# Cash-out User Detection based on Attributed Heterogeneous Information Network with a Hierarchical Attention Mechanism

Binbin Hu[1], Zhiqiang Zhang[2], Chuan Shi[1], Jun Zhou[2], Xiaolong Li[2], Yuan Qi[2]

[1]Beijing University of Posts and Telecommunicatios

[2]AI Department, Ant Financial Service Group

## Background

### Credit Payment Services

Offline credit card services in commercial banks and online credit payments in internet financial institutions

### Cash-out Fraud

Pursue cash gains with illegal or insincere means,

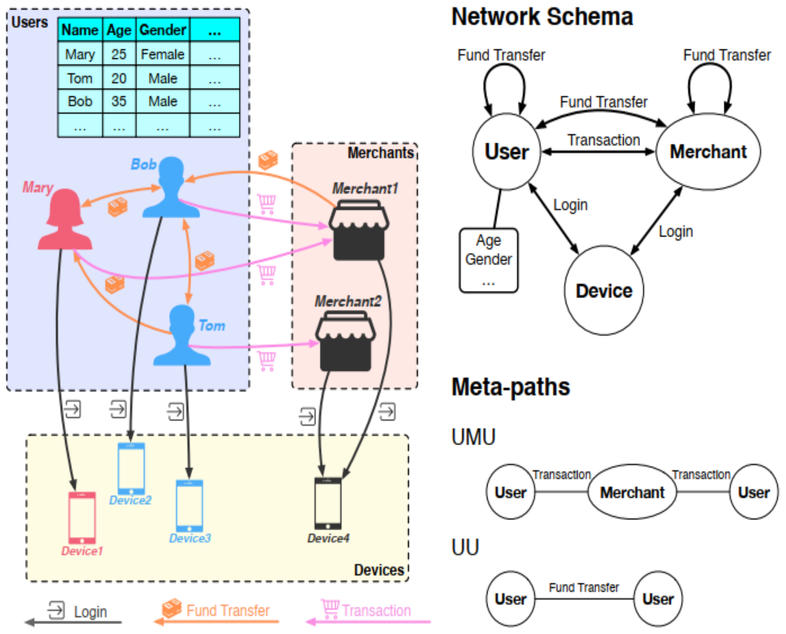
e.g., through buying pre-paid cards or other goods then reselling them.

### Cash-out User Detection

Predict whether a user will do cash-out transactions or not in the future

### Rich Interaction Relations in the Scenario of Credit Payment Service

- Fund transfer relation among users
- Login relation between users and devices
- Transaction relation between users and merchants.




(a) Scenario of credit payment service (b) Network schema and meta-path examples

### Traditional Methods

- Feature engineer + Classifier
- Mainly based on the statistical features of a certain user
- seldom fully exploit the interaction relations between users

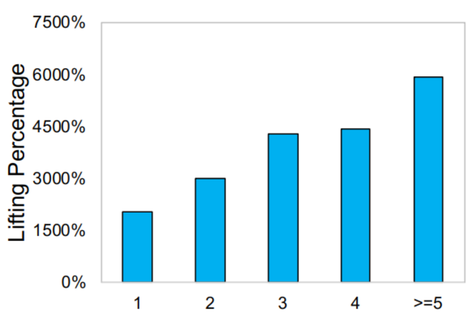
### Attribute Heterogeneous Information Network (AHIN)

- Contain attributes and multiple types of nodes and relations
- Meta-path : Semantic path between two nodes

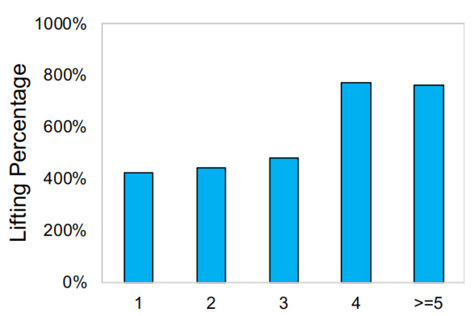
### Our work

- First to study the cash-out users detection problem
- Model the cash-out user detection problem as a classification problem in AHIN
- Propose a novel model HACUD with meta-path based neighbor and hierarchical attention mechanism

