LinkRadar: Assisting the Analysis of Inter-app Page Links via Transfer Learning

Diandian Gu, Ziniu Hu, Shangchen Du, Yun Ma

App Ecosystem
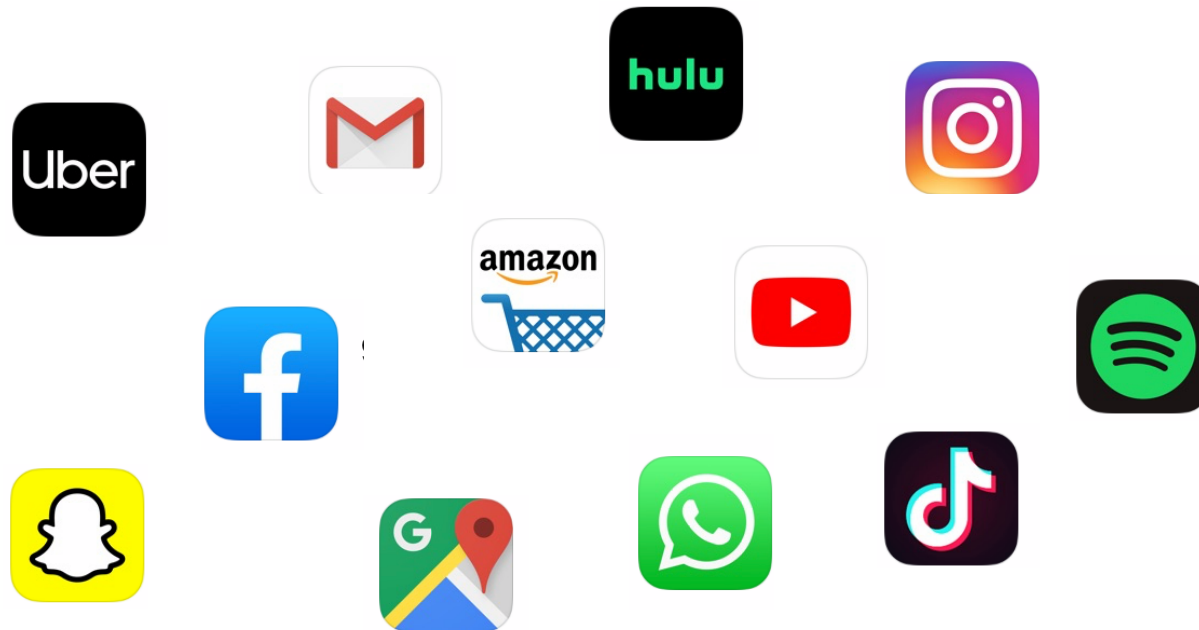