Mining Meaningful Knowledge from User Behavior: Network-based Approach

Chanyoung Park

Assistant Professor

Industrial & Systems Engineering

KAIST

Outline

Part 1: Research Motivation & Background

Part 2: **Multi-modal** User Behavior Analysis

Part 3: **Multi-aspect** User Behavior Analysis

Part 4: Vision for the future

Outline

Part 1: Research Motivation & Background ✓

Part 2: **Multi-modal** User Behavior Analysis

Part 3: **Multi-aspect** User Behavior Analysis

Part 4: Vision for the future

User Behaviors in E-Commerce

“… ***35 percent of what consumers purchase on Amazon*** *and* ***75 percent of what they watch on Netflix*** *come from product recommendations …”*

https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers

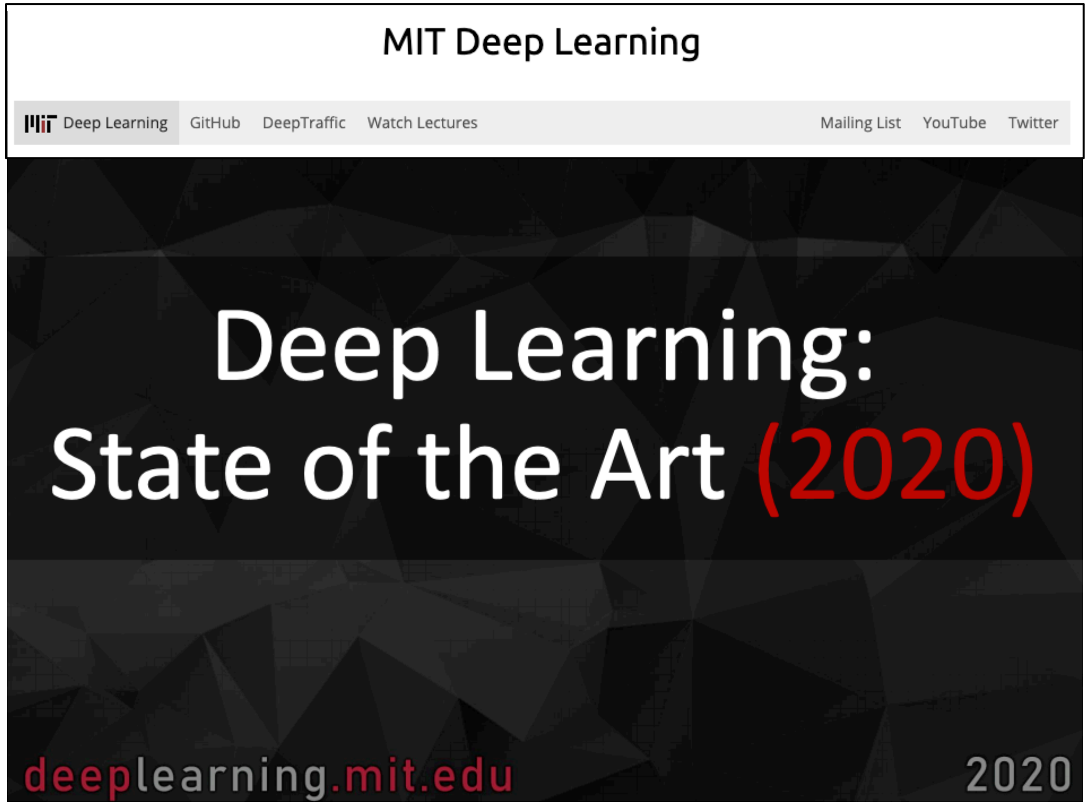
How Netflix’s AI Recommendation Engine Helps It Save $1 Billion A Year

On Wednesday, Aug 7 2019, by Vishnu Subramanian

Over the last decade, Netflix has slowly grown into the world’s most popular subscription-based video streaming service, offering a wide selection of films and TV series including several “Netflix Originals” produced by the company themselves in-house. Netflix has over 150 million subscribers worldwide, a testament to the company’s cross-cultural popularity and market dominance in several countries around the world. While this popularity can be attributed to Netflix’s pioneering model, an affordable subscription fee and top-notch content/programming, Netflix is also known for using techniques from Artificial Intelligence to maintain its market dominance. Chief among these is the Netflix Recommendation Engine, a tool that is reportedly worth over $1 Billion per year to the company in indirect cost savings.

Importance of user behavior analysis

https://artelliq.com/blog/how-netflix-s-ai-recommendation-engine-helps-it-save-1-billion-a-year/



Deep Learning State of the Art (2020)
Lex Fridman
MIT
Date/Time: Mon, Jan 6, 3-4:30pm
Room: E54-100

https://www.youtube.com/watch?v=0VH1Lim8gL8

Recommendation System

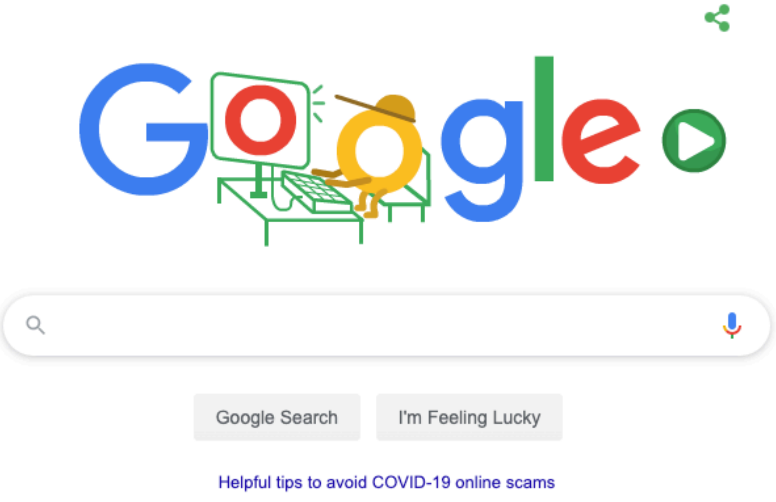
"Recommendation system is the **most important** in terms of **impact part of AI systems**..."

"...the **most powerful AI space for the next a couple of decades** is recommendation systems. They are going to have **the biggest impact on our society** because they **affect the information we see, how we learn, what we think, how we communicate**. **These algorithms are controlling us**..."

Importance of user behavior analysis

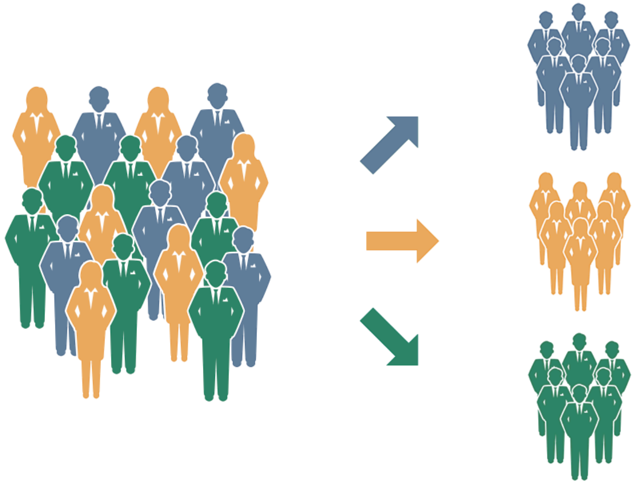
Other Applications

Search engine



Ad retargeting/remarketing

User segmentation



https://www.wordstream.com/google-remarketing

Collecting More Data

More Data → Better Performance

Computer Vision (Image)

Natural Language Processing (Text)

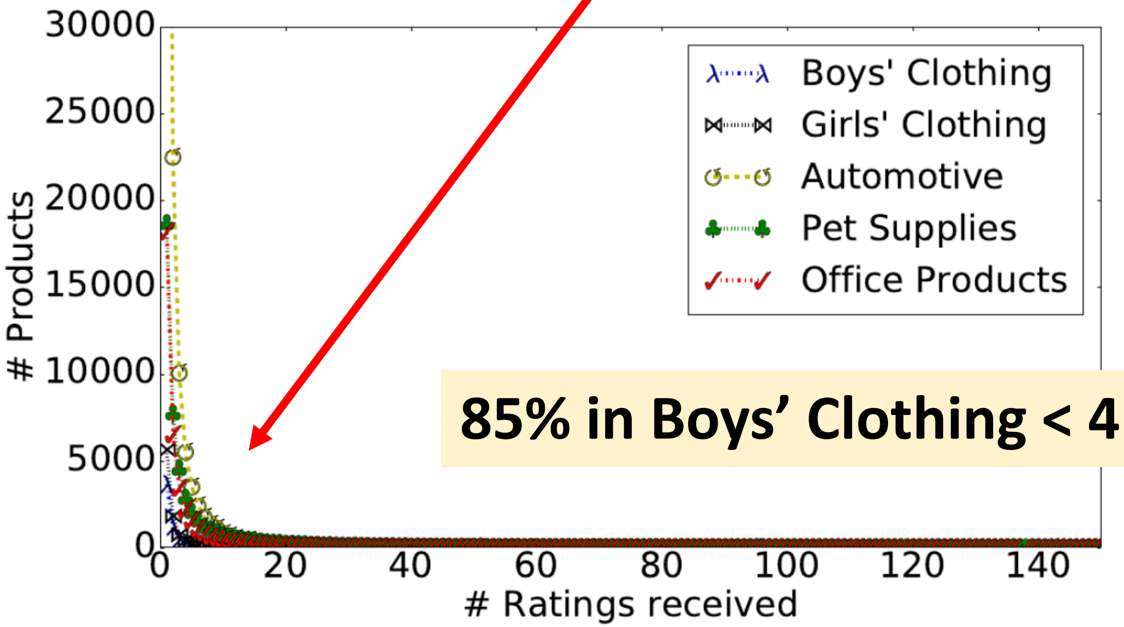
Collecting Data in User Behavior Analysis

User explicit feedback requires user engagement

⬇

Hard to collect!

Most items receive **few** feedback from users



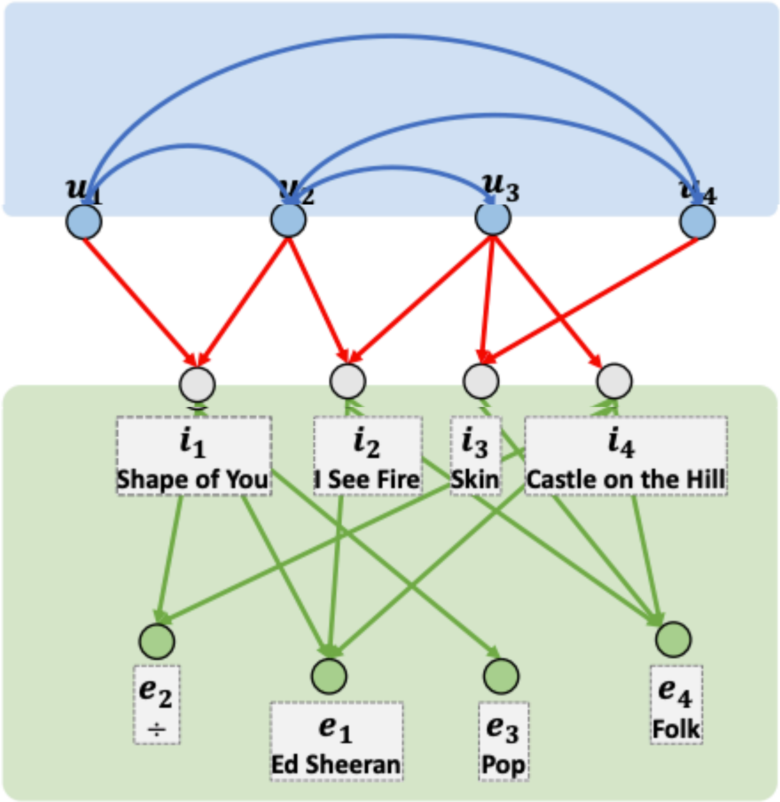
85% in Boys' Clothing < 4 ratings

Data sparsity!

Collecting Data in User Behavior Analysis

We need to make use of
user implicit feedback or auxiliary data

How can we represent user behavior in the real-world?

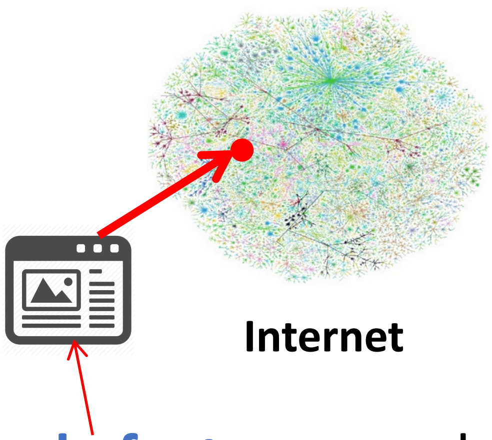


User-User Connections
- Social Relations
- Same Profiles ...

User-Item Interactions
- Implicit Feedback
- Explicit Feedback ...

Item-Item Connections
- Same Attributes
- External Knowledge ...

