KDD2019

Adversarial Substructured Representation Learning for Mobile User Profiling

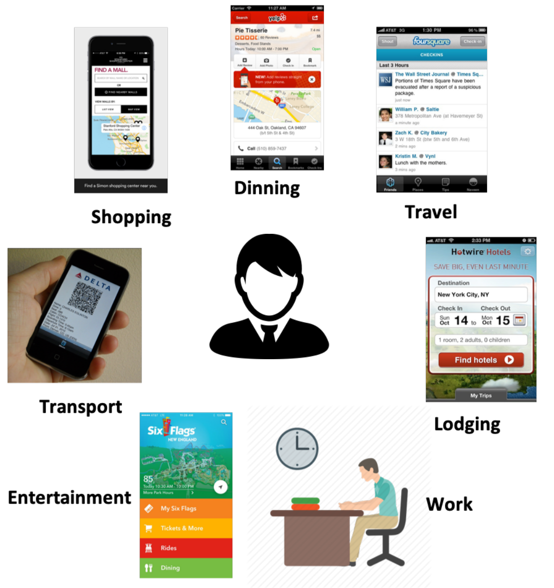
Presenter: Pengyang Wang

Pengyang Wang, Yanjie Fu, Xiaolin Li, Hui Xiong

Outline

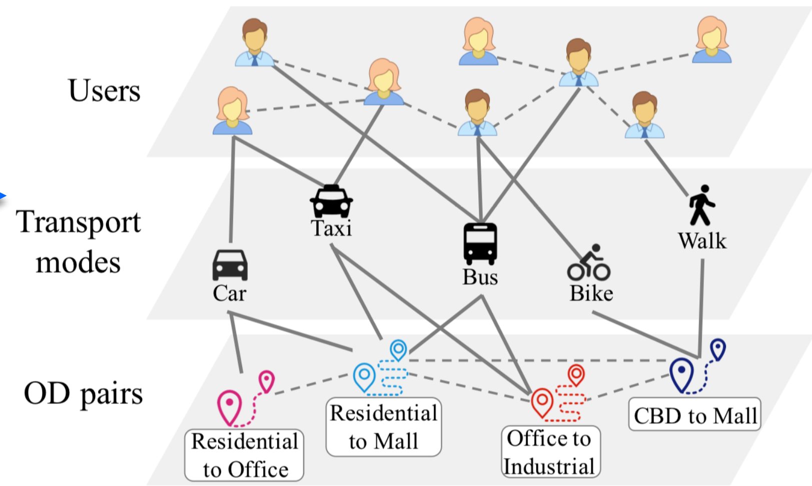
- Background and Motivation
- Definition and Problem Statement
- Methodology
- Evaluation
- Conclusion

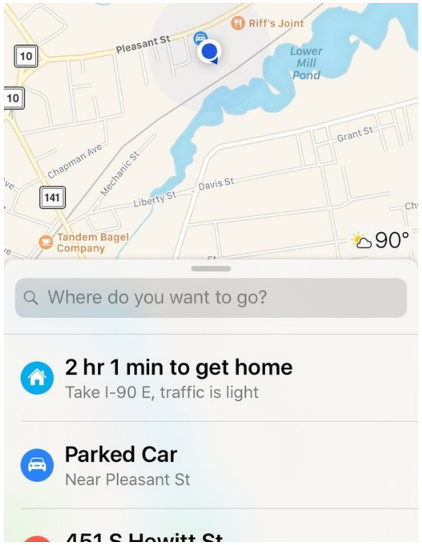
Motivation Application: Toward Adaptive User Interfaces



Mobile user profiling

A similarity graph of users, transportations, OD pairs





Adaptive interfaces by:

(1) inferring trip purposes,

(2) transport modes,

(3) origin-destination pairs

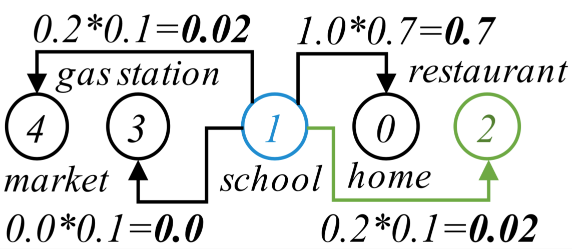
to improve user engagement and performances

Challenge I: Implicit User Patterns in Mobile Activities

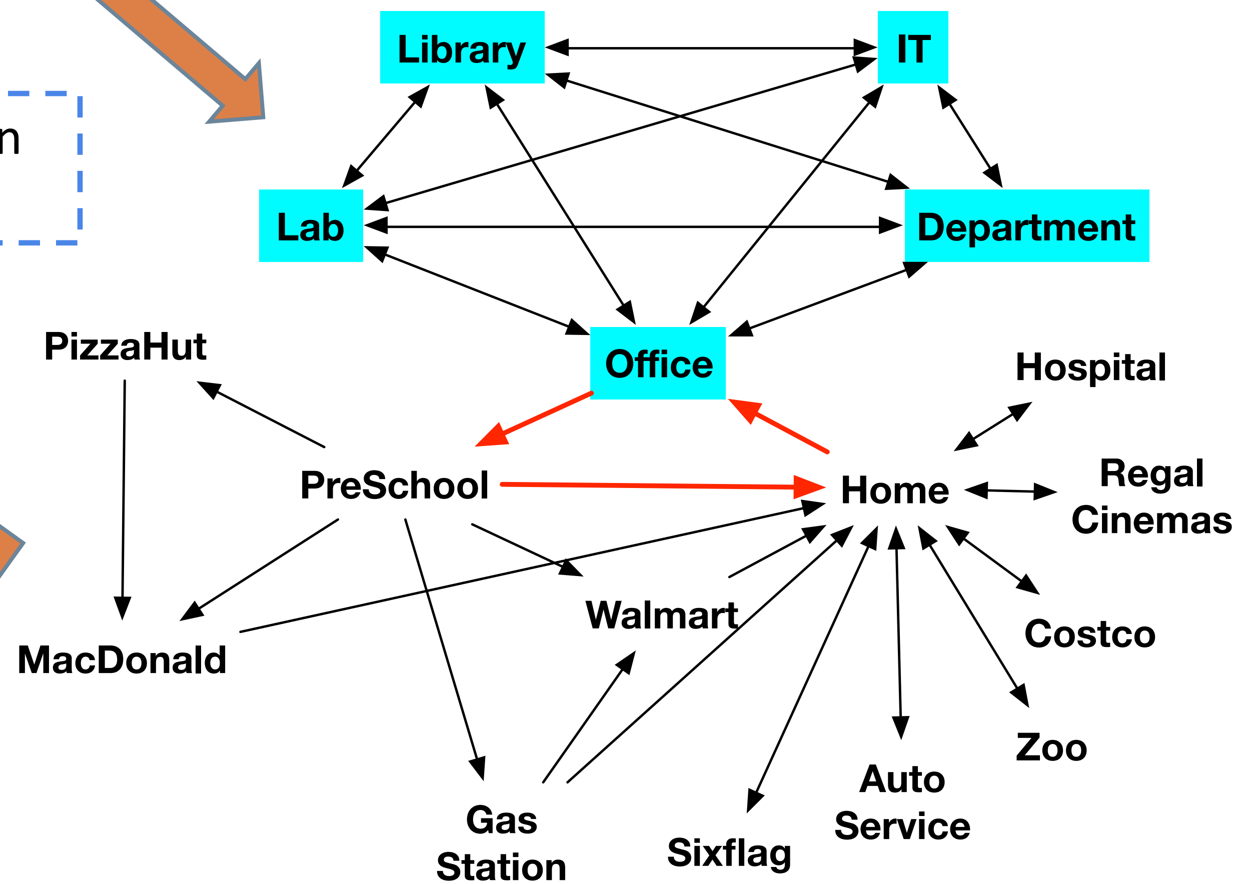
- Human activities are spatially, temporally, and socially structural.

- How can we identify a data structure to better describe a mobile user's activities?

