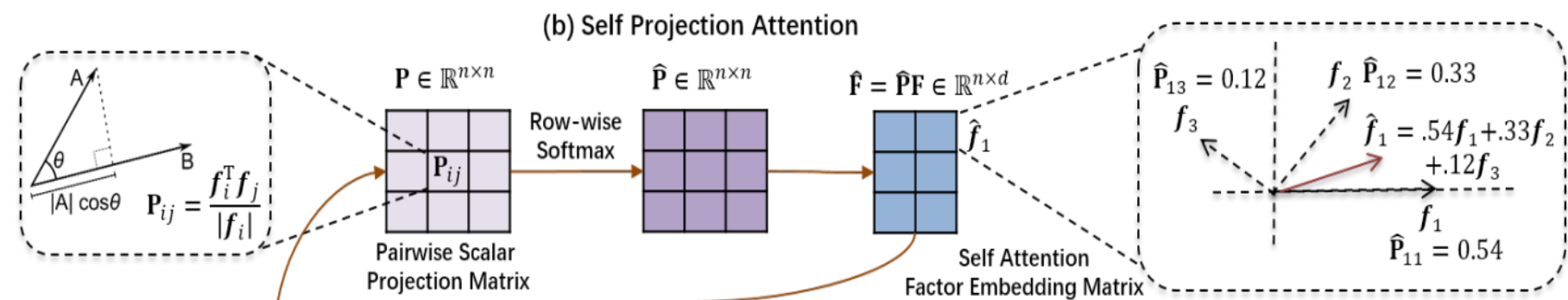
Non-convex
&
NP-hard

The idea is further illustrated by an example in Fig. 2. Suppose we aim to profile a decision $\mathcal{D} = \{f_1, \dots, f_4\}$ by identifying up to two key factors. We learn factor embeddings f_1, \dots, f_4 , and the core is to compute an aggregated key embedding d , e.g., along the direction of $f_1 + f_4$ in our case, such that the sum of scalar projection $\hat{f}^T d / |d|$ is maximized, where $\hat{f} = \sum_{i=1}^4 f_i$ and $|\cdot|$ is the Euclidean norm. Finally, f_1 and f_4 are selected as key factors, while factors f_2 and f_3 only have limited (or opposite) impacts to the decision.

Architecture



Self Projection Attention



Formally, given factors f_1, \dots, f_n in a decision and the corresponding factor embedding matrix $\mathbf{F} = [\mathbf{f}_1, \dots, \mathbf{f}_n]^T$, it first computes a pairwise scalar projection matrix $\mathbf{P} \in \mathbb{R}^{n \times n}$, in which $\mathbf{P}_{ij} = \mathbf{f}_i^T \mathbf{f}_j / |\mathbf{f}_i|$ is the scalar projection of \mathbf{f}_j on \mathbf{f}_i . It then normalizes \mathbf{P} with row-wise softmax:

$$\hat{\mathbf{P}}_{i:} = \text{softmax}(\mathbf{P}_{i:}), \quad i \in \{1, \dots, n\}. \tag{2}$$

Decision Structure Learner

Negetive sampling:

for each positive decision D , we can generate several negative instances D^- by replacing the POI-related factors with factors of other POIs that the user does not decide to visit. Typical alternative POIs for negative instances can be those ***near the visited one*** or those of ***the same category as the visited POI***

$$O = - \sum_{\mathcal{D}} \log \text{VR}(\mathcal{D}) - \sum_{\mathcal{D}^-} \log(1 - \text{VR}(\mathcal{D}^-)).$$

Maximize $\text{VR}(\mathcal{D})$ or minimize $\text{VR}(\mathcal{D}^-)$ \longrightarrow maximize or minimize $\hat{\mathbf{f}}^\top \mathbf{d} / |\mathbf{d}|$

Experiments

- Datasets:

Description	BEIJING	NYC
time spanning	3/20/18~8/30/18	4/12/12~2/16/13
# of users	90,090	1,083
# of POIs	169,528	109,018
# of positive \mathcal{D}	199,106	146,325
# of negative \mathcal{D}^-	1,694,365	1,282,302
# of factors per $\mathcal{D}/\mathcal{D}^-$	20.5	45