- Many types of data can be flexibly formulated as **networks**
  - e.g., Internet, biological network, chemical network, network of neurons



Internet

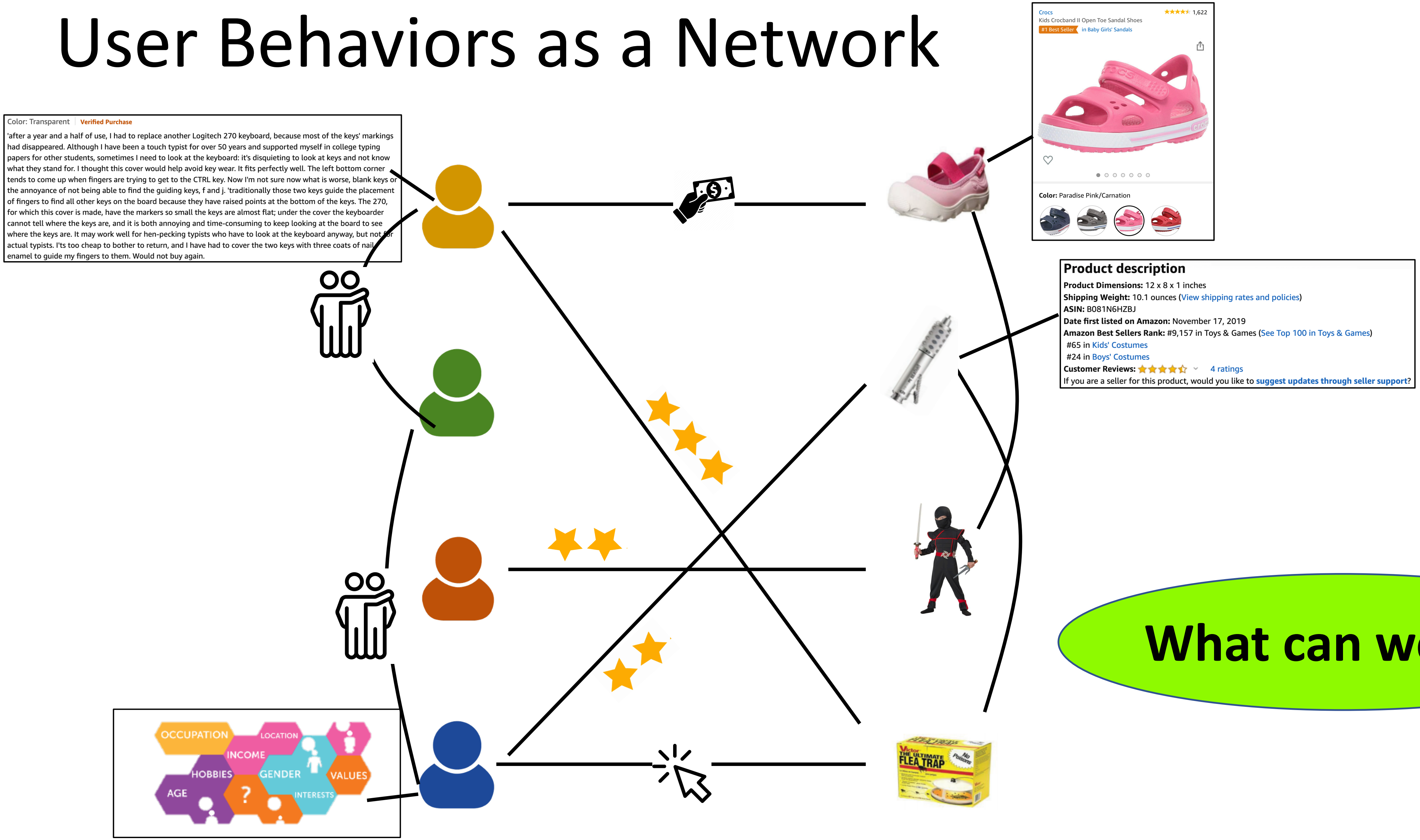
Chemical network

- **Node features** can be easily incorporated
  - Text or image as features of nodes

Network best represents complex user behavior

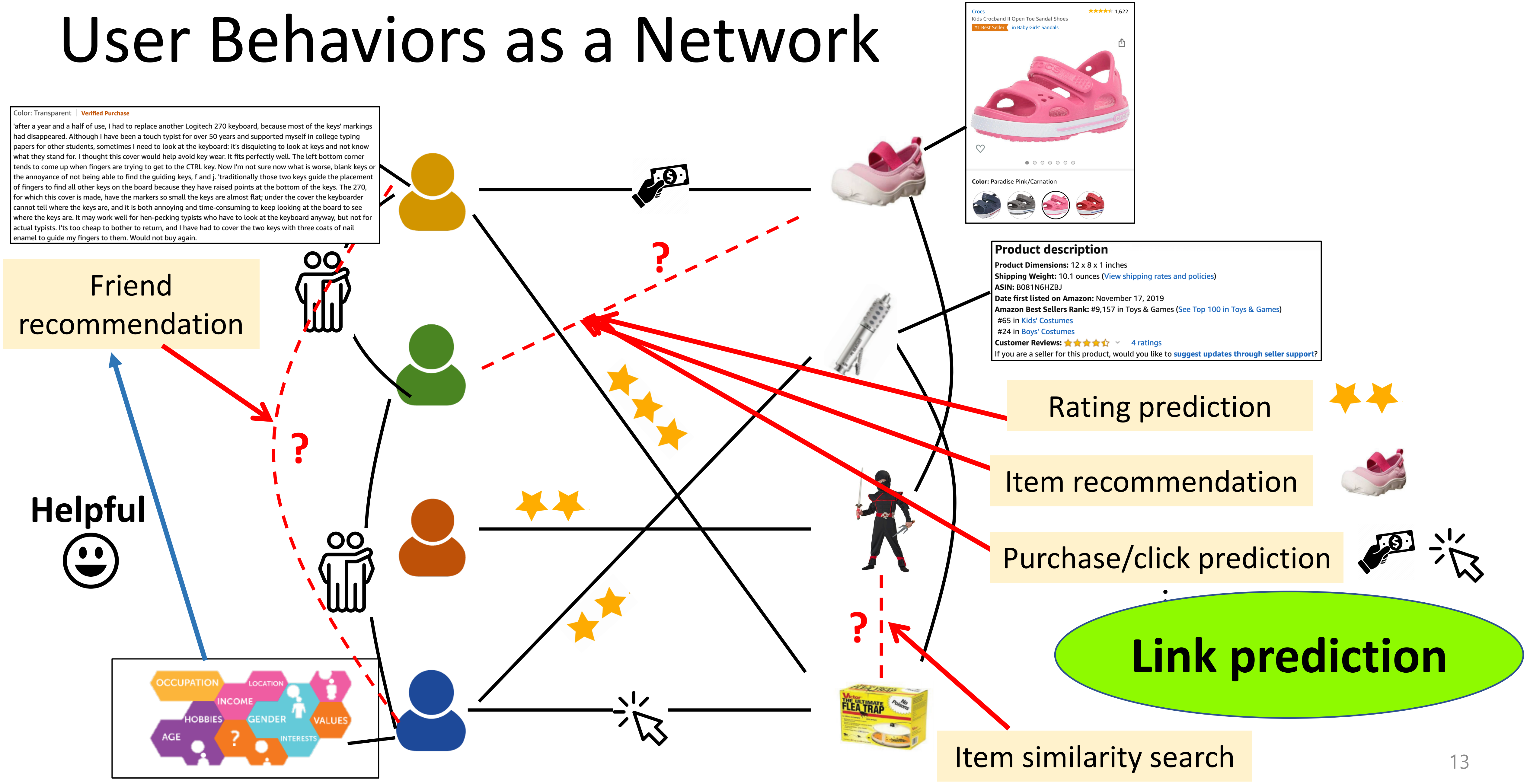
Figure credit: (Xiang Wang et al, 2019) Learning and Reasoning on Graph for Recommendation, CIKM 2019 Tutorial

User Behaviors as a Network

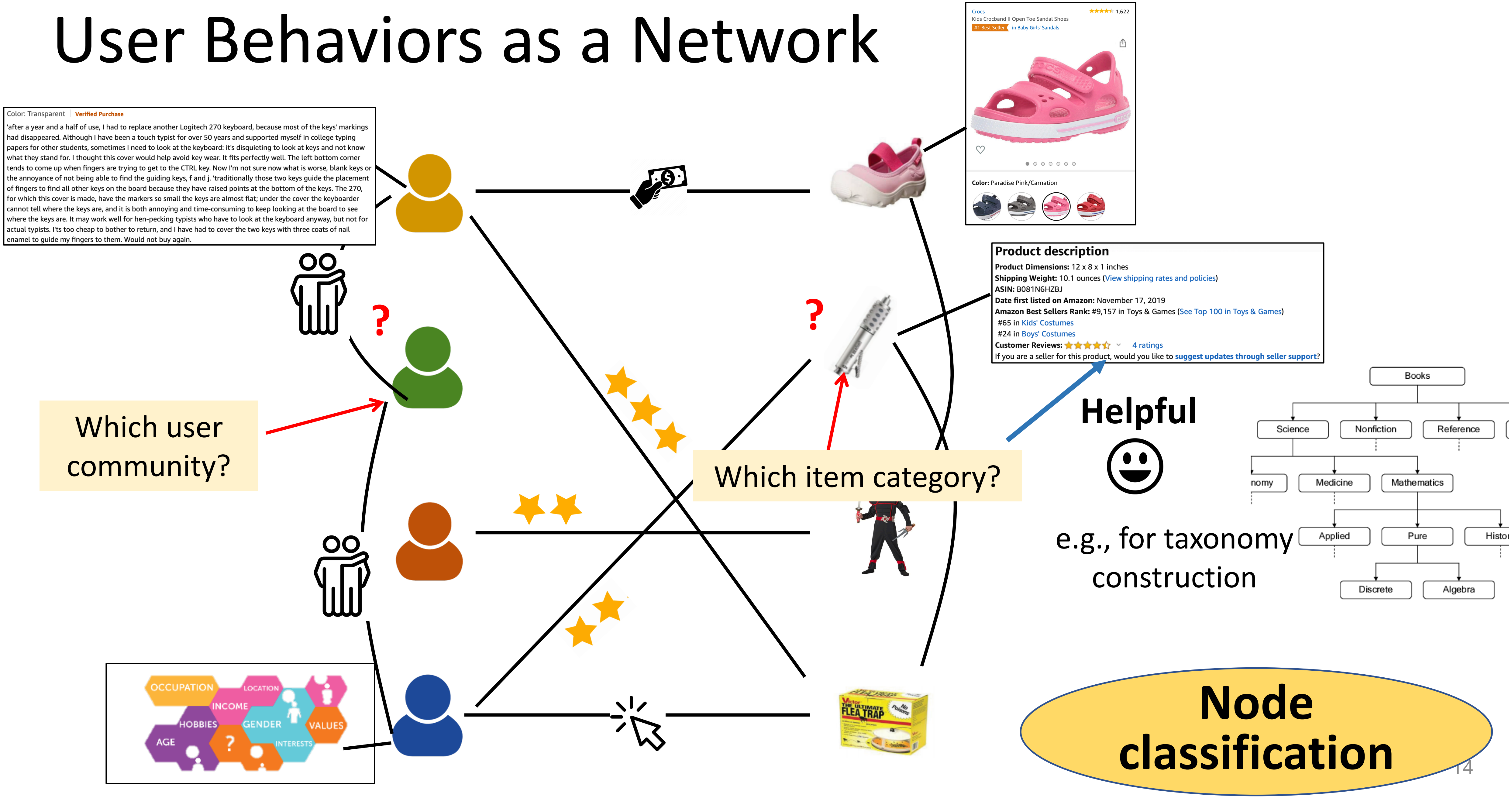


What can we do?