## HACUD : The Proposed Model




(a) UMU

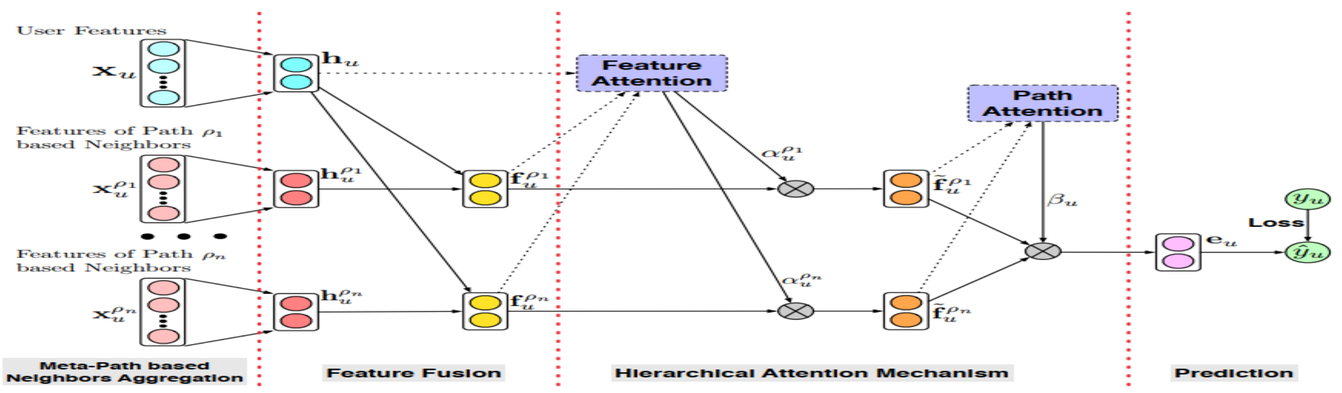



(b) UU

The lifting percentages of cash-out rate in users with different amount of cash-out neighbors against users without any cash-out neighbor in two meta-paths.

### Observations in Real Data

- Users with higher cash-out rate tend to have more cash-out neighbors
  - **Meta-path based Neighbors**
- Different meta-path based neighbors have different impacts on users
  - **Hierarchical Attention Mechanism**




**Meta-path based Neighbors Aggregation**

$$\mathbf{x}_u^\rho = \sum_{j \in \mathcal{N}_u^\rho} w_{uj}^\rho * \mathbf{x}_j$$

**Feature Fusion**

$$\mathbf{h}_u = \mathbf{W}\mathbf{x}_u + \mathbf{b}, \quad \mathbf{h}_u^\rho = \mathbf{W}^\rho \mathbf{x}_u^\rho + \mathbf{b}^\rho$$

$$\mathbf{f}_u^\rho = \mathrm{ReLU}(\mathbf{W}_F^\rho g(\mathbf{h}_u, \mathbf{h}_u^\rho) + \mathbf{b}_F^\rho)$$

**Hierarchical Attention**

Feature Attention:

$$\boldsymbol{v}_u^\rho = \mathrm{ReLU}(\mathbf{W}_f^1[\mathbf{h}_u; \mathbf{f}_u^\rho] + \mathbf{b}_f^1),$$

$$\boldsymbol{\alpha}_u^\rho = \mathrm{ReLU}(\mathbf{W}_f^2 \boldsymbol{v}_u^\rho + \mathbf{b}_f^2),$$

Path Attention:

$$\beta_{u,\rho} = \frac{\exp(\mathbf{z}^{\rho\mathrm{T}} \cdot \widetilde{\mathbf{f}}_u^C)}{\sum_{\rho' \in \mathcal{P}} \exp(\mathbf{z}^{\rho'\mathrm{T}} \cdot \widetilde{\mathbf{f}}_u^C)},$$