From Users To Activity Graphs

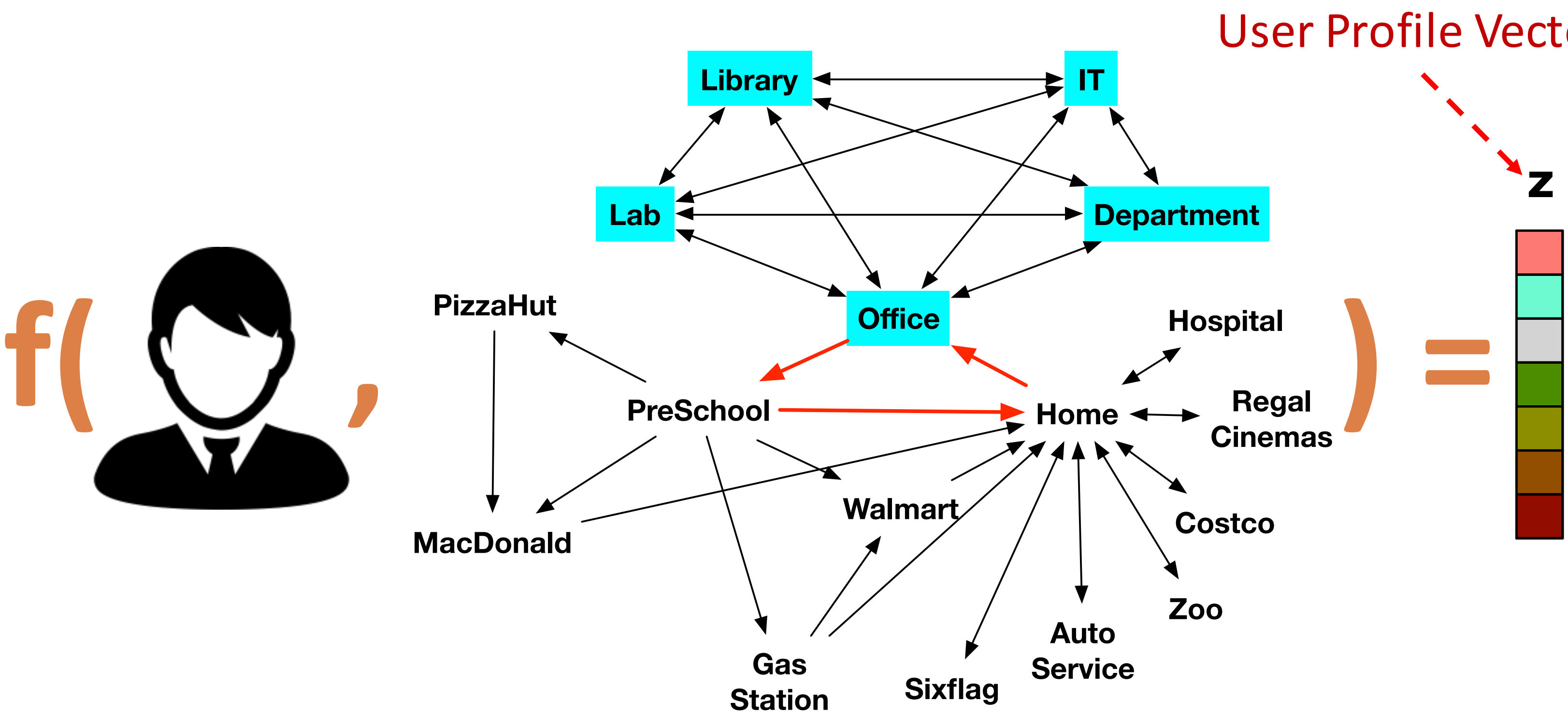


Spatial-temporal transition patterns



User Activity Graph

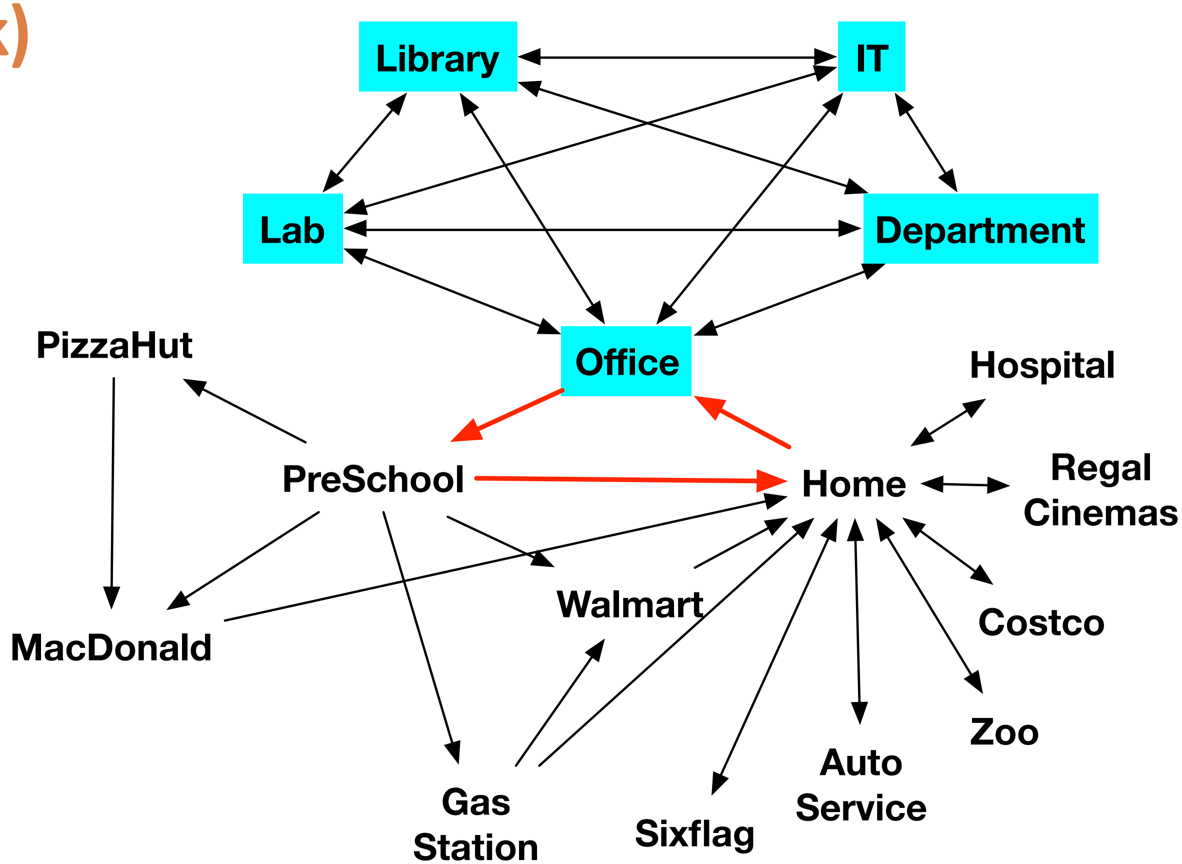
Problem Formulation: Representation Learning with Activity Graphs



- Given a user and corresponding user activity graph, we aim to map the user to a profile vector

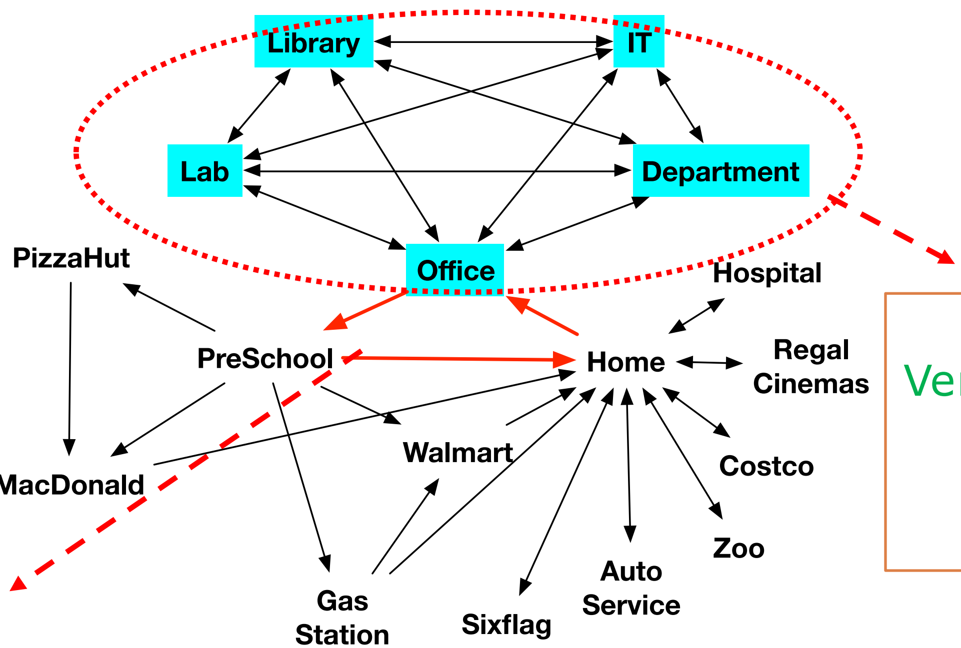
Global Behavioral Pattern

- Entire structures: how a user's activities globally interact with each other (strongly link, weakly link, no link)