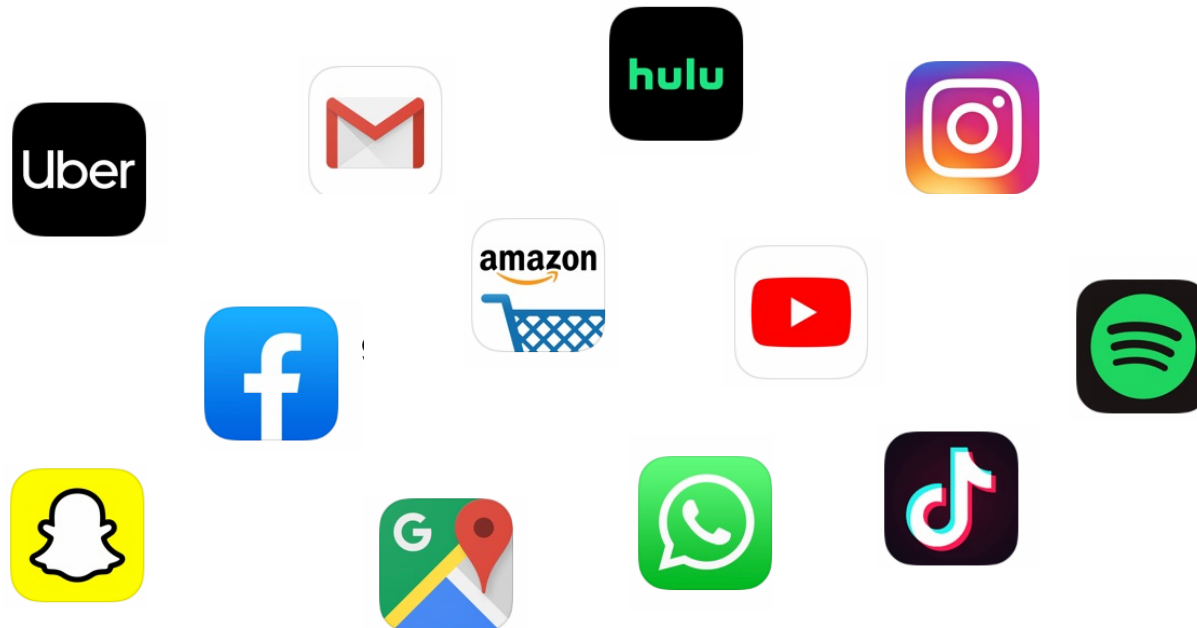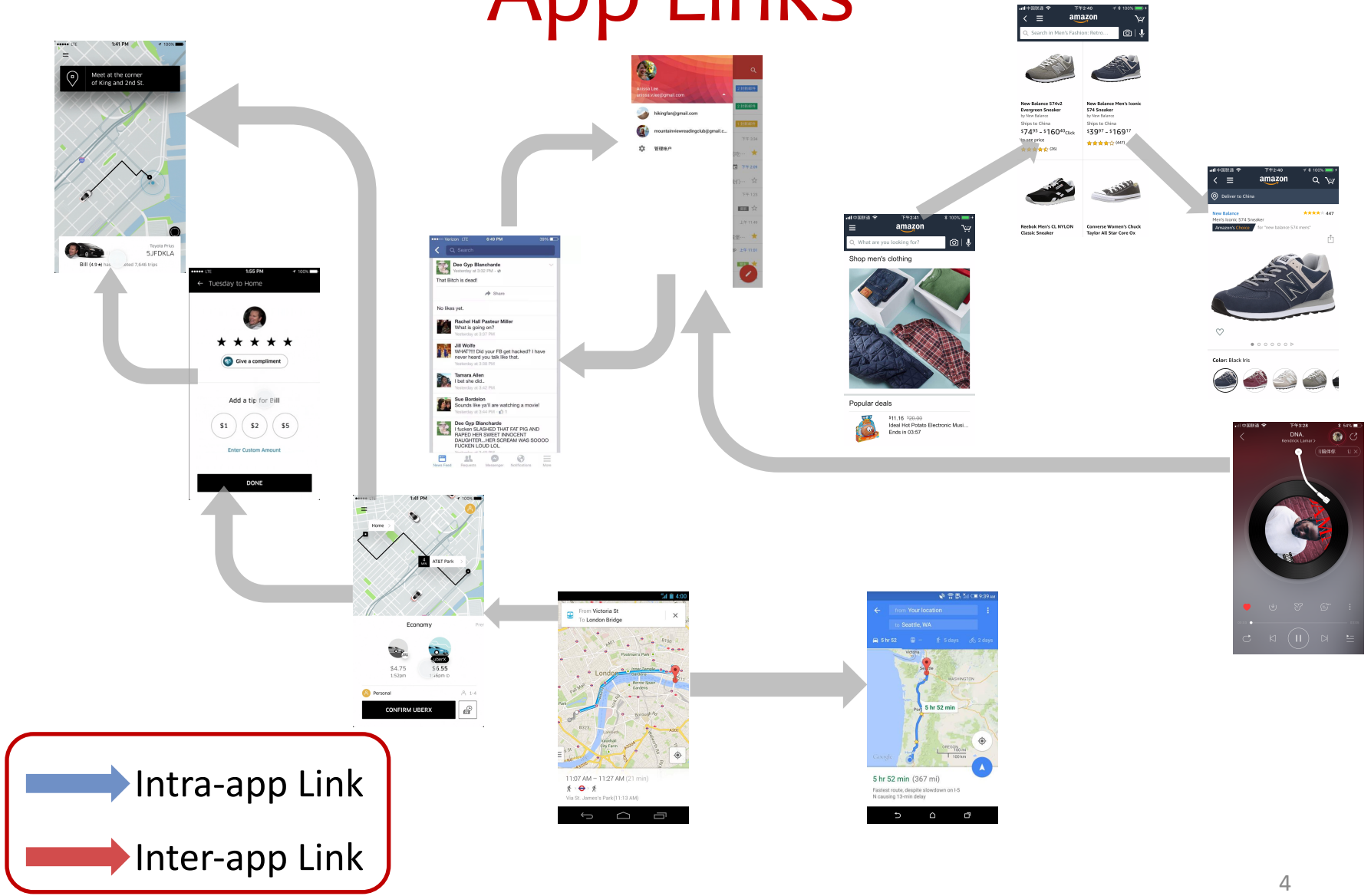
- Mobile applications have become the major medium of information retrieval