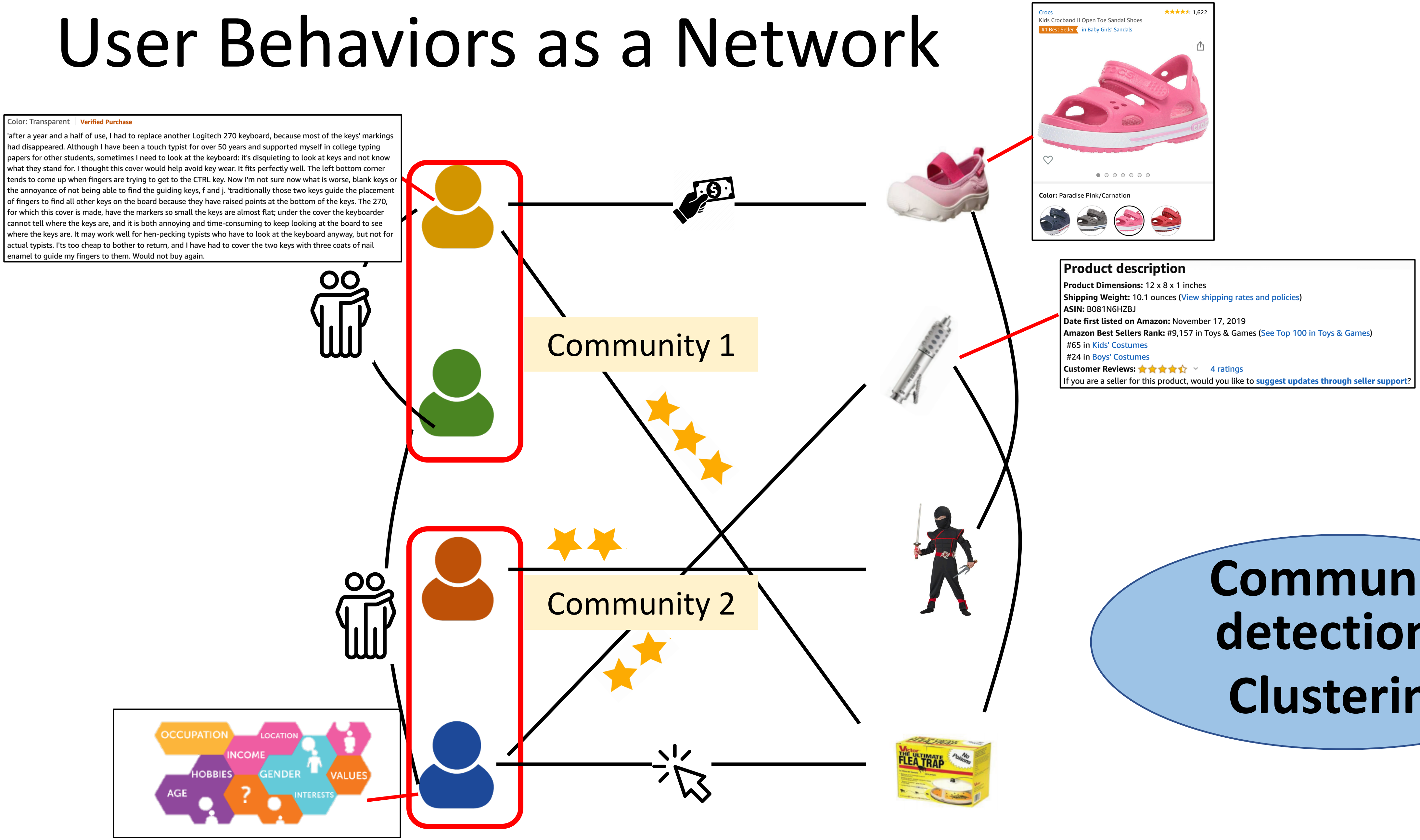
User Behaviors as a Network



User Behaviors as a Network



User Behaviors as a Network



Community detection / Clustering

User Behaviors as a Network

Understanding user behavior network

→

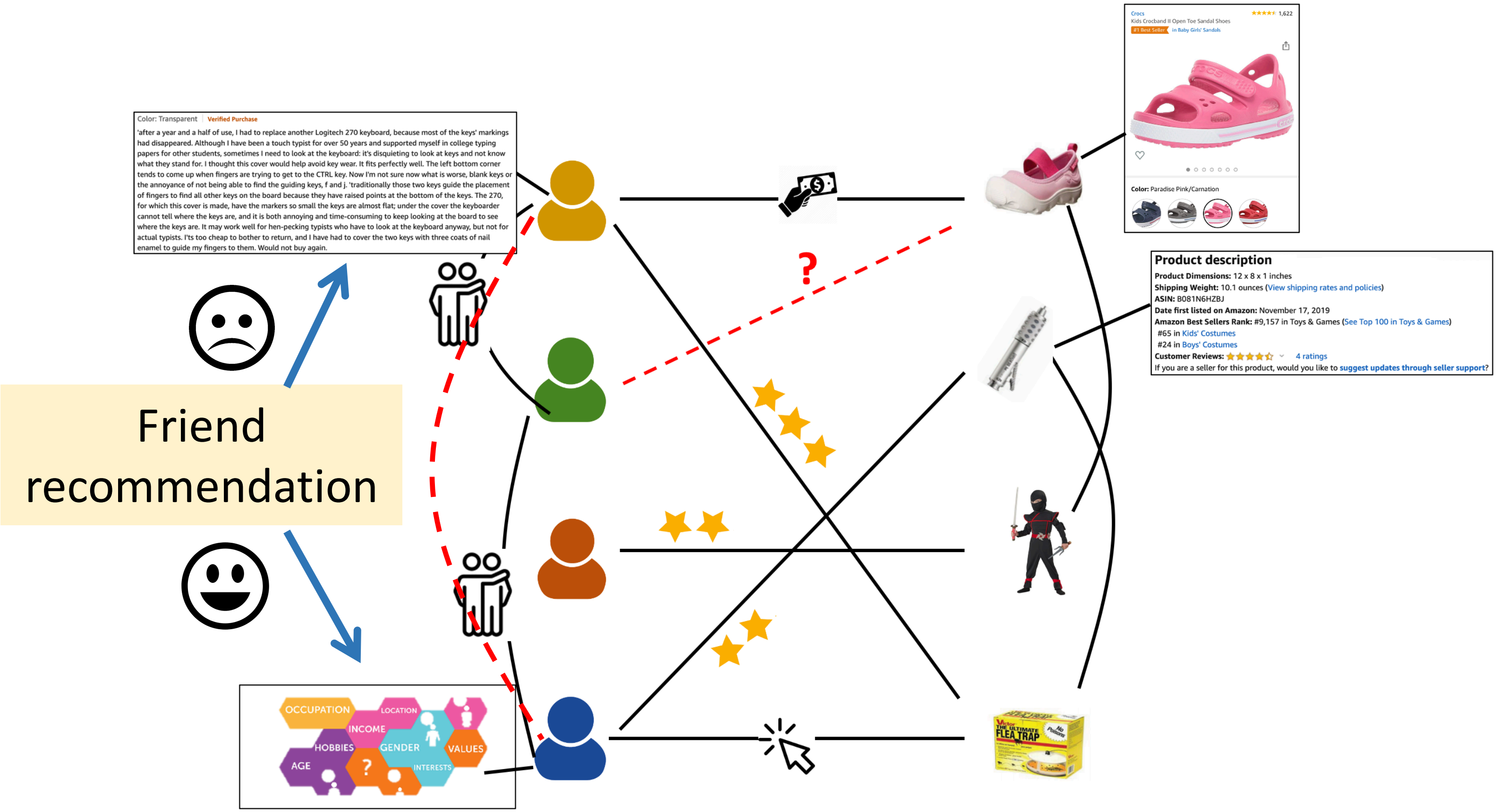
The key for accurate user behavior analysis

=

Learning **representations for nodes** and understanding **relationship between nodes** in user behavior network

Challenges of User Behavior Analysis

1. User behavior data is **multi-modal**



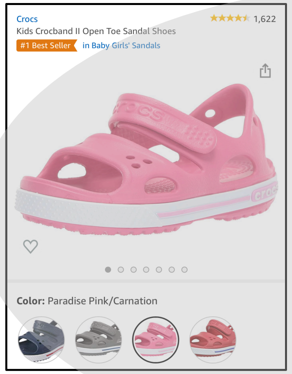
- Click
- Purchase
- Friend
- Item image
- Item description text
- User reviews
- User demographics
- Also-viewed, also-bought

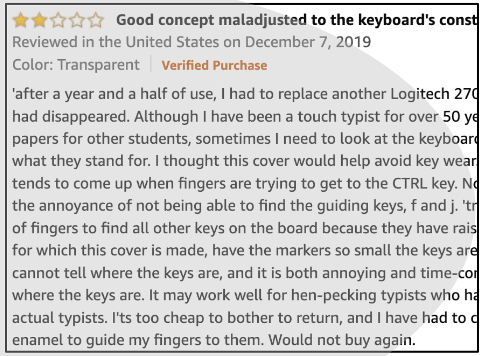
…

How to extract knowledge from multi-modal user behavior data?

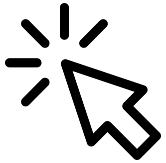
vs. Modeling data in other domains?

	Computer vision	Natural language processing (NLP)	User behavior analysis
Data type			

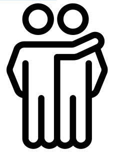
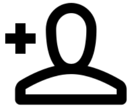



Good concept maladjusted to the keyboard's const
Reviewed in the United States on December 7, 2019
Color: Transparent | Verified Purchase
'after a year and a half of use, I had to replace another Logitech 270
had disappeared. Although I have been a touch typist for over 50 ye
papers for other students, sometimes I need to look at the keyboar
what they stand for. I thought this cover would help avoid key wear
tends to come up when fingers are trying to get to the CTRL key. No
the annoyance of not being able to find the guiding keys, f and j. 'tr
of fingers to find all other keys on the board because they have rais
for which this cover is made, have the markers so small the keys are
cannot tell where the keys are, and it is both annoying and time-cor
where the keys are. It may work well for hen-pecking typists who ha
actual typists. I'ts too cheap to bother to return, and I have had to c
enamel to guide my fingers to them. Would not buy again.

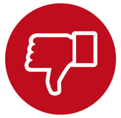
Various sources


Click

Add-to-cart

Purchase

Friend

Follow

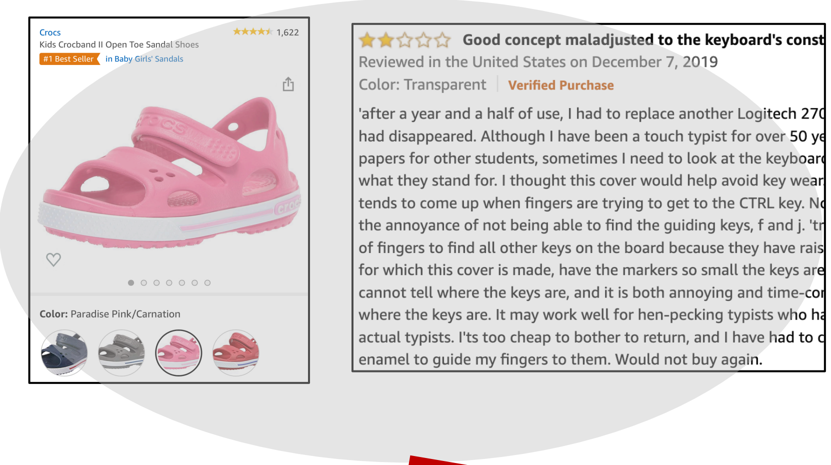

Rating

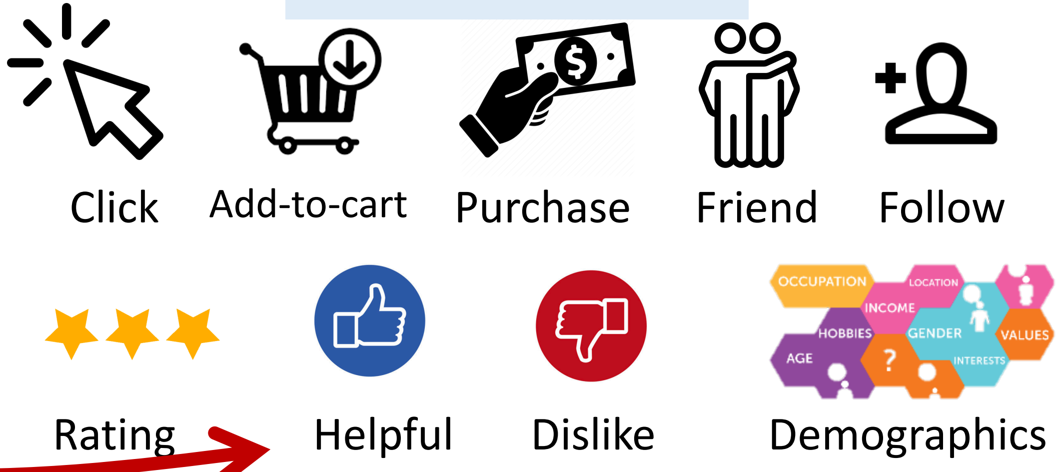
Helpful

Dislike


Demographics

vs. Modeling data in other domains?

	Computer vision	Natural language processing (NLP)	User behavior analysis
Data type			Various sources
Methodology	Pretrained deep learning model + Finetuning		Insight from data analysis + Application specific model

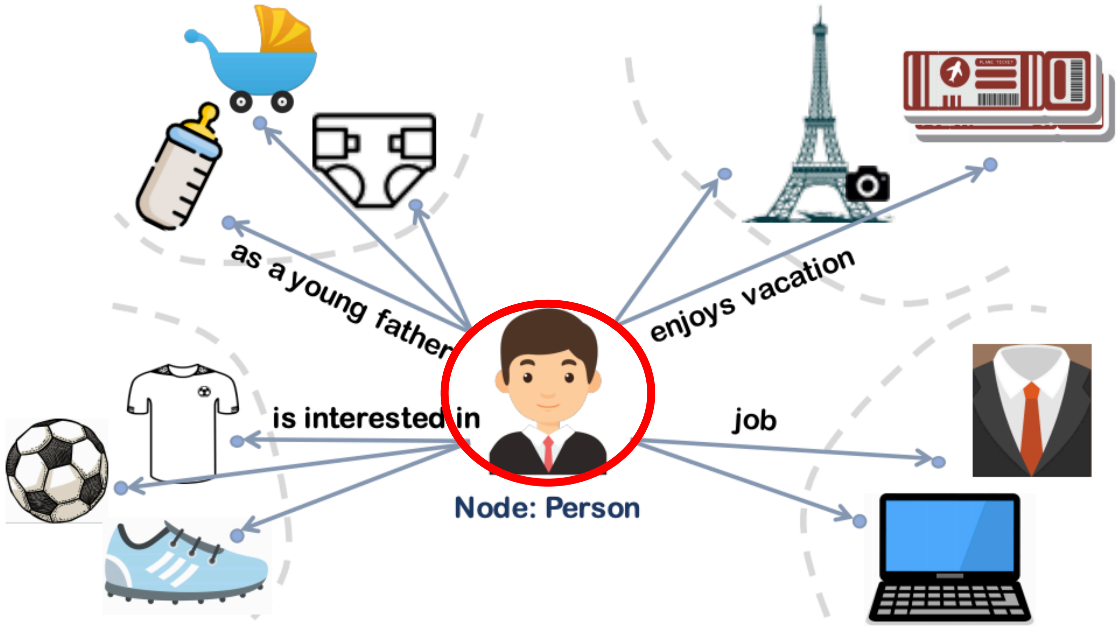




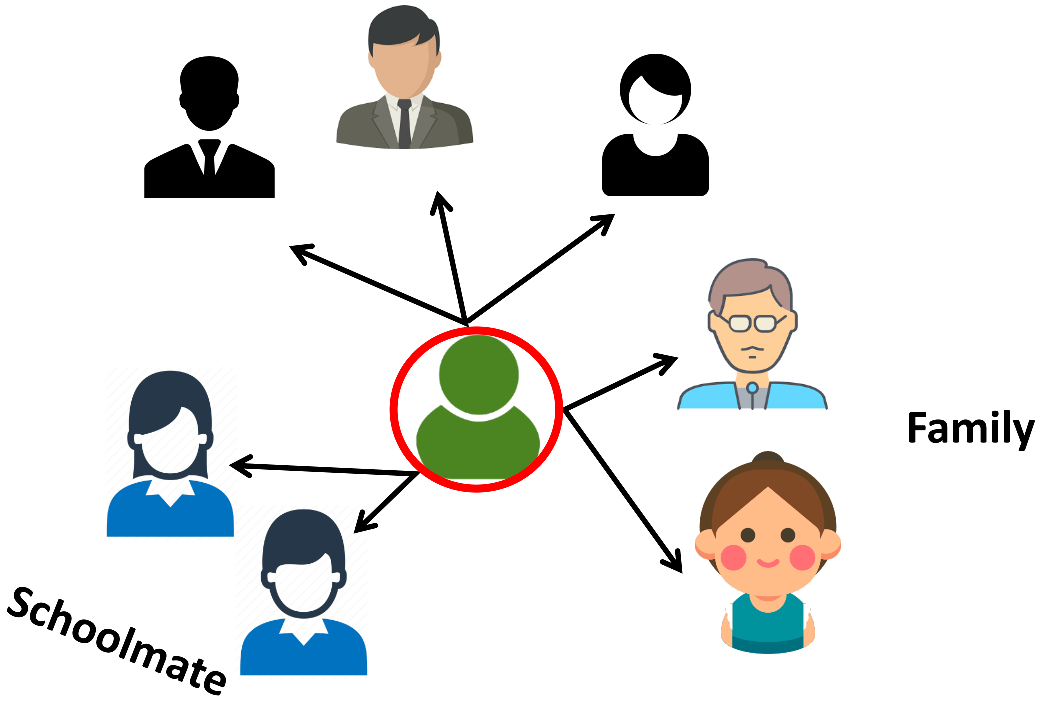
Ability to deal with various types of data is the key for user behavior analysis

Challenges of User Behavior Analysis

2. User behavior has **multi-aspect**



Purchase history



Social network

How to differentiate among multiple aspects?

Figure credit: Liu. et al, KDD19

Recap: User Behavior Analysis

Goal: To understand and extract meaningful knowledge from user behaviors

Challenge ①

User behavior data is **multi-modal**

How to extract knowledge from different modalities?

Challenge ②

User behavior has **multi-aspect**

How to differentiate multiple aspects?