Substructure Behavioral Pattern

- **Substructures: topology of subgraphs that feature the unique behavioral patterns of a user's activities**



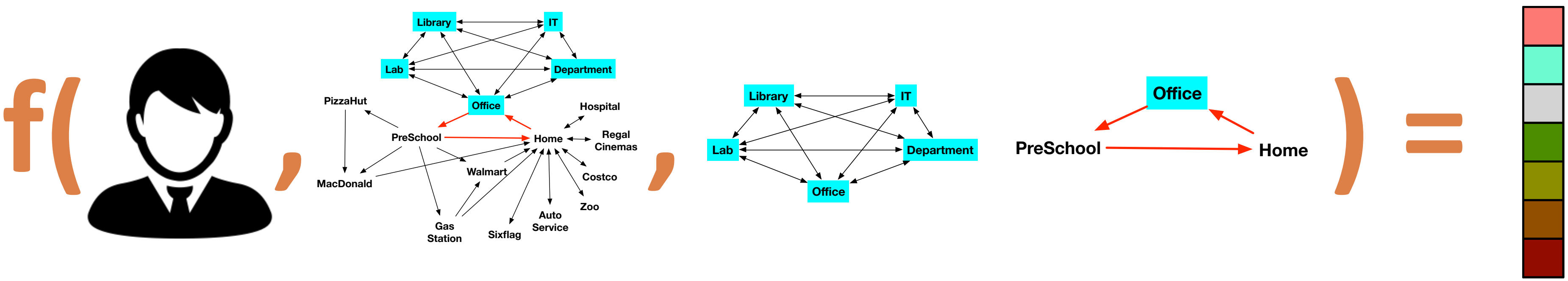
High-frequency Vertexes: Personalized Interests For A Specific Type of Activities

Circle: Personalized Preference For A Close-loop Consecutive Activity Pattern

A Young Faculty with Young Kids

Problem Reformulation: Representation Learning with Global and Sub-Structure Awareness

User Profile Vector



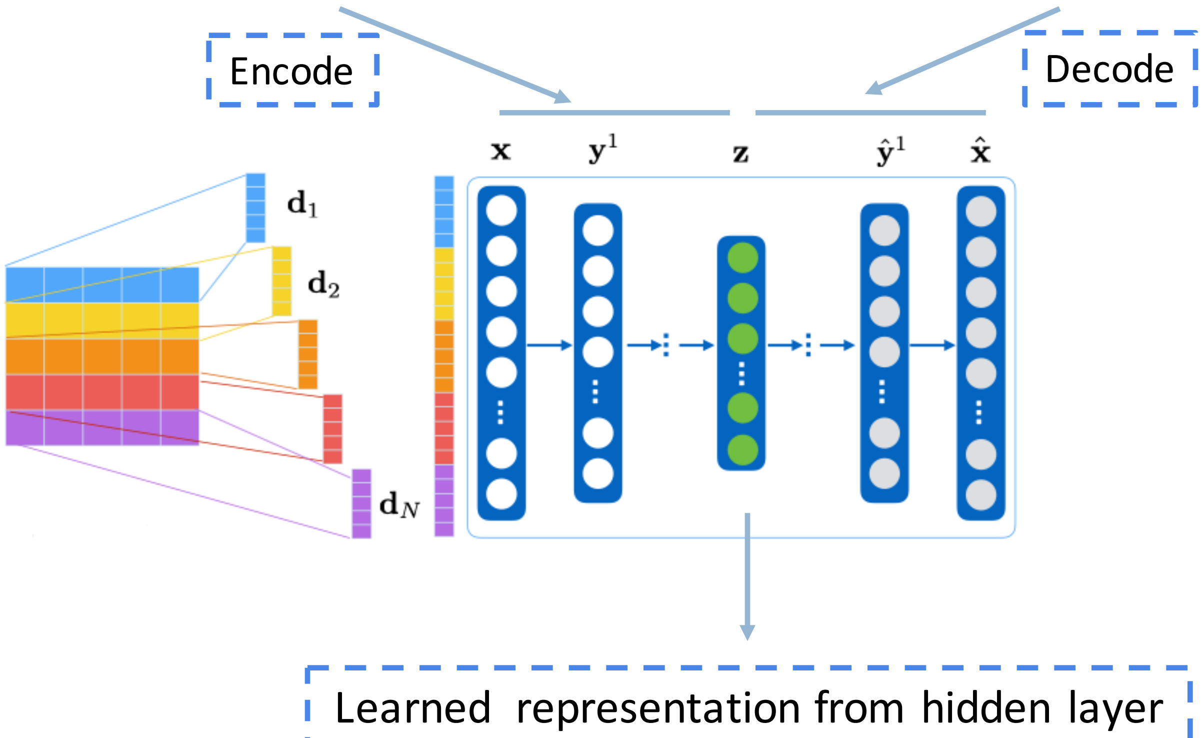
Entire Structure Patterns

Substructure Patterns

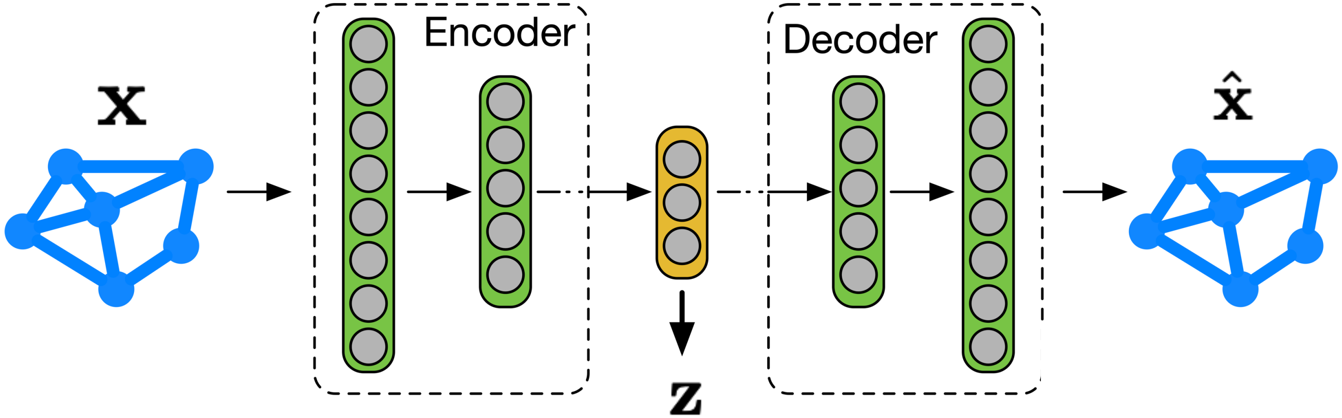
Preserving Entire-Structures

$$\begin{cases} \mathbf{y}_i^1 & = \sigma(\mathbf{W}^1 \mathbf{x}_i + \mathbf{b}^1), \\ \mathbf{y}_i^k & = \sigma(\mathbf{W}^k \mathbf{y}_i^{k-1} + \mathbf{b}^k), \forall k \in \{2, 3, \cdots, o\}, \\ \mathbf{z_i} & = \sigma(\mathbf{W}^{o+1} \mathbf{y}_i^o + \mathbf{b}^{o+1}). \end{cases}$$