App Pages

- App pages are linked with one another to provide services for mobile users

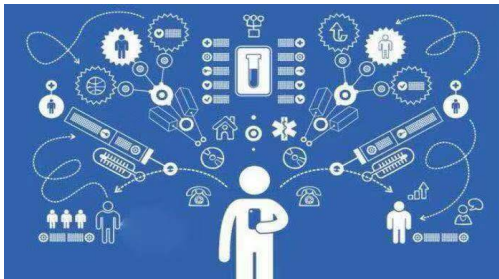# App Links

Intra-app Link

Inter-app Link

4

Why Inter-App Link Analysis?

- A critical component in many fields

Recommender Systems

Sensitive Data Theft
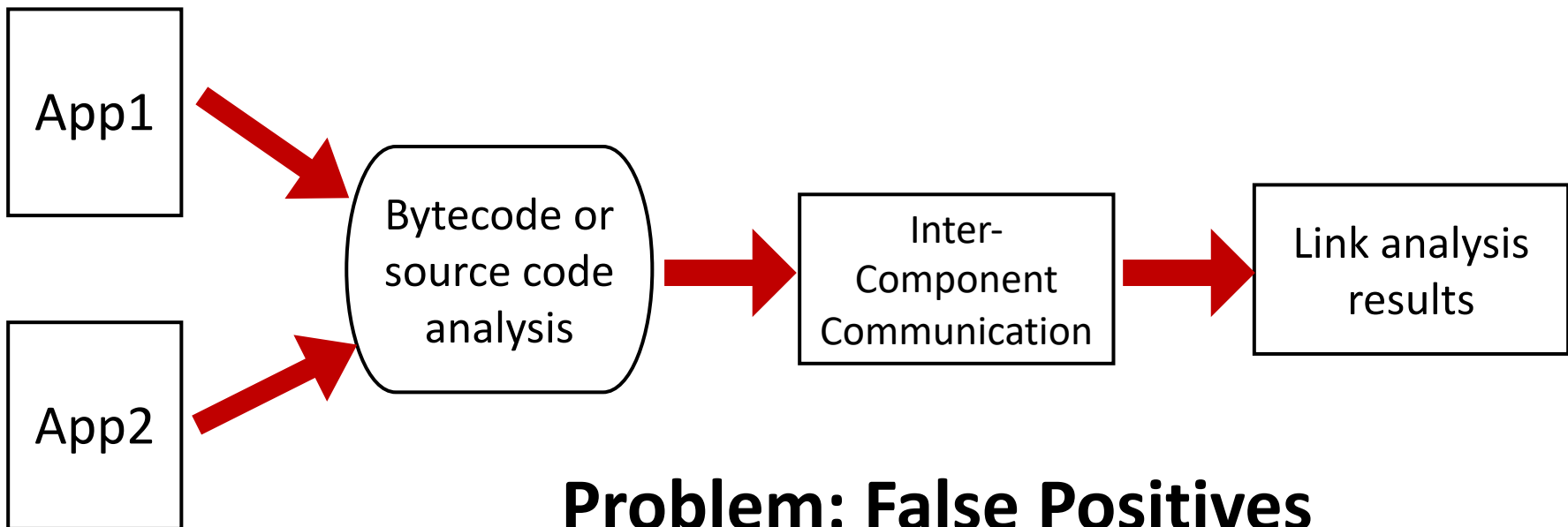
- **Service recommendation**
- **Malicious data access detection**
- **……**