Overview: User Behavior Analysis

Part2: Multi-modal

User
- Demographics [RecSys15]
- Social network [WWW16, InforSci16]
- User Review [InforSci19, SIGIR18]

Item
- Meta-data [RecSys15, ICDM18]
- Item description [Recsys16, InforSci17]
- Image [WWW17]
- Affinity relation [WWW17]

User-Item
- Rating [RecSys16, WWW17, InforSci17,19, SIGIR18,ICDM18]
- Click/Purchase/ Add-to-cart/ Add-to-favorite [IJCAI19b, InforSci20,KNOSYS20]
- Bookmark [ICDM18]
- Time information [IJCAI19a, KNOSYS20, sub_b]

Part 3: Multi-aspect

- Homogeneous Network [KDD20]
- Multiplex Network [AAAI20, KNOSYS20]
- Heterogeneous Network [ICDM18, CIKM19, KDD20, sub_c]

Outline

Part 1: Research Motivation & Background

Part 2: **Multi-modal** User Behavior Analysis

Part 3: **Multi-aspect** User Behavior Analysis

Part 4: Vision for the future

Overview: User Behavior Analysis

Part2: Multi-modal

User
- Demographics [RecSys15]
- Social network [WWW16, InforSci16]
- User Review [InforSci19, SIGIR18]

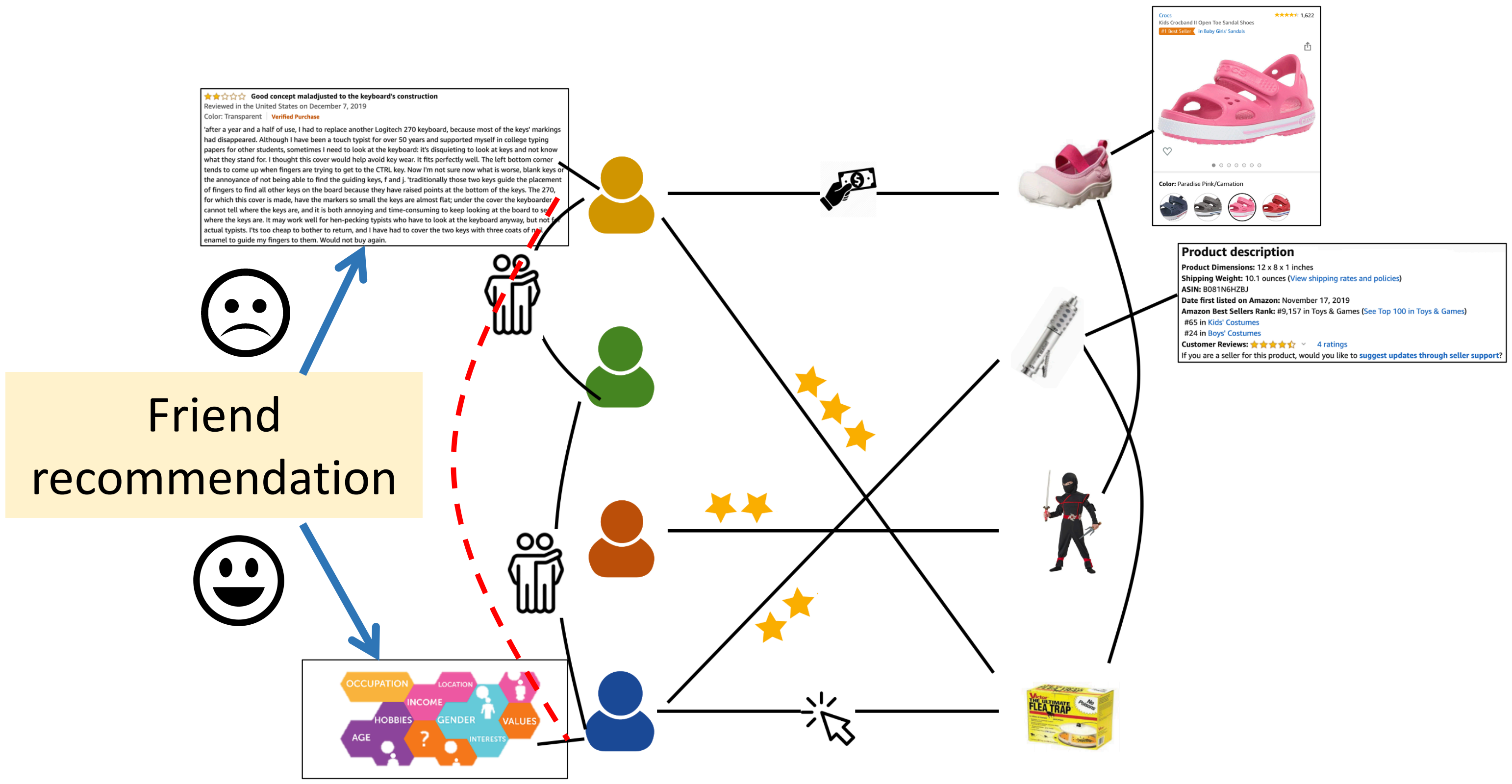
Item
- Meta-data [RecSys15, ICDM18]
- Item description [Recsys16, InforSci17]
- Image [WWW17]
- Affinity relation [WWW17]

User-Item
- Rating [RecSys16, WWW17, InforSci17,19, SIGIR18,ICDM18]
- Click/Purchase/Add-to-cart/Add-to-favorite [IJCAI19b, InforSci20,KNOSYS20]
- Bookmark [ICDM18]
- Time information [IJCAI19a, KNOSYS20, sub_b]

Part 3: Multi-aspect

- Homogeneous Network [KDD20]
- Multiplex Network [AAAI20, KNOSYS20]
- Heterogeneous Network [ICDM18, CIKM19, KDD20, sub_c]

User behavior data is **multi-modal**



- Click
- Purchase
- Friend
- Item image
- Item description text
- User reviews
- User demographics
- Also-viewed, also-bought

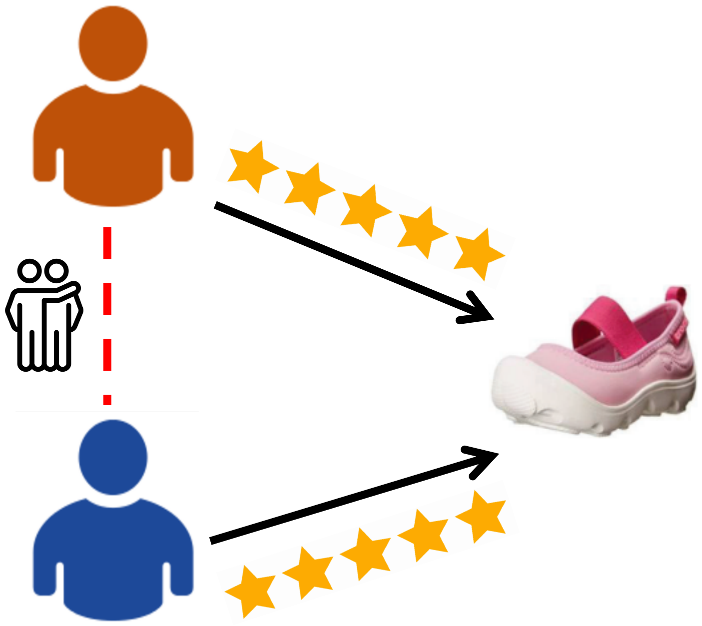
…

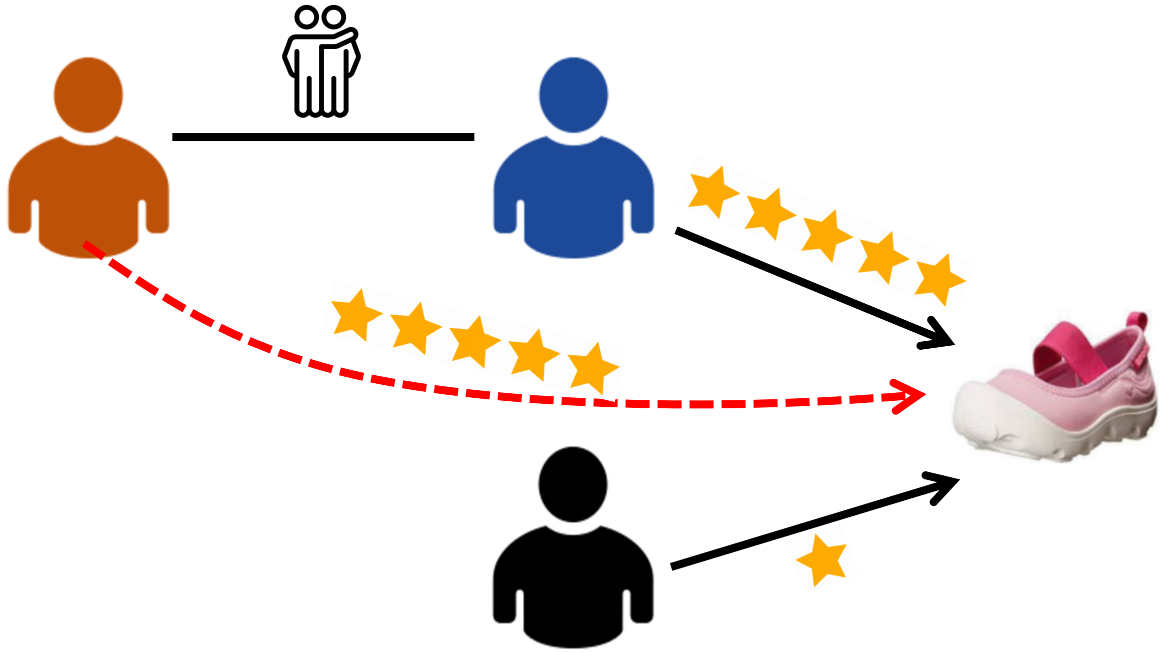
How to extract knowledge from multi-modal user interaction?

How Does User Social Network Help?

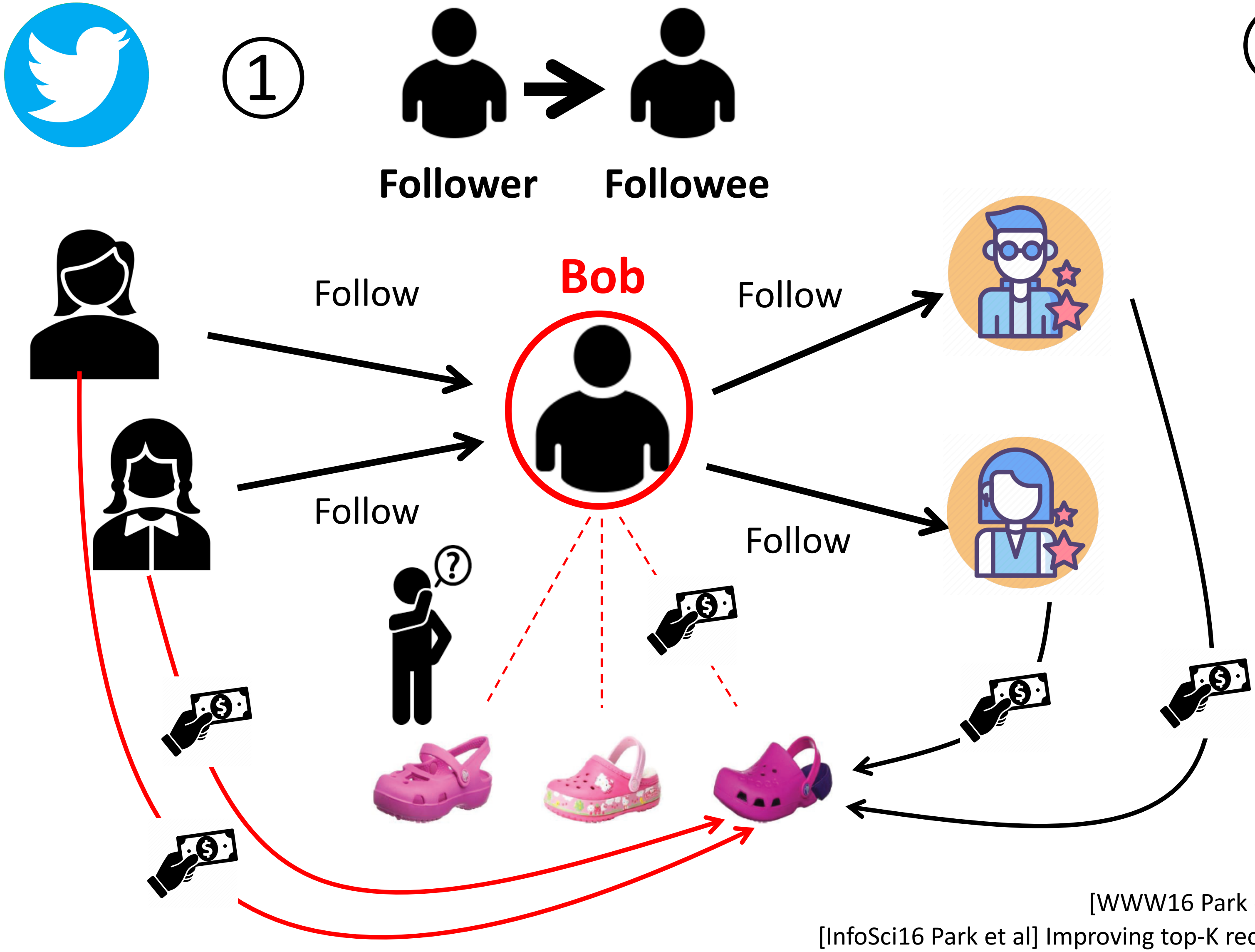
Homophily effect

Likely to be friends with users with similar interest

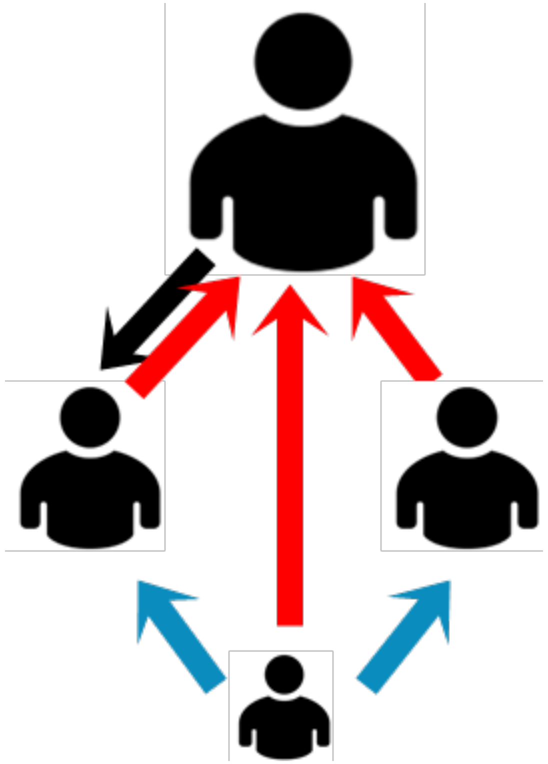




Social information of users [WWW'16, InforSci'16]