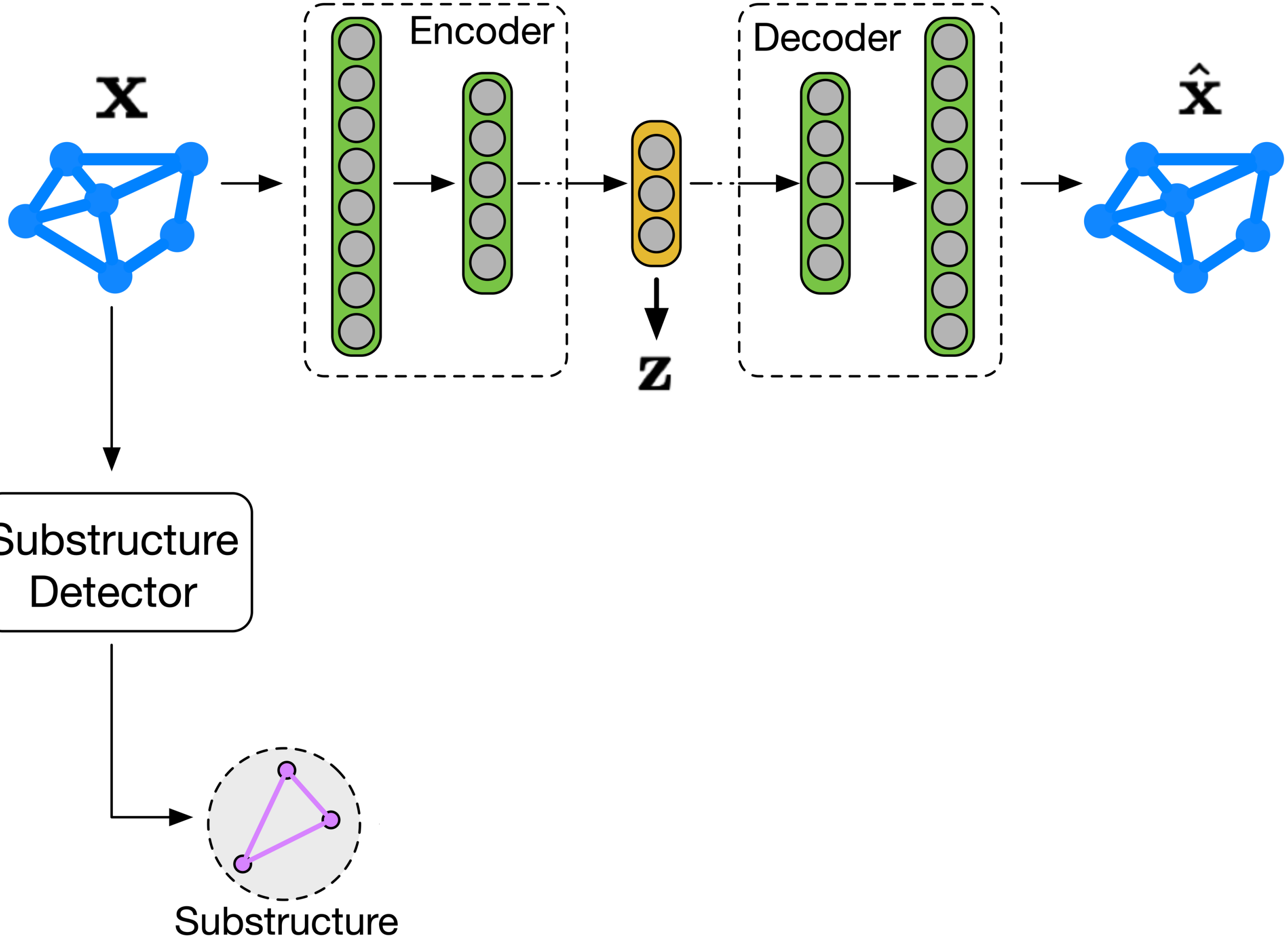
$$\begin{cases} \hat{\mathbf{y}}_i^o & = \sigma(\hat{\mathbf{W}}^{o+1} \mathbf{z}_i + \hat{\mathbf{b}}^{o+1}), \\ \hat{\mathbf{y}}_i^{k-1} & = \sigma(\hat{\mathbf{W}}^k \hat{\mathbf{y}}_i^k + \hat{\mathbf{b}}^k), \forall k \in \{2, 3, \cdots, o\}, \\ \hat{\mathbf{x}}_i & = \sigma(\hat{\mathbf{W}}^1 \hat{\mathbf{y}}_i^1 + \hat{\mathbf{b}}^1). \end{cases}$$