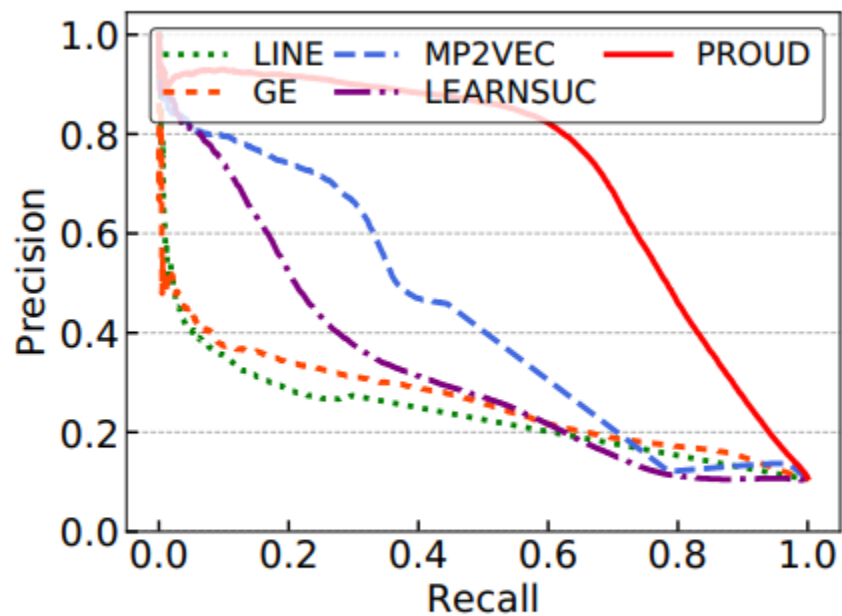
- Metrics: Prec (Precision), Recall, F1, and AUC (Area Under the ROC Curve)

Experiments

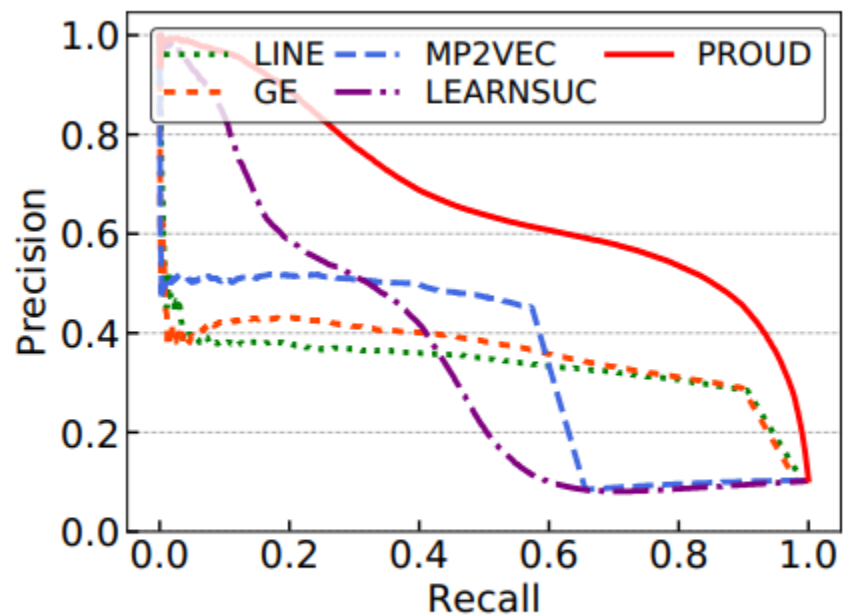
Algorithm	BEIJING			
	Prec	Recall	F1	AUC
LINE	0.2684	0.2433	0.2552	0.5827
GE	0.3005	0.3499	0.3234	0.6271
MP2VEC	0.4594	0.4180	0.4377	0.6801
LEARNSUC	0.2222	0.5849	0.3221	0.6791
PROUD	0.7637*	0.6375*	0.6949*	0.9248*

NYC			
Prec	Recall	F1	AUC
0.3528	0.3835	0.3261	0.6459
0.3767	0.5262	0.4391	0.7134
0.4665	0.5247	0.4939	0.7281
0.3984	0.4011	0.3993	0.5600
0.5487*	0.7743*	0.6422*	0.9439*