Current Inter-App Link Analysis work

- Static analysis of mobile applications

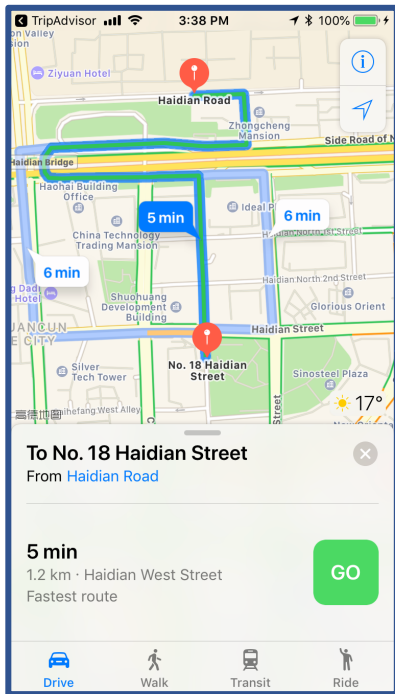

Problem: False Positives

Significant manual effort is required to verify the links!
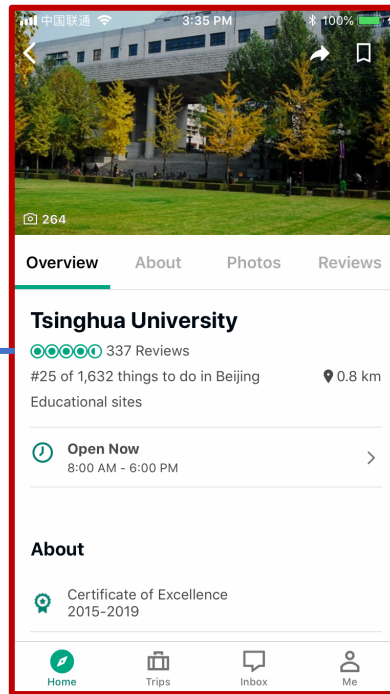
Our Method
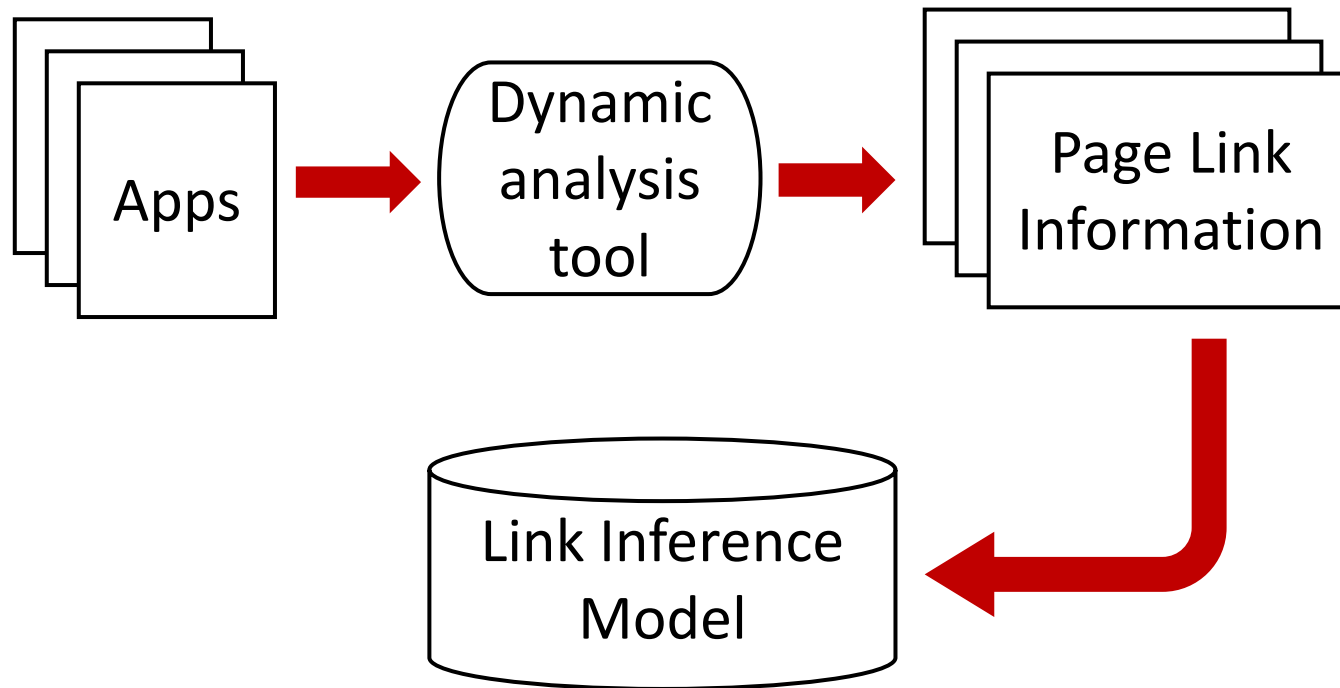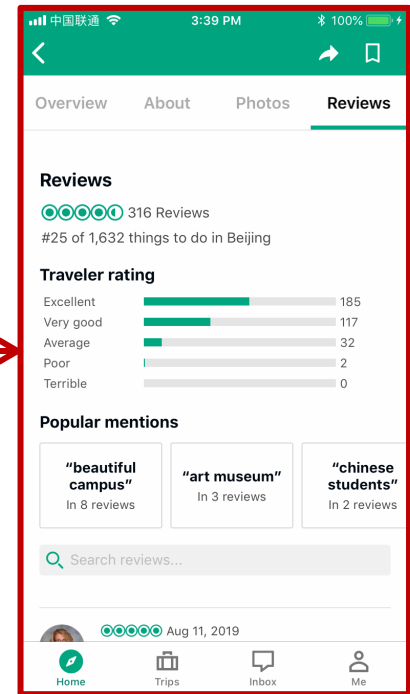

Challenge: few inter-app links in the data set

Motivating Example

Route

Detail of a place

Comment

The Transfer Learning Method

Source Labels

Source Model

Source Data

Large amount of data (intra-app links)

Transfer Learned Knowledge

Target Labels

Target Model

Target Data

Small amount of data (inter-app links)