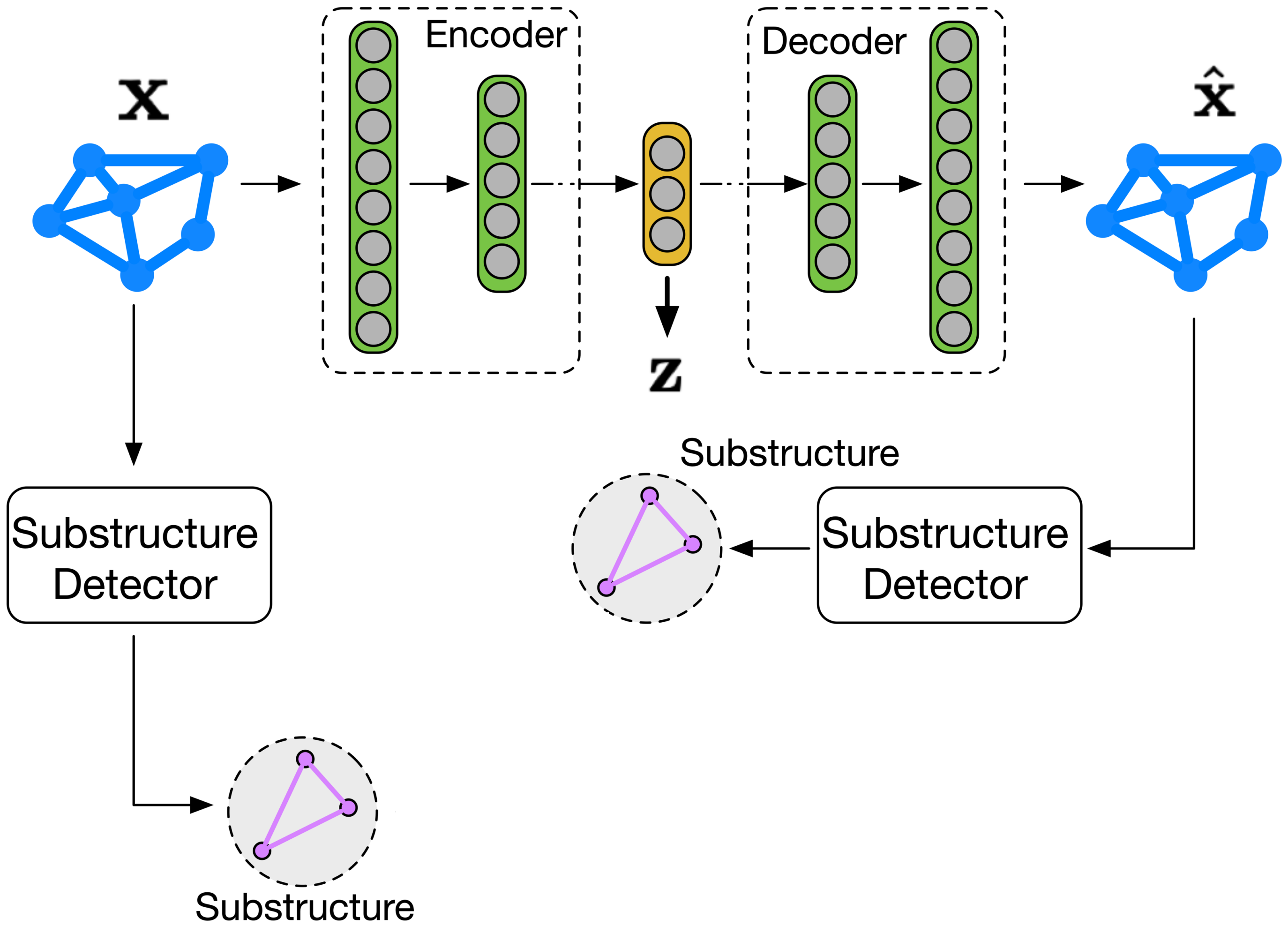
Preserving Substructures



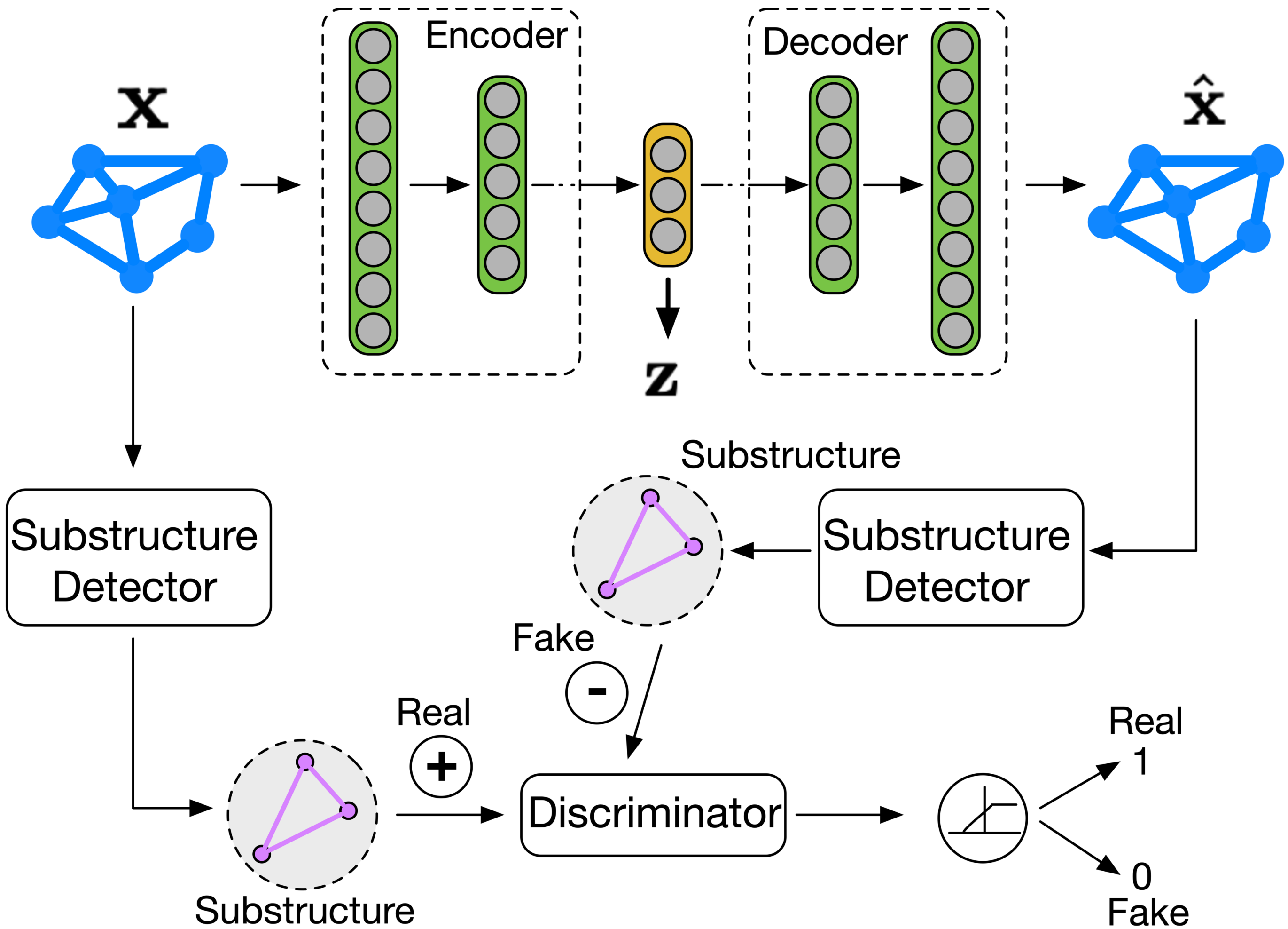
Preserving Substructures



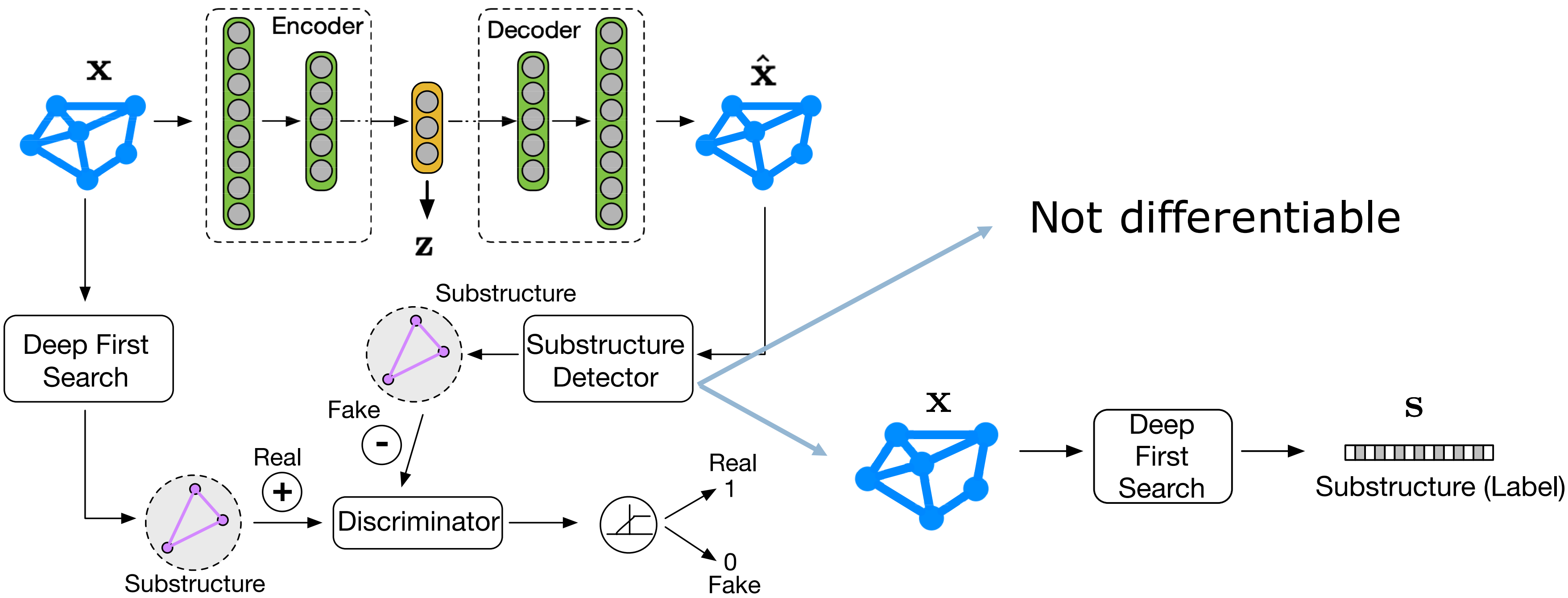
Preserving Substructures



Preserving Substructures

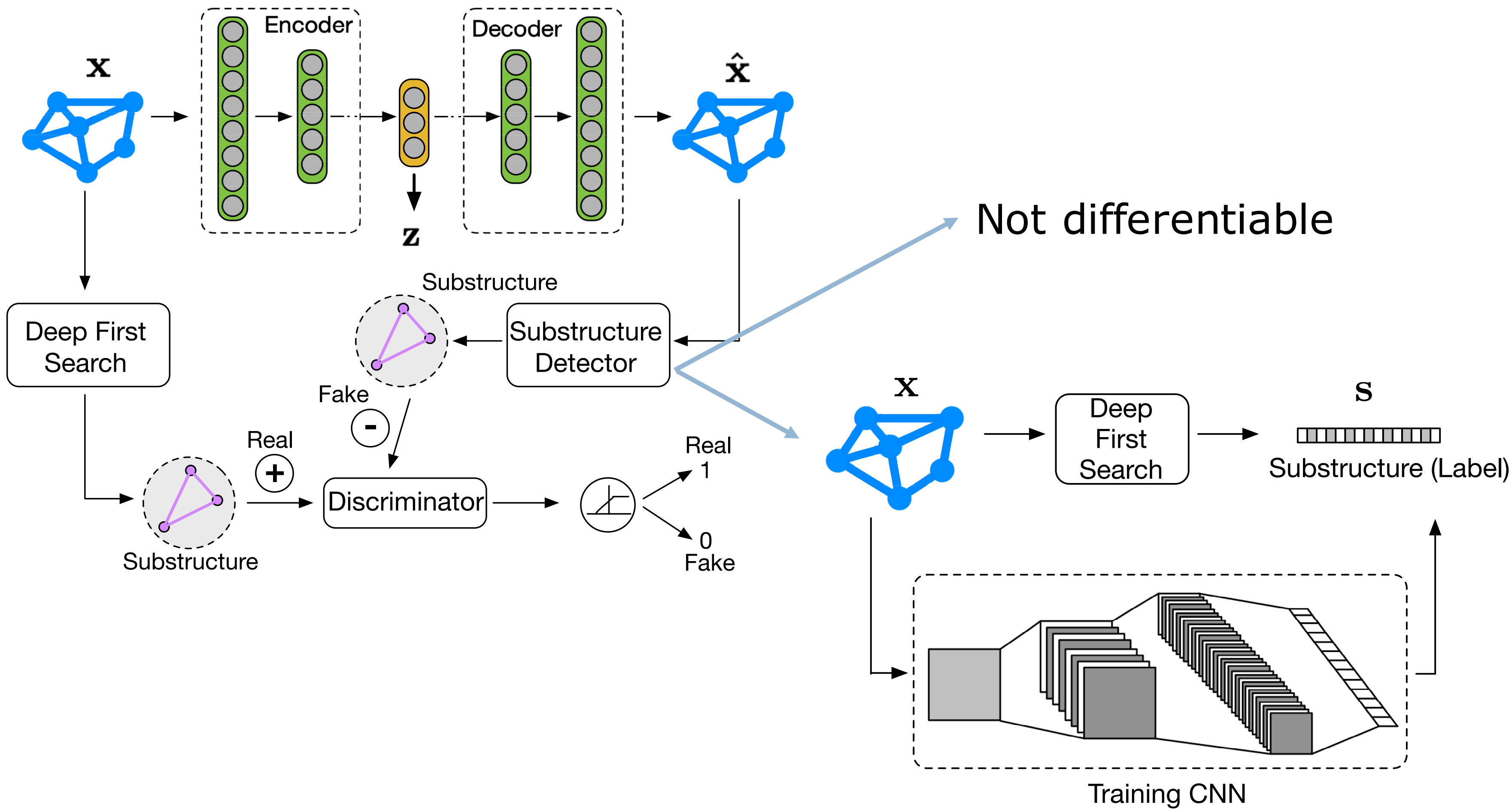


Approximating Substructure Detector



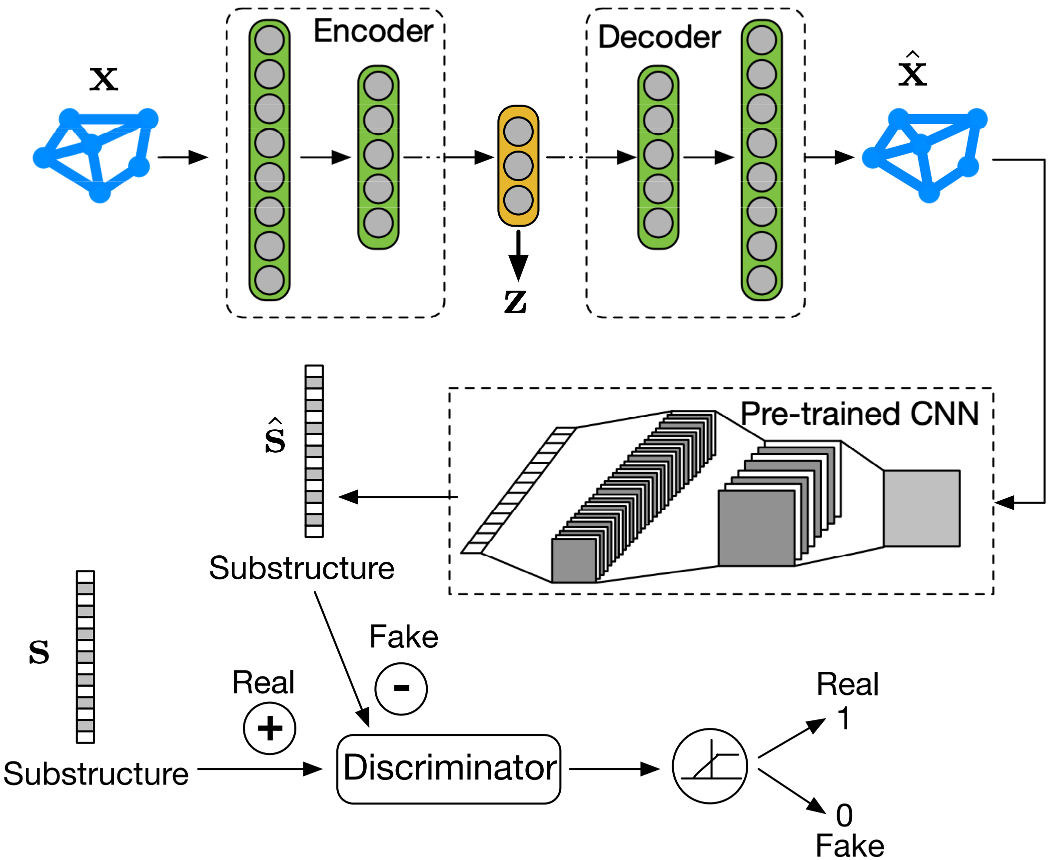
Pre-train a Convolutional Neural Network (CNN) to approximate the traditional substructure detector

Approximating Substructure Detector



Pre-train a Convolutional Neural Network (CNN) to approximate the traditional substructure detector

Summary



- **Generator**
  Autoencoder linked with an approximated substructure detector (pre-trained CNN)