Experiments



(a) BEIJING



(b) NYC

Experiments

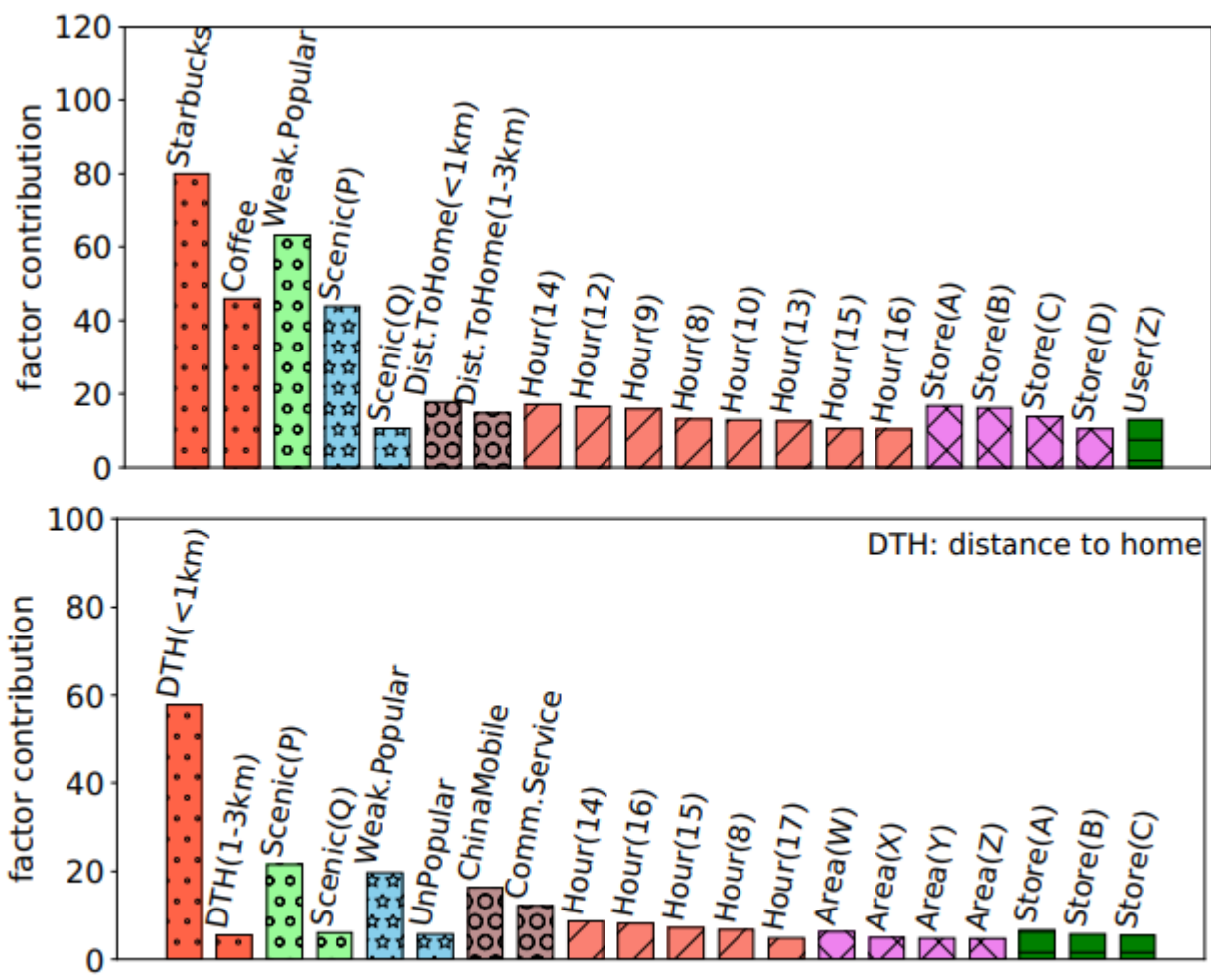


Figure 5: Top-20 key factors of Starbucks and ChinaMobile