Large Amount of Page Pairs with Intra-app Links

Small Amount of Page Pairs with Inter-app Links

Labels

Labels

Page Embedding Model

Page Embedding Model

Frozen

Transfer Learned Knowledge as Initialized Weight

Page Embedding

Page Embedding

Page Embedding

Page Embedding

Link Inferencer

Link Inferencer

Fine Tuned

Cross Entropy Loss

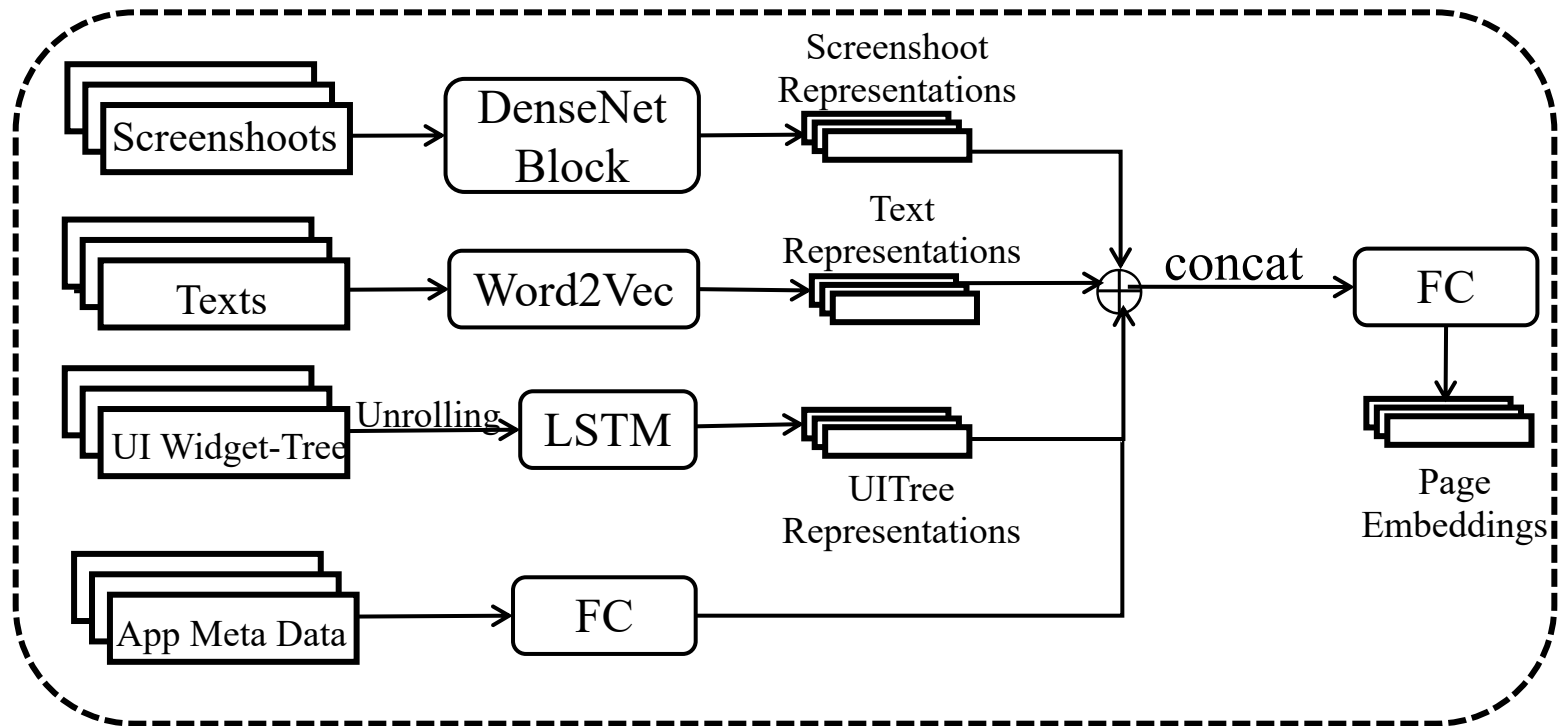Cross Entropy Loss

The Structure of Feature Extractor

Evaluation Methods

- Precision and recall

- Compare the inference result of LinkRadar with manually validated result

Baseline Methods

- **Random**
  - Random numbers from 0 to 1
- **LinkRadar-NoF**
  - The LinkRadar model without fine tuning
- **LinkRadar-NoN**
  - The LinkRadar Model without neighborhood information

Results

	Precision	Recall	MAP	MRR
LinkRadar	**0.71**	**1.00**	**0.257**	**0.412**
Random	0.07	0.44	0.089	0.103
LinkRadar-NoF	0.33	0.88	0.206	0.253
LinkRadar-NoN	0.32	0.95	0.227	0.280

- LinkRadar outperforms all three baselines

Take Away

- App link analysis for service recommendation and data theft detection

- Transfer learning method and Representation learning method for link inference

- Actionable insights for other link analysis works that also suffer from large false positives

Q&A