- **Discriminator**
  A multilayer percetron
- **Adversarial Training**
  - Discriminator accuracy

$$\mathcal{L}_D = \frac{1}{m}\sum_{i=1}^{m}[\log D(\mathbf{s}_i) + \log(1 - D(G(\mathbf{x}_i)))]$$

  - Generator loss

$$\mathcal{L}_G = \frac{1}{m}\sum_{i=1}^{m}\log(1 - D(G(\mathbf{x}_i)))]$$

Optimization

- ## Training

Reconstruction Loss

$$\mathcal{L}_{AE} = \frac{1}{2}\sum_{i=1}^{m}\|(\mathbf{x}_i - \hat{\mathbf{x}}_i)\|_2^2$$

$$\mathcal{L} = -\lambda_D \mathcal{L}_D + \lambda_G \mathcal{L}_G + \lambda_{AE} \mathcal{L}_{AE}$$

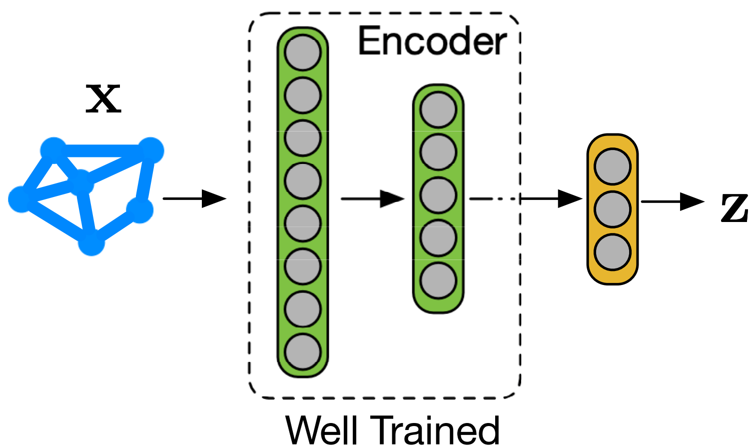
Discriminator Loss

$$\mathcal{L}_D = \frac{1}{m}\sum_{i=1}^{m}[\log D(\mathbf{s}_i) + \log(1 - D(G(\mathbf{x}_i)))]$$