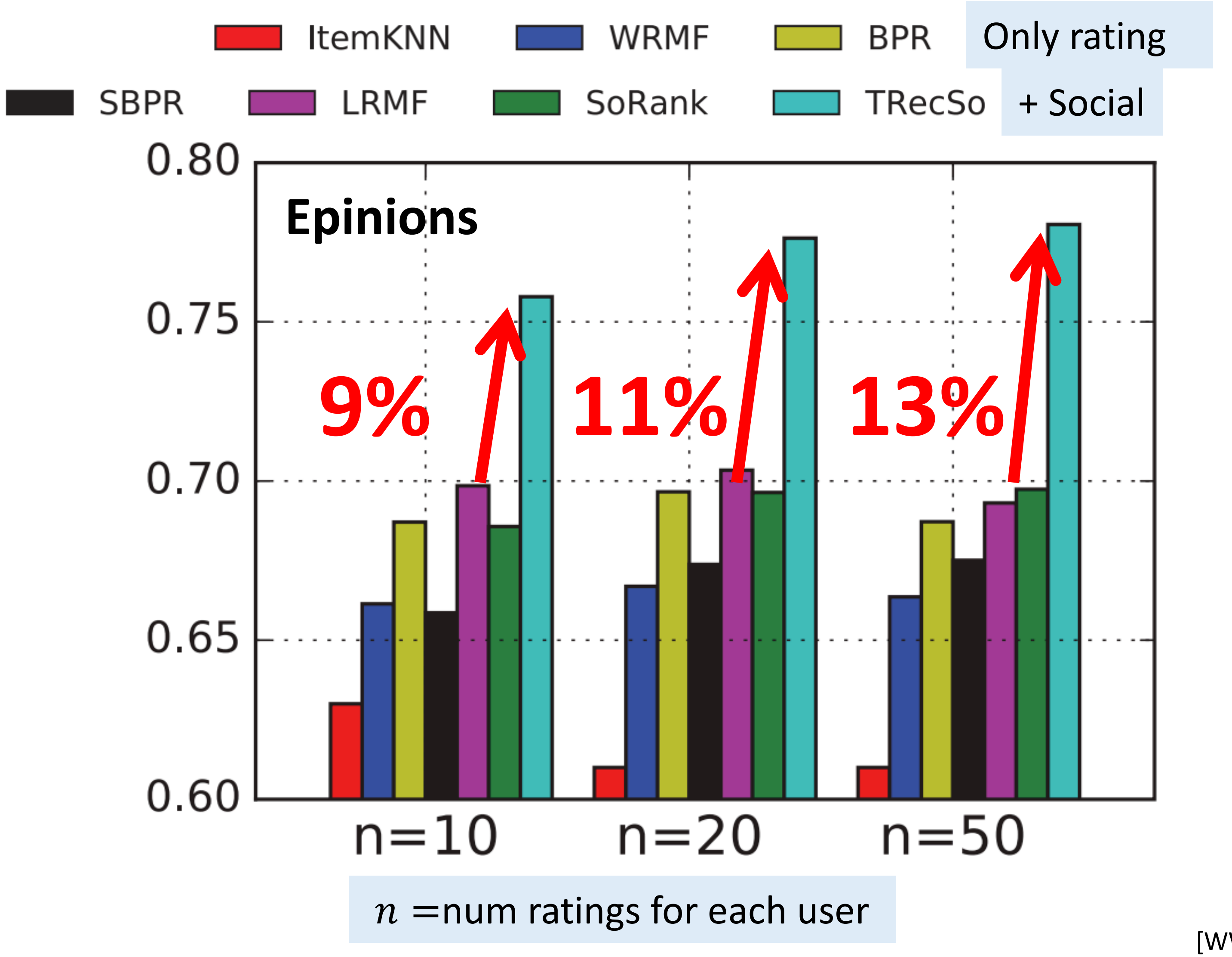
② Structural importance

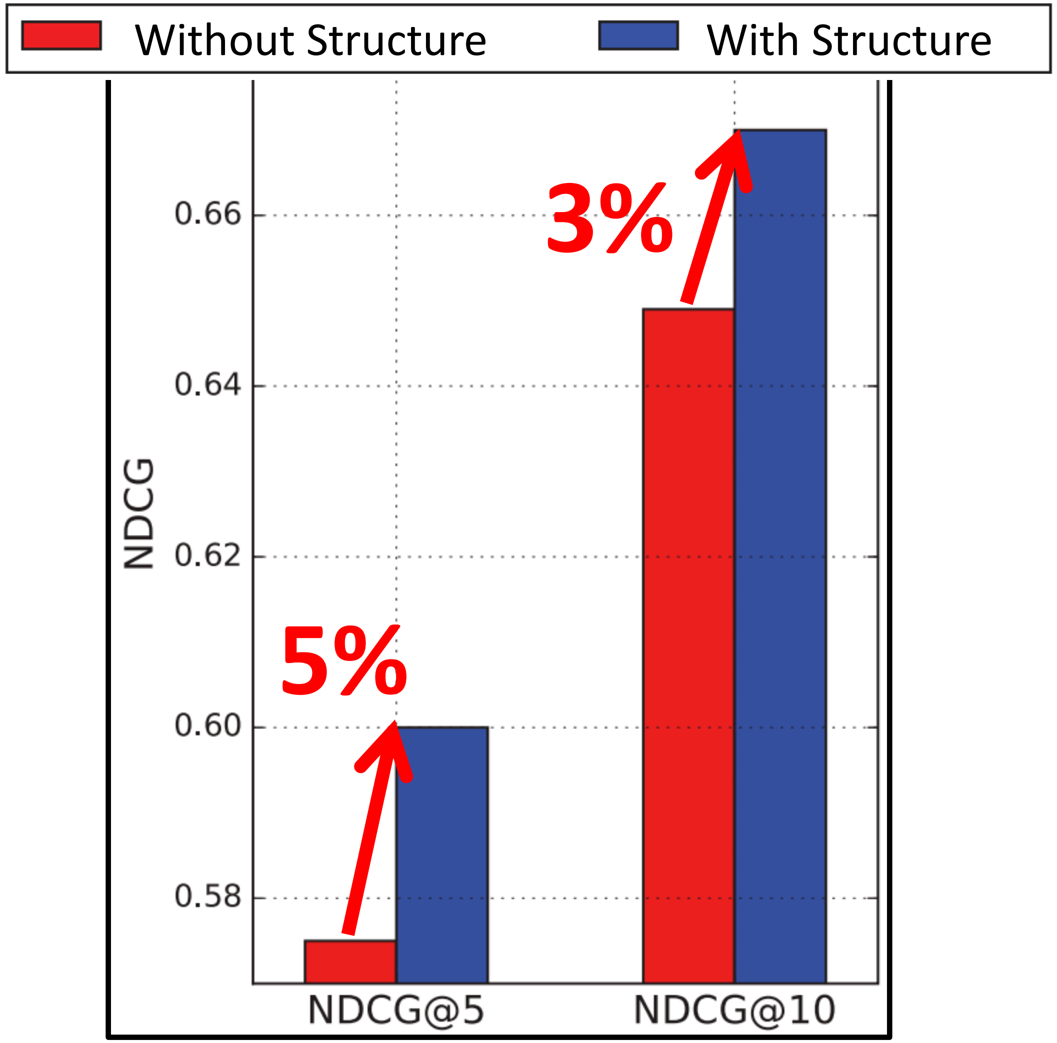


[WWW16 Park et al] TRecSo: Enhancing Top-k Recommendation With Social Information

[InfoSci16 Park et al] Improving top-K recommendation with truster and trustee relationship in user trust network

Social information of users [WWW16, InforSci16]





[WWW16 Park et al] TRecSo: Enhancing Top-k Recommendation With Social Information
[InfoSci16 Park et al] Improving top-K recommendation with truster and trustee relationship in user trust network

How Does Item Relational Network Help? [WWW'17]

Clothing Domain

Non-clothing Domain

Visually related

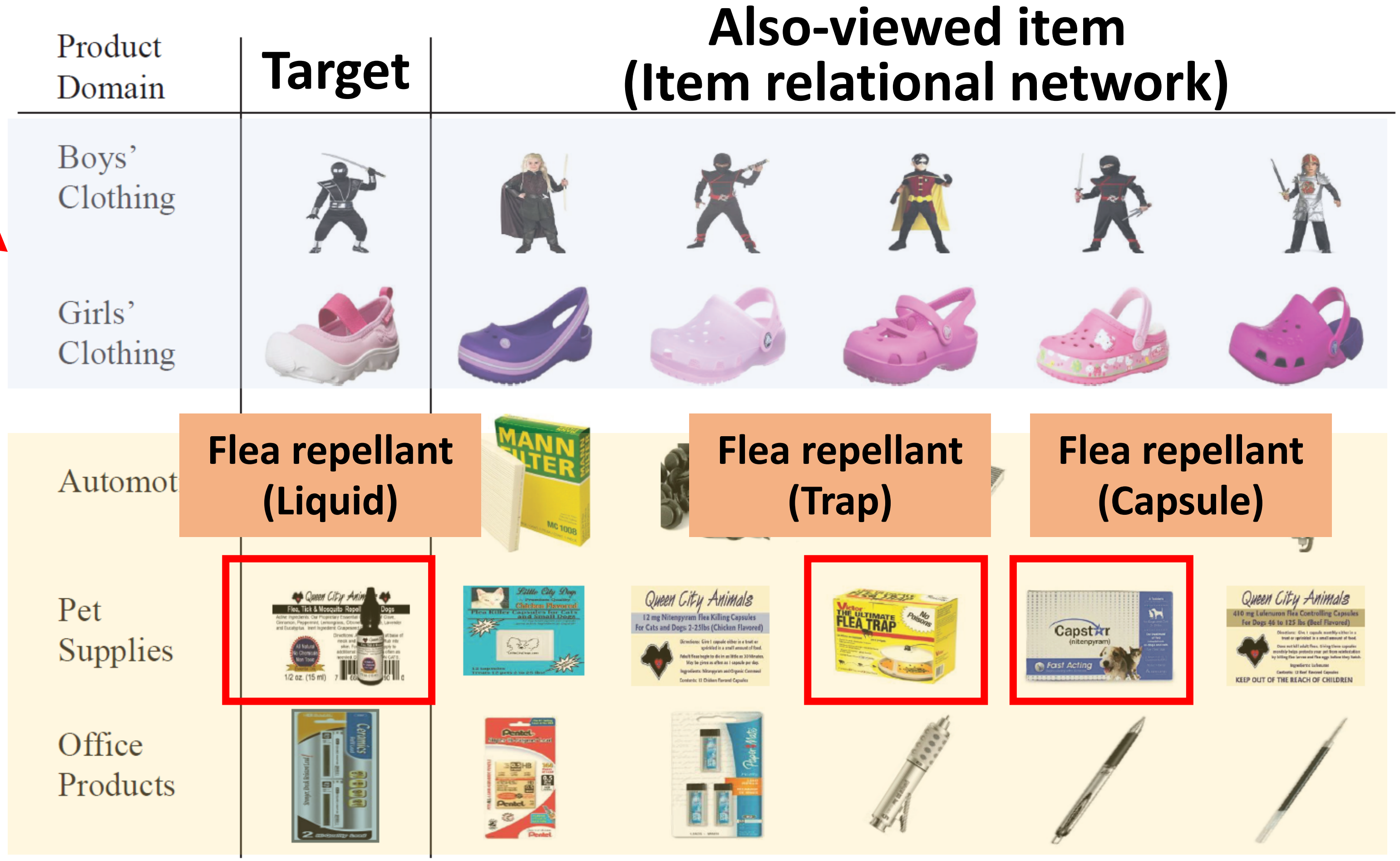
Functionally related

[WWW17 Park et al] Do "Also-Viewed" Products Help User Rating Prediction?

How Does Item Relational Network Help? [WWW’17]

Clothing Domain

Non-clothing Domain



Visually related

Functionally related

Functionality cannot be captured by item images

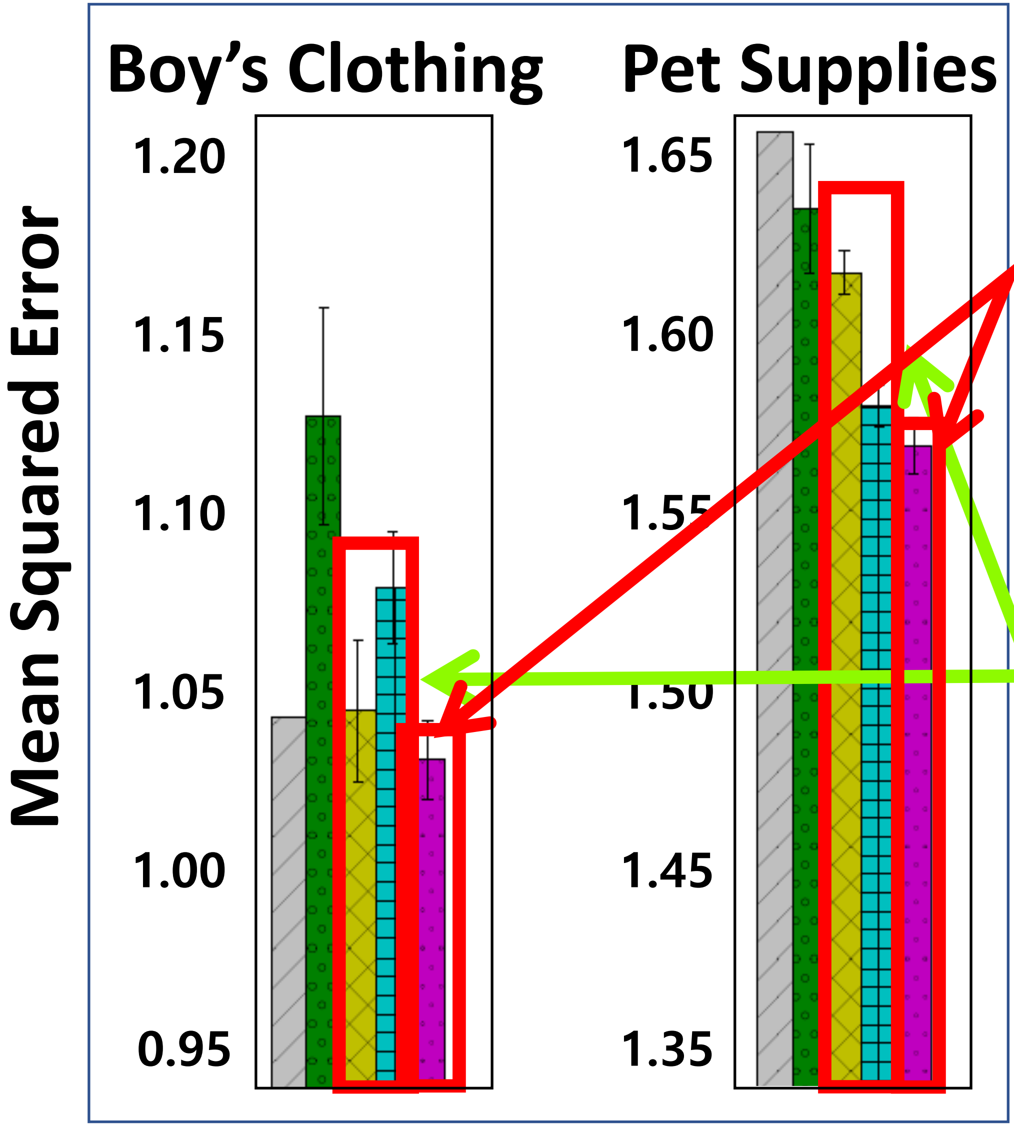
[WWW17 Park et al] Do "Also-Viewed" Products Help User Rating Prediction?