## Experiments

### Dataset

- **Ten Days Dataset**

  1.99 million users (2018/03/21~2018/03/31)

- **One Month Dataset**

  5.16 million users (2018/03/01~2018/03/31)

### AHIN

- #User : 56.75 millions
- #Merchants : 0.51millions
- #Transfer relation : 77.40 millions
- #Transaction : 20.64 millions
- #Attribute : 123

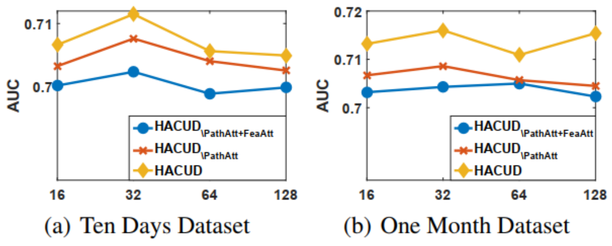
### Metric

$$AUC = \frac{\sum_{u \in \mathcal{U}^+} rank_u - \frac{|\mathcal{U}^+| \times (|\mathcal{U}^+|+1)}{2}}{|\mathcal{U}^+| \times |\mathcal{U}^-|}$$

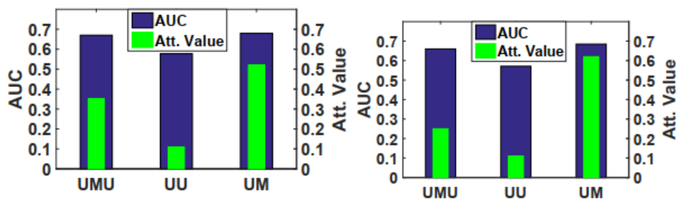
### Performance Comparison

| Algorithm | AUC | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Ten Days Dataset | | | | One Month Dataset | | | |
| | $d=16$ | $d=32$ | $d=64$ | $d=128$ | $d=16$ | $d=32$ | $d=64$ | $d=128$ |
| Node2vec | 0.5893 | 0.5913 | 0.5926 | 0.5930 | 0.5980 | 0.6063 | 0.6009 | 0.6021 |
| Metapath2vec | 0.5914 | 0.5903 | 0.5917 | 0.5920 | 0.6005 | 0.5976 | 0.5995 | 0.5983 |
| Node2vec + Feature | 0.6455 | 0.6464 | 0.6510 | 0.6447 | 0.6541 | 0.6561 | 0.6607 | 0.6518 |
| Metapath2vec + Feature | 0.6456 | 0.6429 | 0.6469 | 0.6485 | 0.6550 | 0.6552 | 0.6523 | 0.6545 |
| Structure2vec | 0.6537 | 0.6556 | 0.6598 | 0.6545 | 0.6641 | 0.6632 | 0.6657 | 0.6678 |
| GBDT | 0.6389 | 0.6389 | 0.6389 | 0.6389 | 0.6467 | 0.6467 | 0.6467 | 0.6467 |
| $\mathrm{GBDT}_{Struct}$ | 0.6948 | 0.6948 | 0.6948 | 0.6948 | 0.6968 | 0.6968 | 0.6968 | 0.6968 |
| HACUD | **0.7066** | **0.7115** | **0.7056** | **0.7049** | **0.7132** | **0.7160** | **0.7109** | **0.7154** |

### Effects of Hierarchical Attention

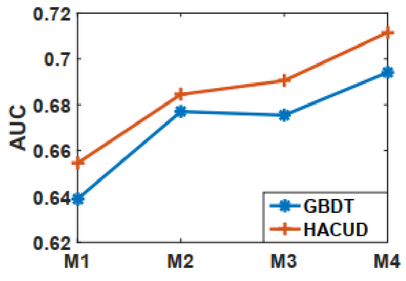



(a) Ten Days Dataset (b) One Month Dataset




(a) Ten Days Dataset (b) One Month Dataset

### Impact of Different Meta-paths




### Parameter Tuning