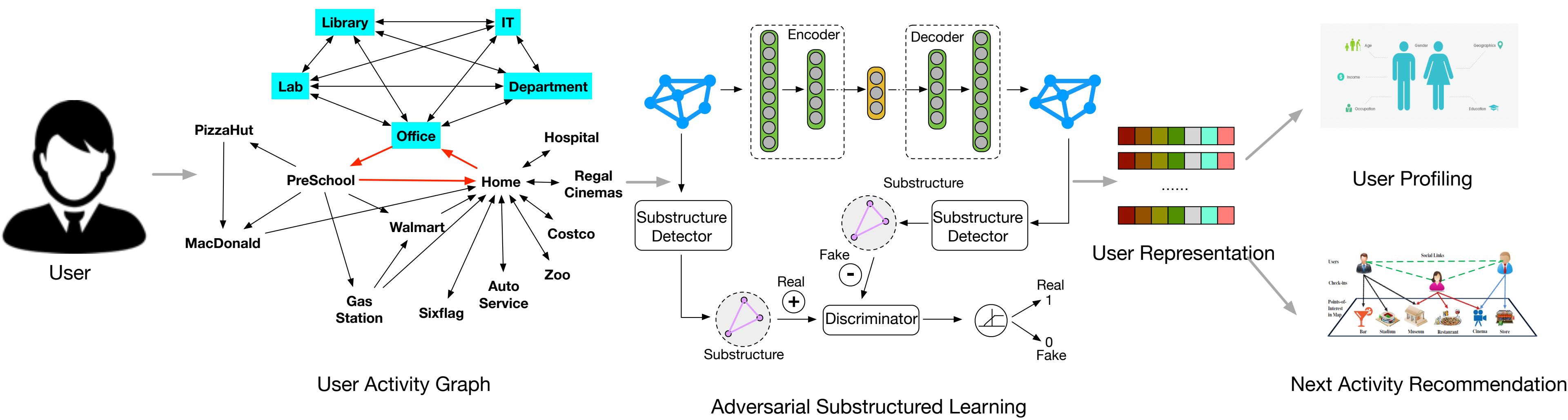
Generator Loss

$$\mathcal{L}_G = \frac{1}{m}\sum_{i=1}^{m}\log(1 - D(G(\mathbf{x}_i)))]$$

- ## Testing



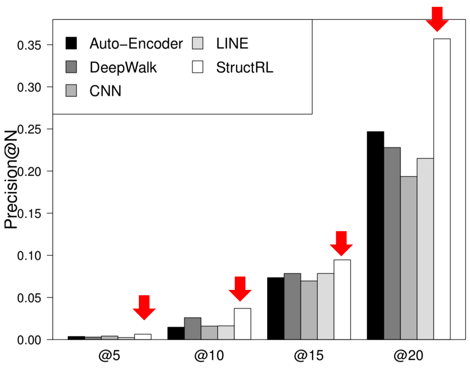
What To Do Next: Inferring Next Activity Type



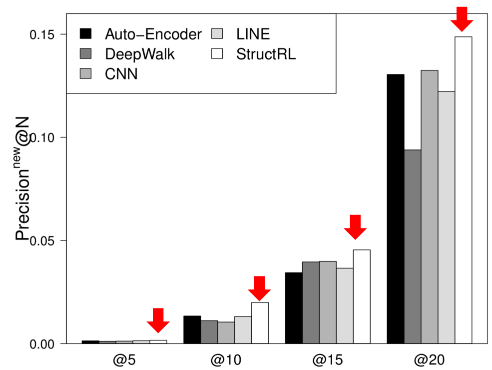
1. Given a time period, learn a user's profiles from corresponding user activity graph
2. Exploit user profiles to forecast next activity type

Overall Comparisons on New York and Tokyo Activity Check-in Data

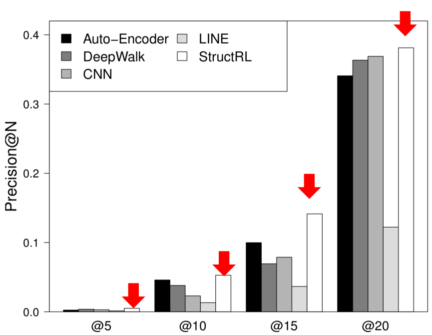
Apply the learned representations to predict next activity type (next POI category)



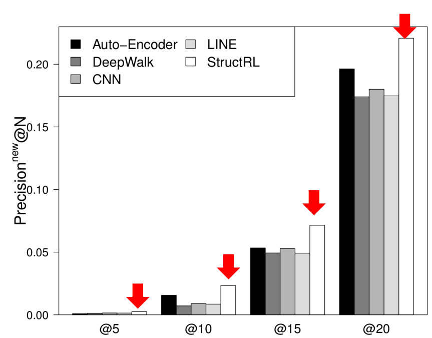
(a) Precision@N with New York dataset



(b) Precision^{New}@N with New York dataset



(c) Precision@N with Tokyo dataset



(d) Precision^{New}@N with Tokyo dataset

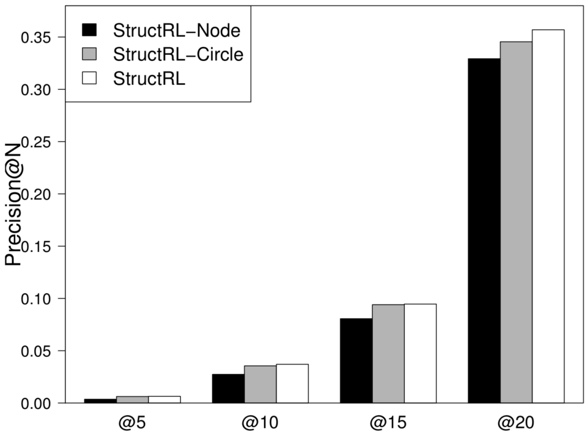
- Our model achieves the best performances on user profiling
- Substructures in a graph are essential for user behavior patterns

- **Data**
  - Mobile activity checkin data of NYC and Tokyo

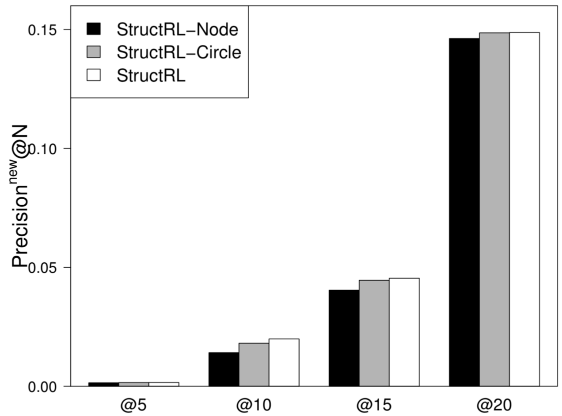
City	# Check-ins	# POI Categories	Time Period
New York	227428	251	12 April 2012 to 16 February 2013
Tokyo	573703	247	12 April 2012 to 16 February 2013

- **Evaluation Metrics**
  - The precision@N of activity category prediction
  - The precision@N of new activity recommendation
- **Baselines**
  - Autoencoder
  - DeepWalk: use truncated random walks to learn latent representations
  - LINE: preserve both local and global network structures with an edge-sampling algorithm
  - CNN: Convolutional Neural Network

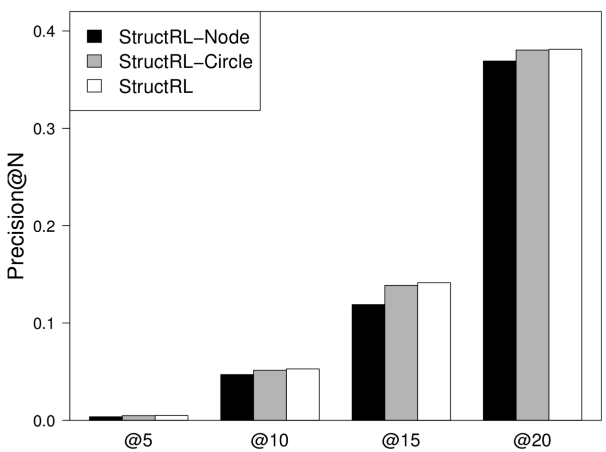
Study of Node and Circle Substructures



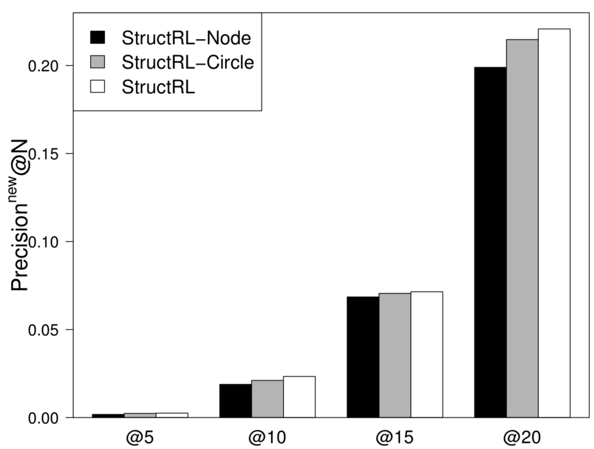
(a) Precision@N with New York dataset



(b) Precision^{New}@N with New York dataset



(c) Precision@N with Tokyo dataset



(d) Precision^{New}@N with Tokyo dataset

- **Evaluation Metrics**
  - The precision@N of activity category prediction
  - The precision@N of new activity recommendation
- **Baselines**
  - StructRL: consider node and circle substructures
  - StructRL-Node: only consider node substructures
  - StrucctRL-Circle: only consider circle substructure
- **Findings**
  - Circle substucture are more effective
  - Capturing more subgraph topologies can help