How Does Item Affinity Network Help? [WWW'17]

Rating Prediction Task

PMF	Rating only
VMF	+ Visual
MCF	+ Also-viewed
VMCF	+ Both

The smaller the better



Combining both visual information and item affinity network **performs the best regardless of the domains**

Clothing domain
→ Visual information

Non-clothing domain
→ also-viewed information

[WWW17 Park et al] Do "Also-Viewed" Products Help User Rating Prediction?

[RecSys16 Kim and Park et al] Convolutional Matrix Factorization for Document Context-Aware Recommendation
[InfoSci17 Kim and Park et al] Deep Hybrid Recommender Systems via Exploiting Document Context and Statistics of Items
[SIGIR18 Hyun and Park et al] Review Sentiment-Guided Scalable Deep Recommender System
[InforSci18 Hyun, Park et al] Target-Aware Convolutional Neural Network for Target-Level Sentiment Analysis
[IJCAI19a Lee, Park et al] Action Space Learning for Heterogeneous User Behavior Prediction
[IJCAI19b Kim, Kim, Park et al] Sequential and Diverse Recommendation with Long Tail

Others

- **User demographics / Item metadata** [RecSys'15]
  - Age, resident, purchased time, category, price, quantity, etc
  - **User purchase prediction competition (Top 1.1% out of 850 teams. Paper invited)**
- **User review / Item description text** [**RecSys'16**, InforSci'17,19, SIGIR'18]
  - **Field-Weighted Citation Impact (FWCI): 54.34 (Top 1% worldwide)**
  - **Top-2 Cited paper in RecSys'16 (Cited 311 times)**
    - c.f.) Top-1: Deep neural networks for YouTube recommendations (work by Google)
- **Temporal dynamics (sequence of clicks/purchases)** [IJCAI'19b, KNOSYS'20, sub]
- **Heterogeneous behaviors** [IJCAI'19a, InforSci'20]
  - Click, Purchase, Add-to-Cart, Favorite

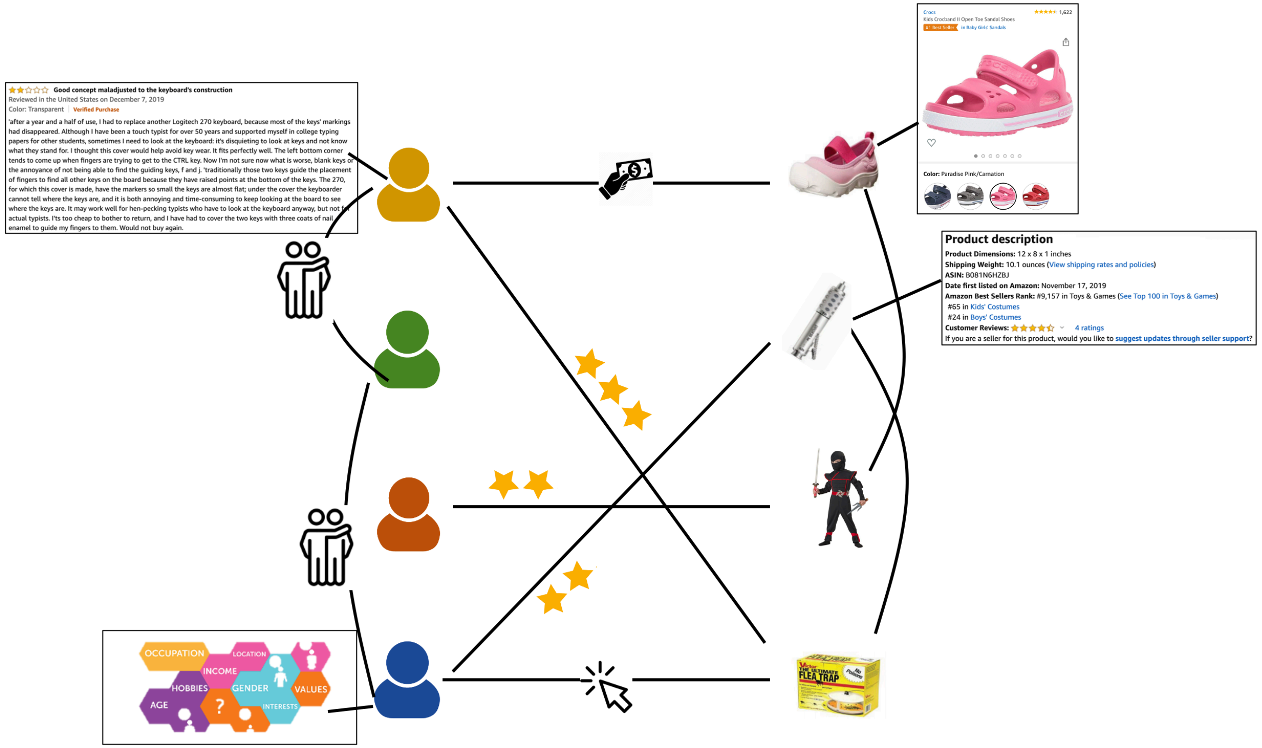
[KNOSYS20 Park et al] An Encoder-Decoder Switch Network for Purchase Prediction
[InforSci20 Park et al] Click-aware Purchase Prediction with Push at the Top
[RecSys15 Park et al] Predicting User Purchase in E-commerce by Comprehensive Feature Engineering and Decision Boundary Focused Under-Sampling
[sub Hyun, Cho, Park et al] Time-variant Review Representation for Recommender System

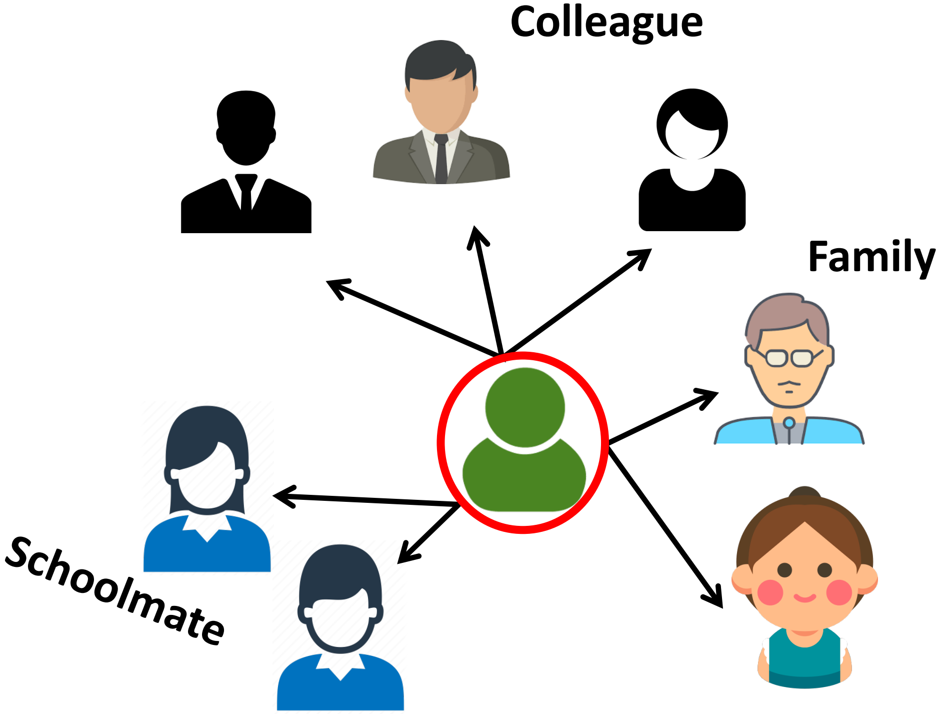
Part 2

So far,
different types of user behavior

Part 3

Now, we will go
deeper into each user





Social network

What?
How?

Why?

Outline

Part 1: Research Motivation & Background

Part 2: **Multi-modal** User Behavior Analysis

Part 3: **Multi-aspect** User Behavior Analysis

Part 4: Vision for the future

Overview: User Behavior Analysis

Part2: Multi-modal

User
- Demographics [RecSys15]
- Social network [WWW16, InforSci16]
- User Review [InforSci19, SIGIR18]

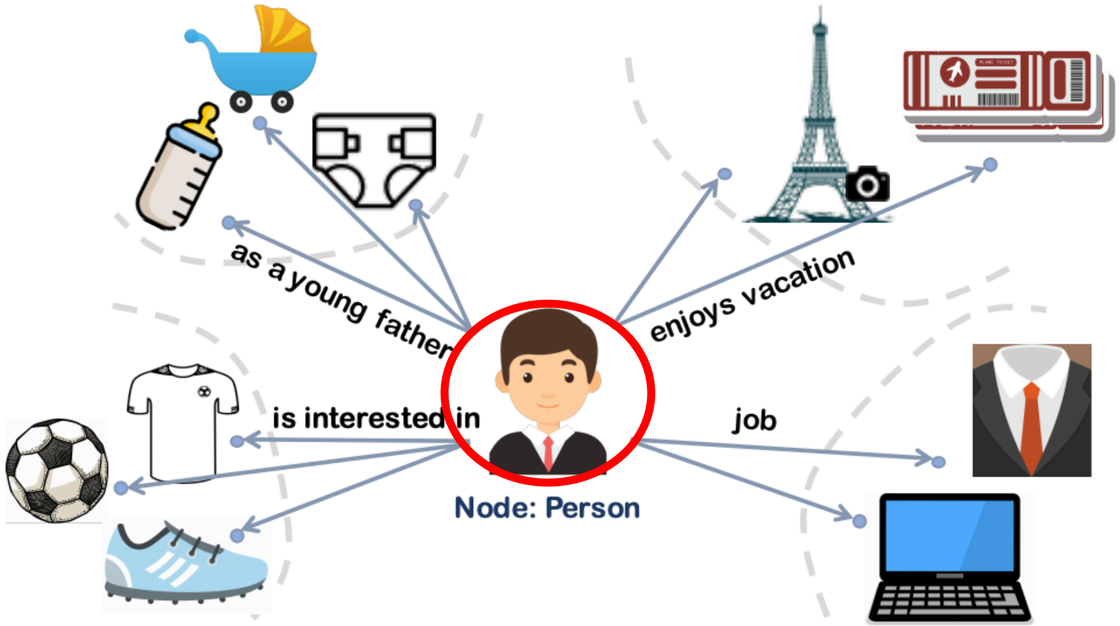
Item
- Meta-data [RecSys15, ICDM18]
- Item description [Recsys16, InforSci17]
- Image [WWW17]
- Affinity relation [WWW17]

User-Item
- Rating [RecSys16, WWW17, InforSci17,19, SIGIR18,ICDM18]
- Click/Purchase/ Add-to-cart/ Add-to-favorite [IJCAI19b, InforSci20,KNOSYS20]
- Bookmark [ICDM18]
- Time information [IJCAI19a, KNOSYS20, sub_b]

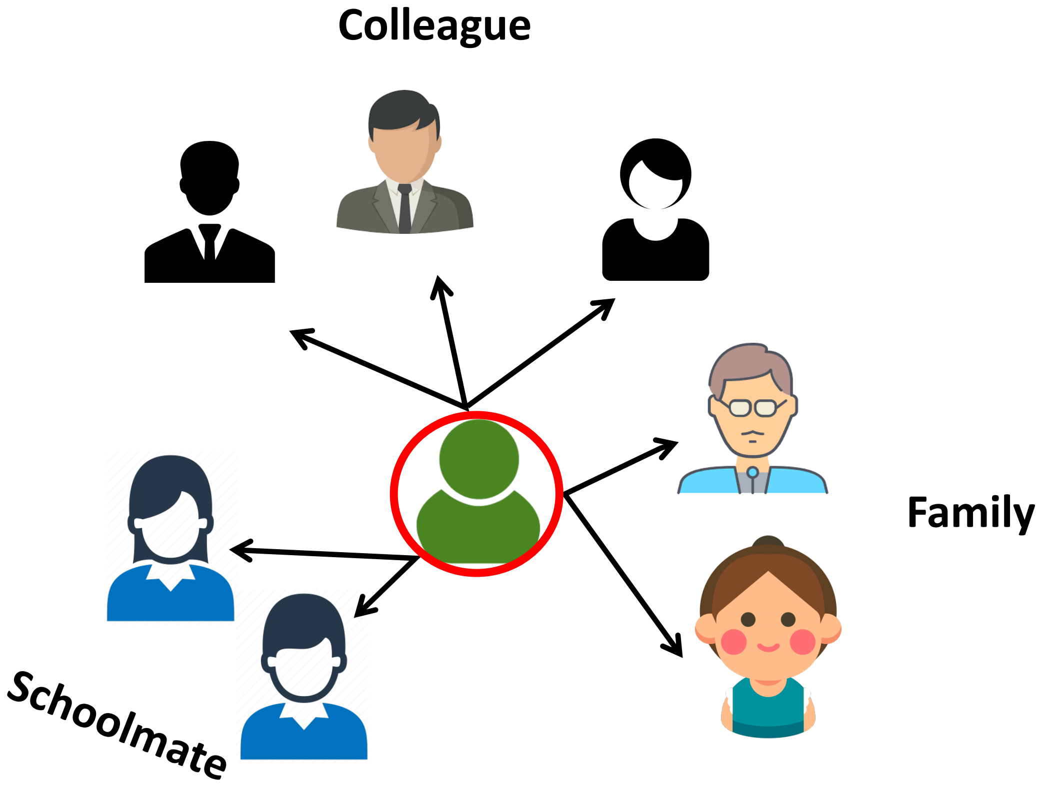
Part 3: Multi-aspect

- Homogeneous Network [KDD20]
- Multiplex Network [AAAI20, KNOSYS20]
- Heterogeneous Network [ICDM18, CIKM19, KDD20, sub_c]

User behavior has **multi-aspect**



Purchase history



Social network

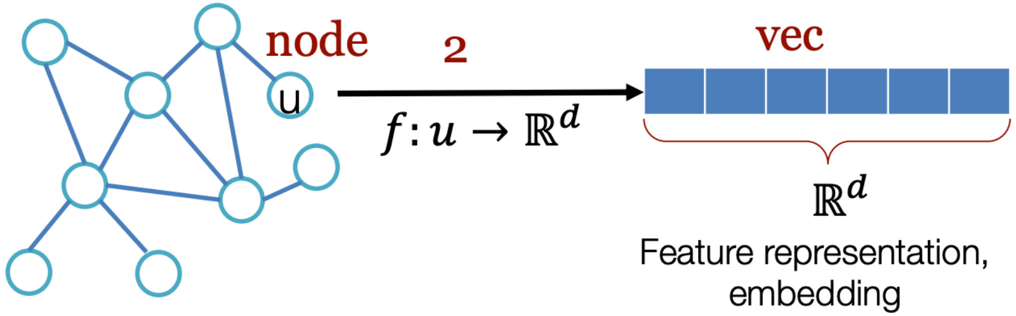
How to differentiate among multiple aspects?