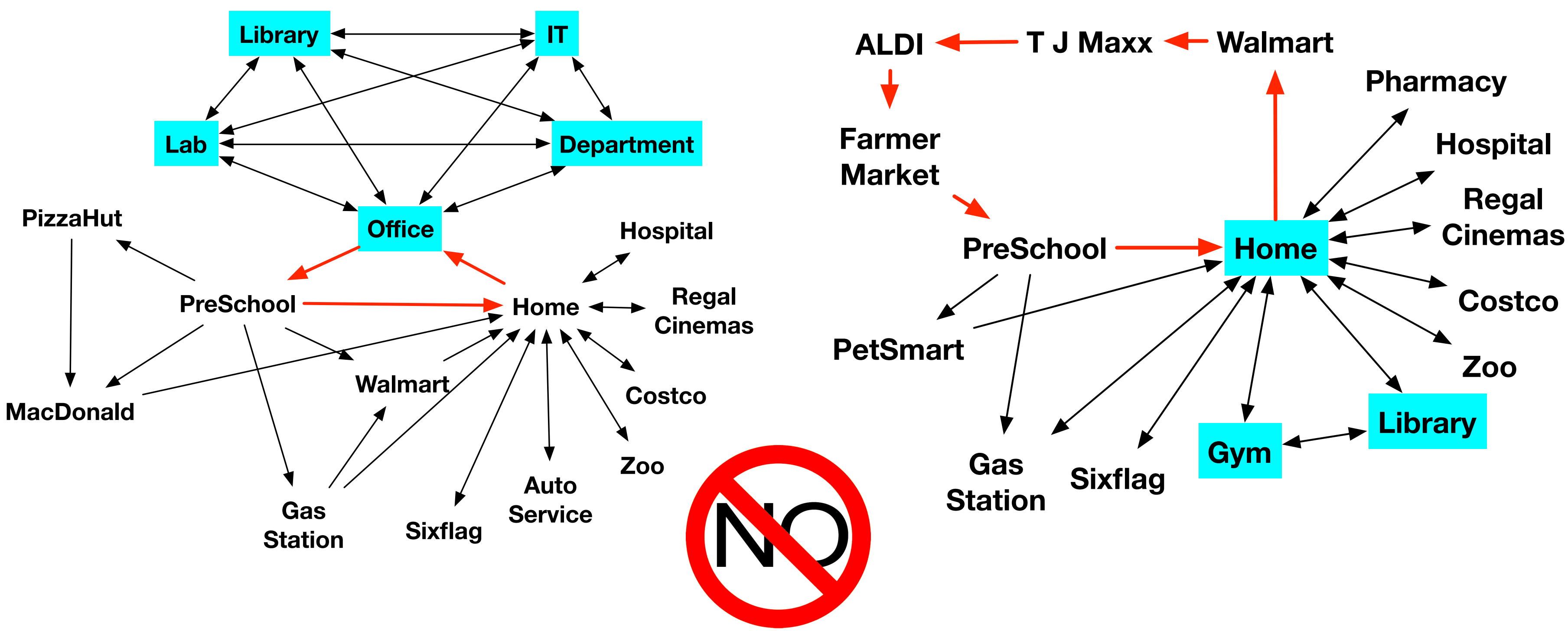
Conclusion

- Research Problem
  - Learn to profile users by both considering general interests and specific interests for certain activity types
- Method
  - Users as Activity Graphs
  - Formulate modeling specific interests as preserving substructures of user activity graphs
  - Propose an adversarial substructured learning model to integrate substructure into representation learning
- Take Away Messages
  - Adversarial learning plays the role of regularization
  - Substructure is very important for quantifying user behavior patterns
  - Pre-train neural networks to approximate undifferentiable algorithms
  - Circle is more effective than independent vertexes for profiling users

Thanks!

Questions?

Will The Traditional Solution Work?



Topologies, contents, locations of subgraphs will dynamically change over users

Will The Traditional Solution Work?

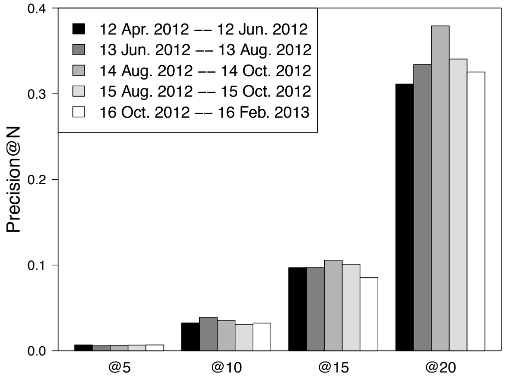
0	0	1	1
0	0	1	1
0	0	0	0
0	0	0	0

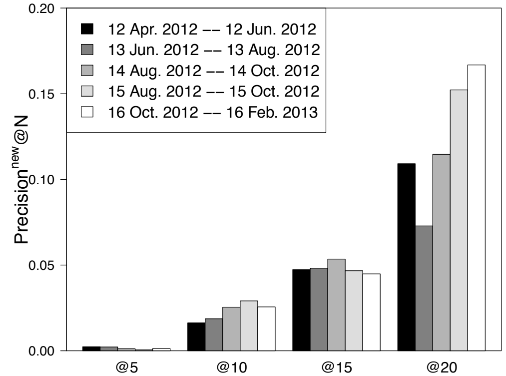
0	0	0	0
0	0	0	0
0	1	1	0
0	1	1	0

Example: Dynamic binary indicator of subgraphs in the activity matrix/graphs of two users

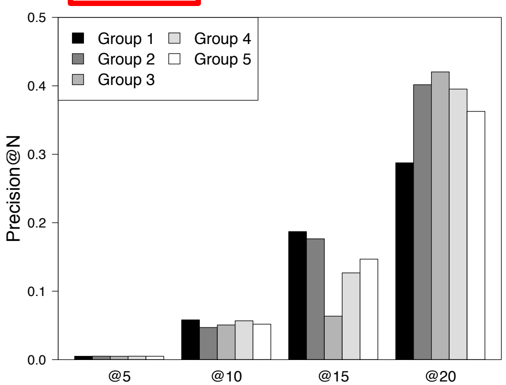
Robustness Check

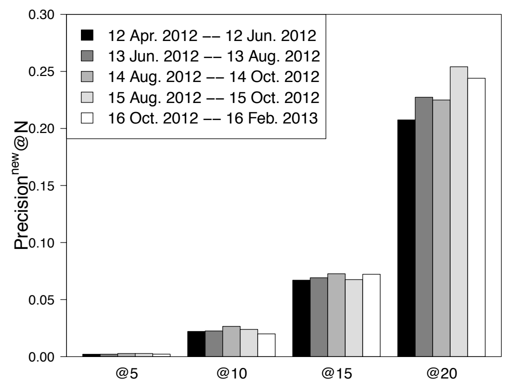
New York





Tokyo





- **Five Periods**
  - 12 Apr. 2012 – 12 Jun. 2012
  - 13 Jun. 2012 – 13 Aug. 2012
  - 14 Aug. 2012 – 14 Oct. 2012
  - 15 Aug. 2012 – 15 Oct. 2012
  - 16 Oct. 2012 – 16 Feb. 2013
- **Prediction**
  - set the last day's activities of each time period as a predictive target

- The performances of our method can achieve a small variance and are relatively stable, especially when K is small.