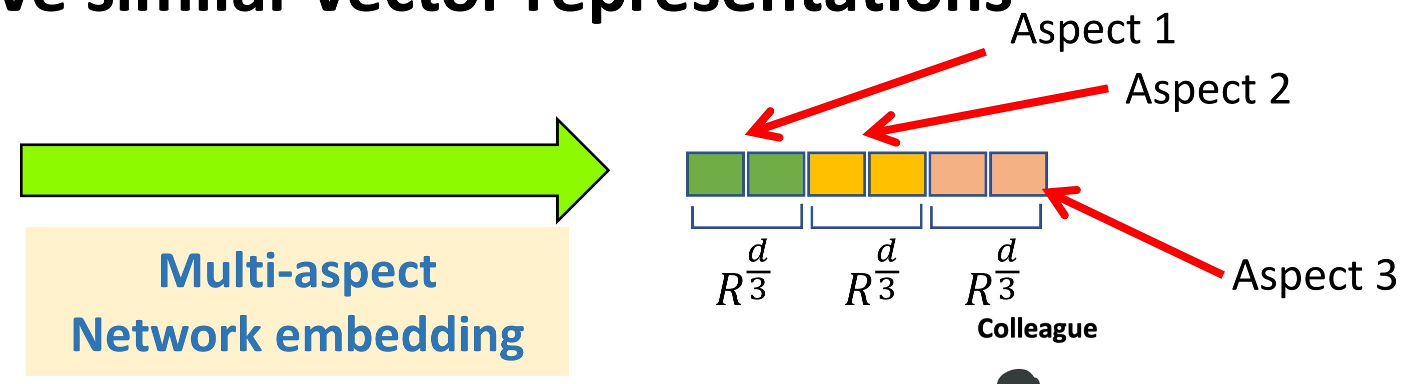
Figure credit: Liu. et al, KDD19

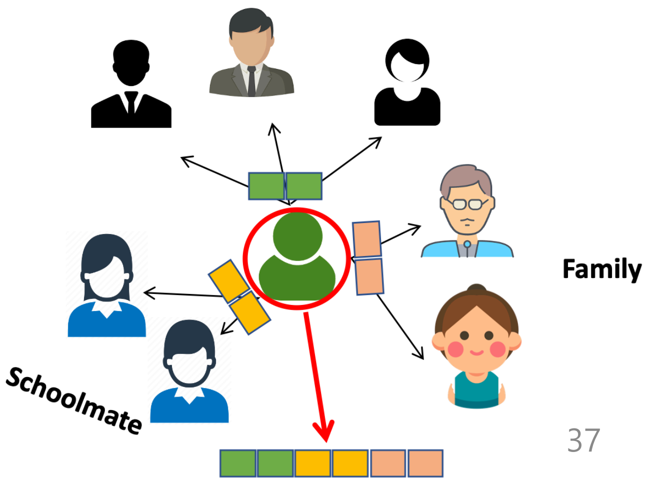
Figure credit: Jure Leskovec's slide

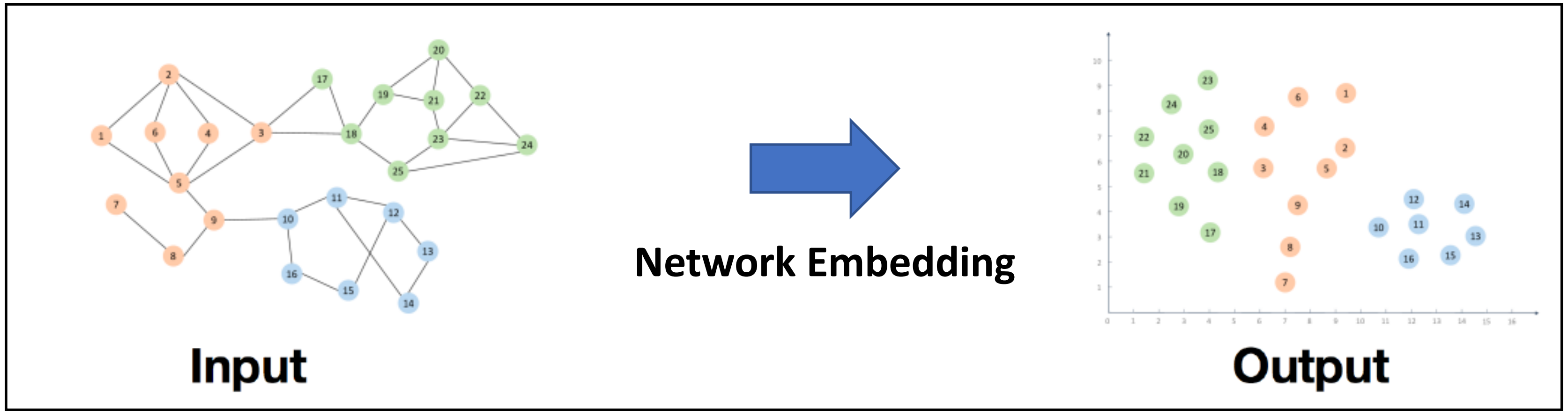
What is Network Embedding?

- Encode nodes so that **similarity in the embedding space** approximates **similarity in the original network**
- **Similar nodes in a network have similar vector representations**

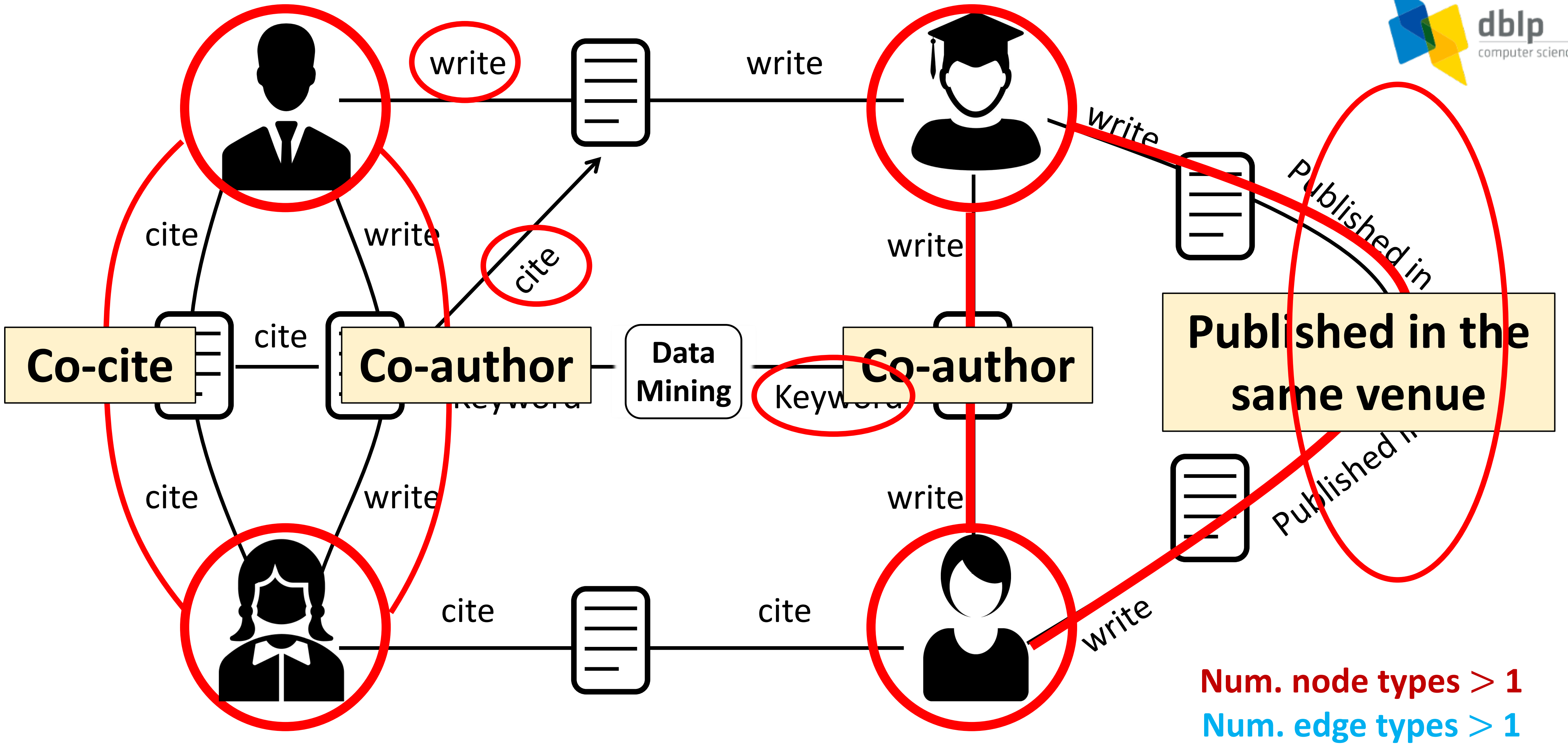








Network of User Behavior



Num. node types > 1

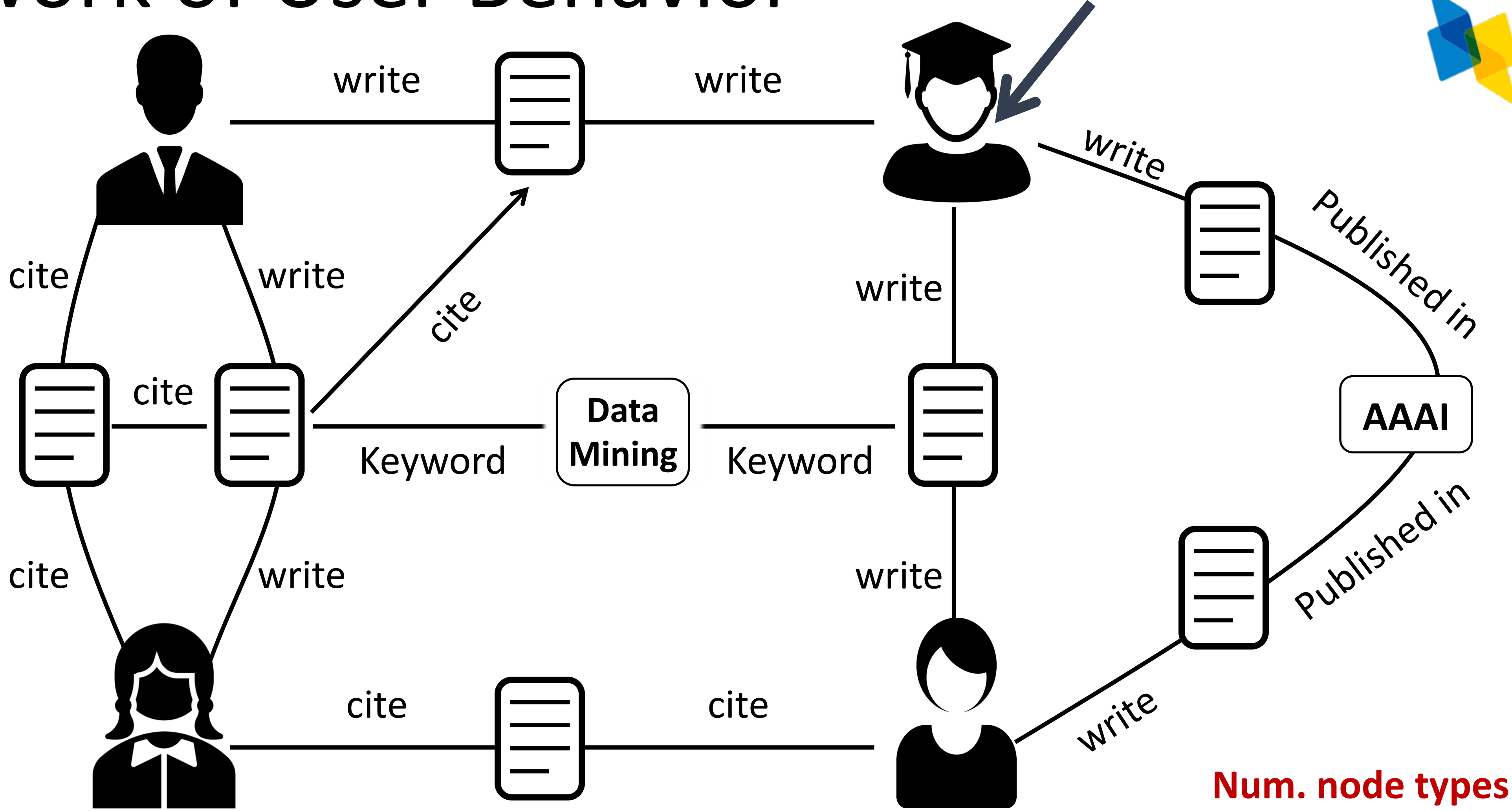
Num. edge types > 1

Heterogeneous Information Network

Network of User Behavior

Task: Research interest?

dblp computer science bibliography



Num. node types > 1

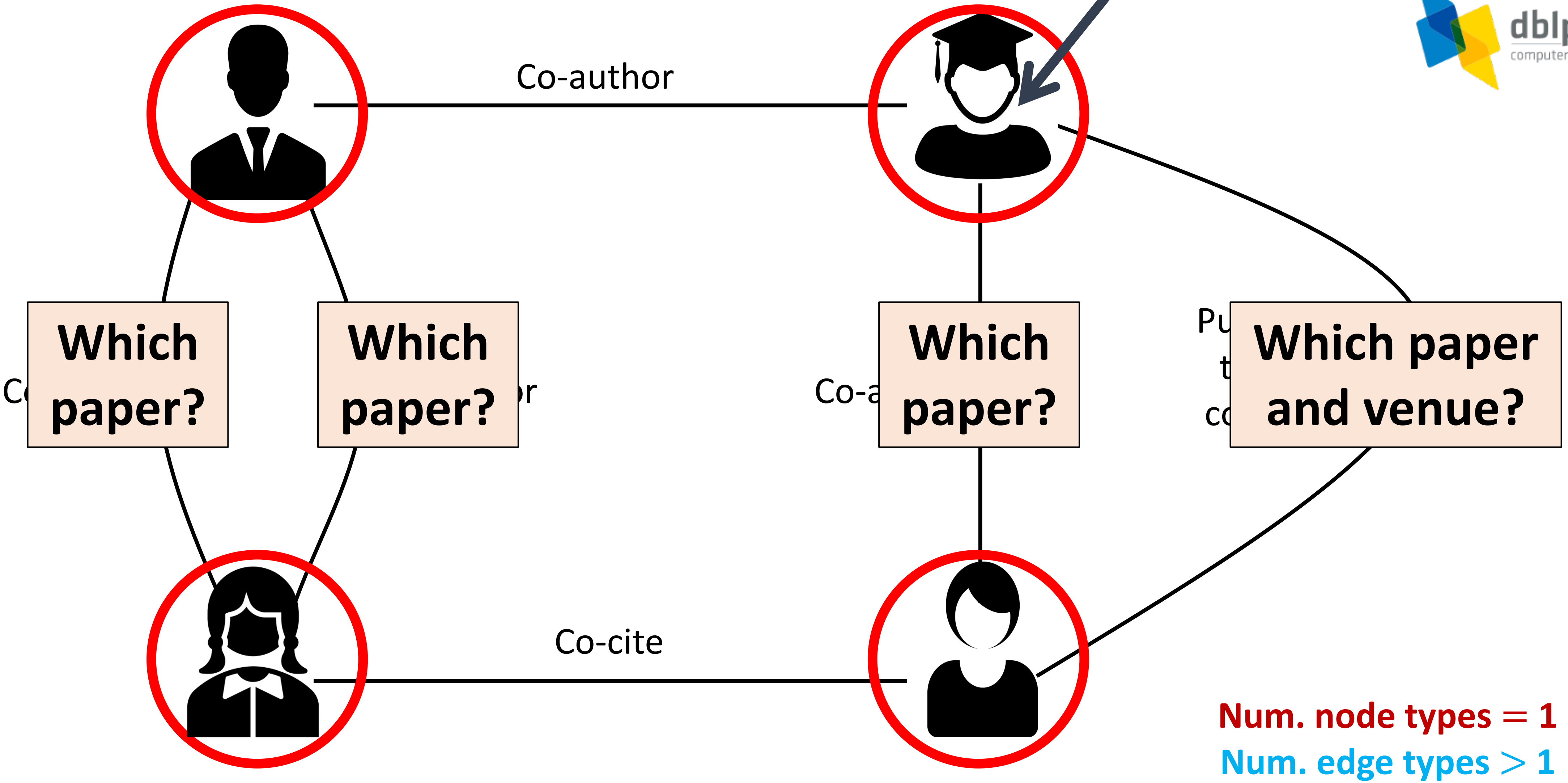
Num. edge types > 1

Heterogeneous Information Network

Network of User Behavior

Task: Research interest?

dblp computer science bibliography



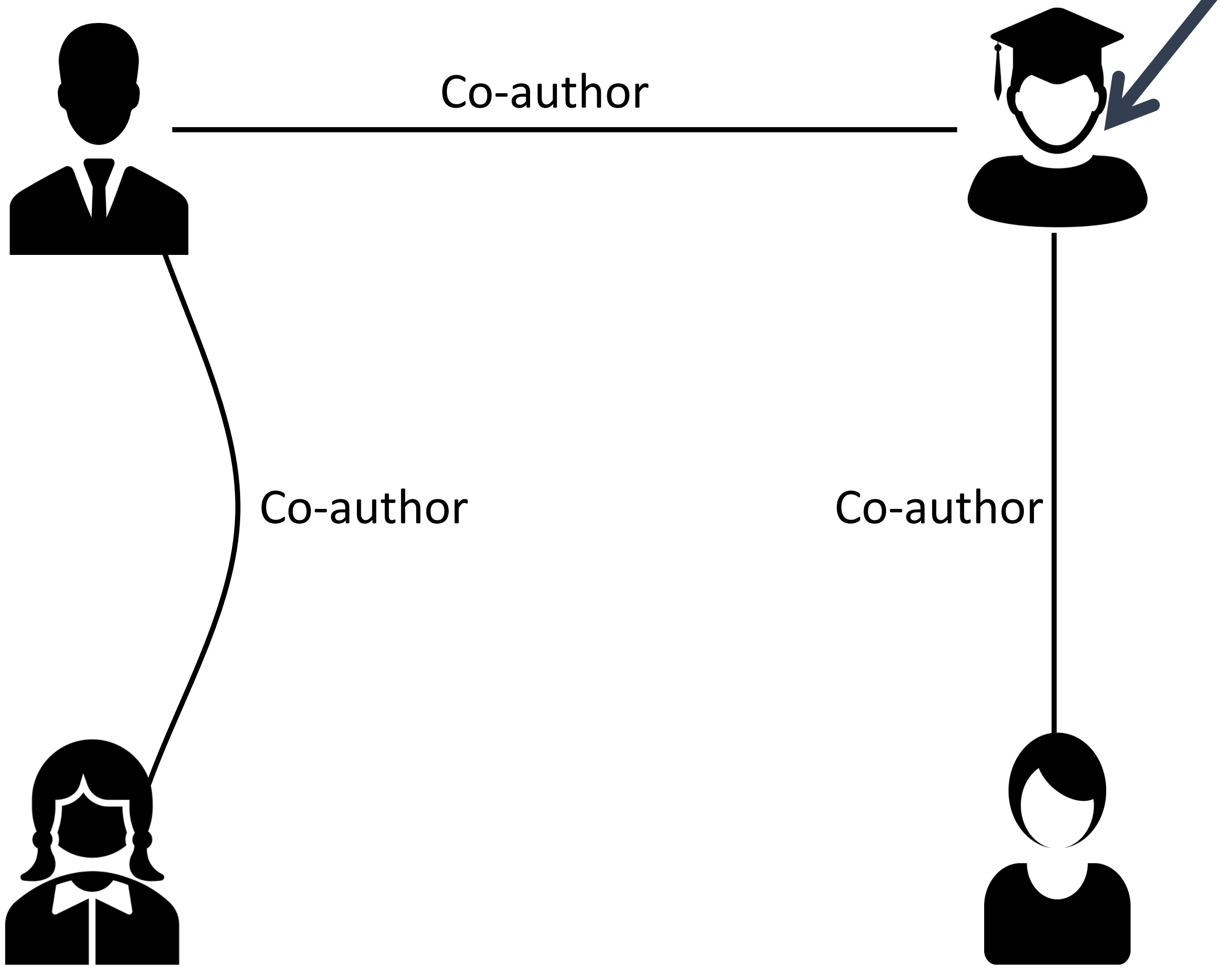
Num. node types = 1

Num. edge types > 1

Multiplex Network

Network of User Behavior

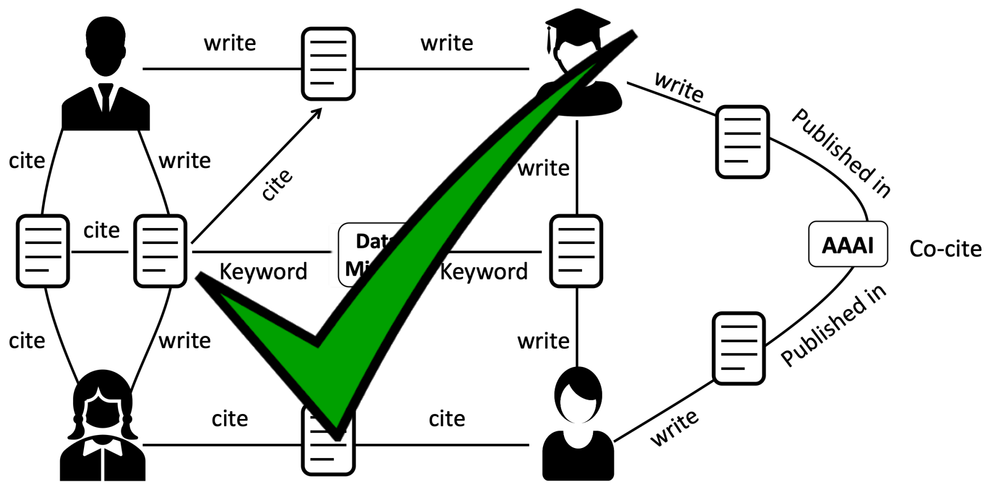
Task: Research interest?



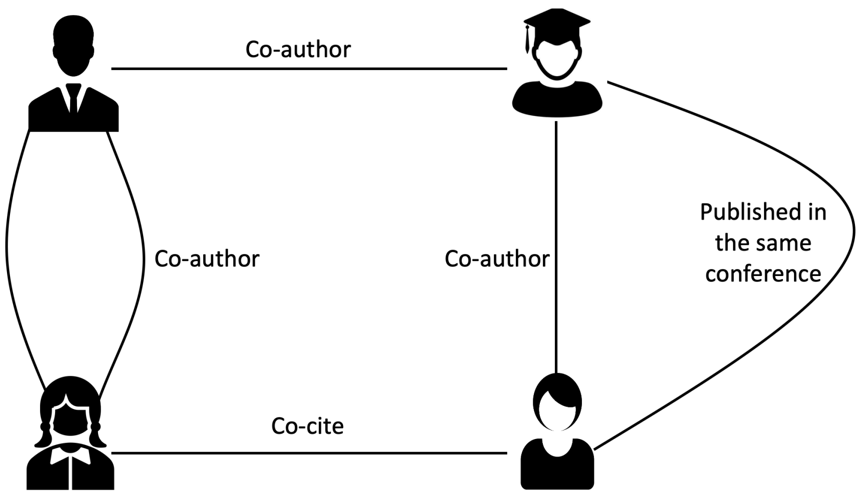
Num. node types = 1
Num. edge types = 1

Homogeneous Network

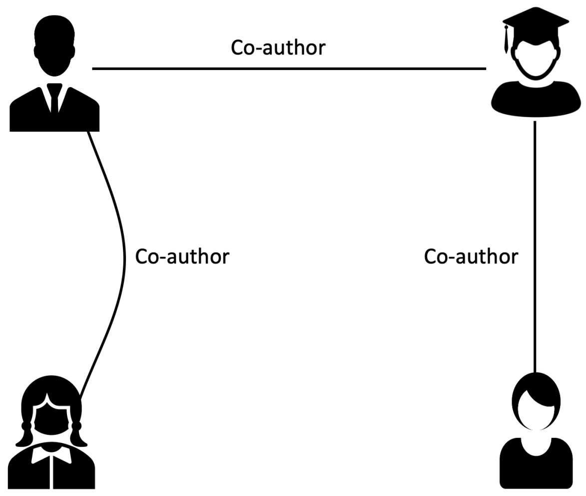
Recap: Multi-aspect User Behavior



Heterogeneous network
[ICDM18, CIKM19, KDD20, sub_c]



Multiplex network
[AAAI20, KNOSYS20]

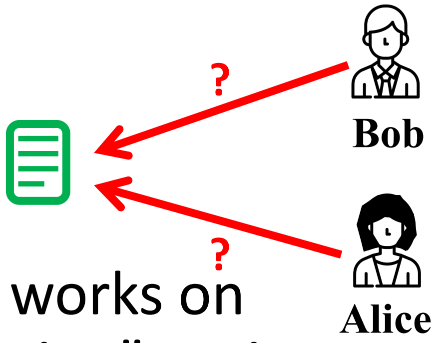


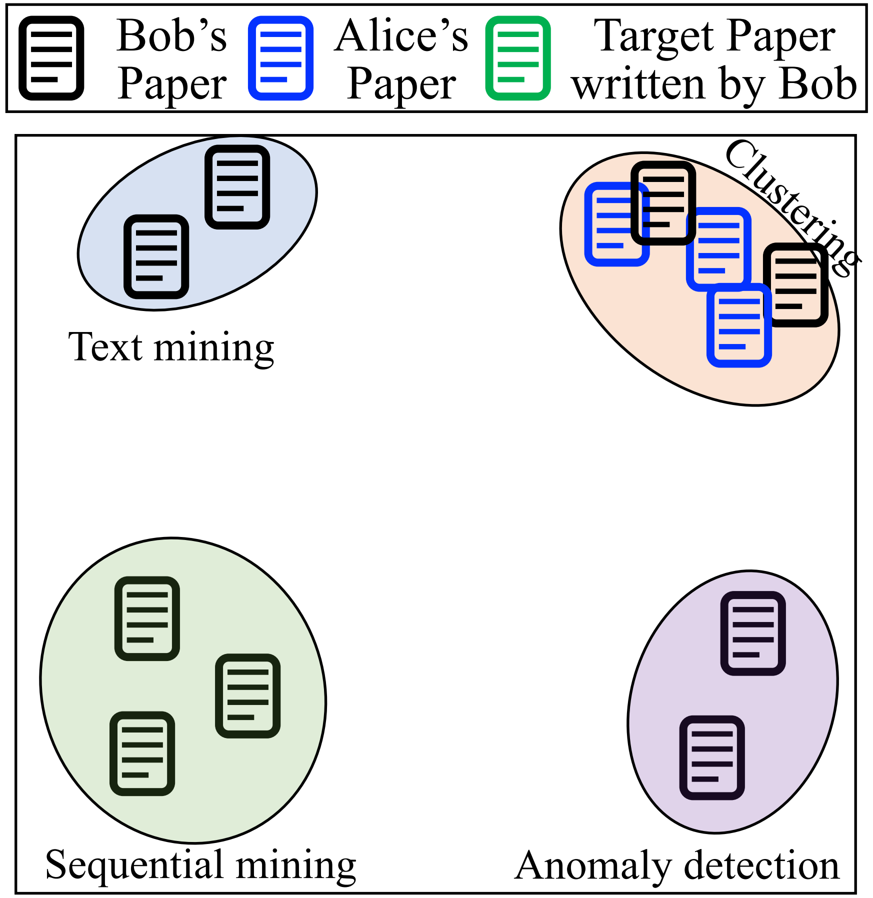
Homogeneous network
[KDD20]

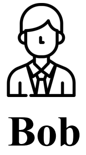
Rich **The amount of Information** Sparse

Easy **Difficulty of understanding multi-aspect user behavior** Difficult

Why does it matter? ex) Author identification







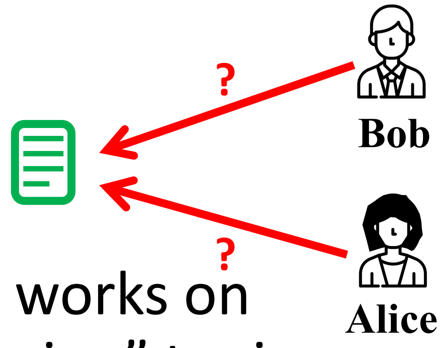
Has **multiple** papers in various research areas

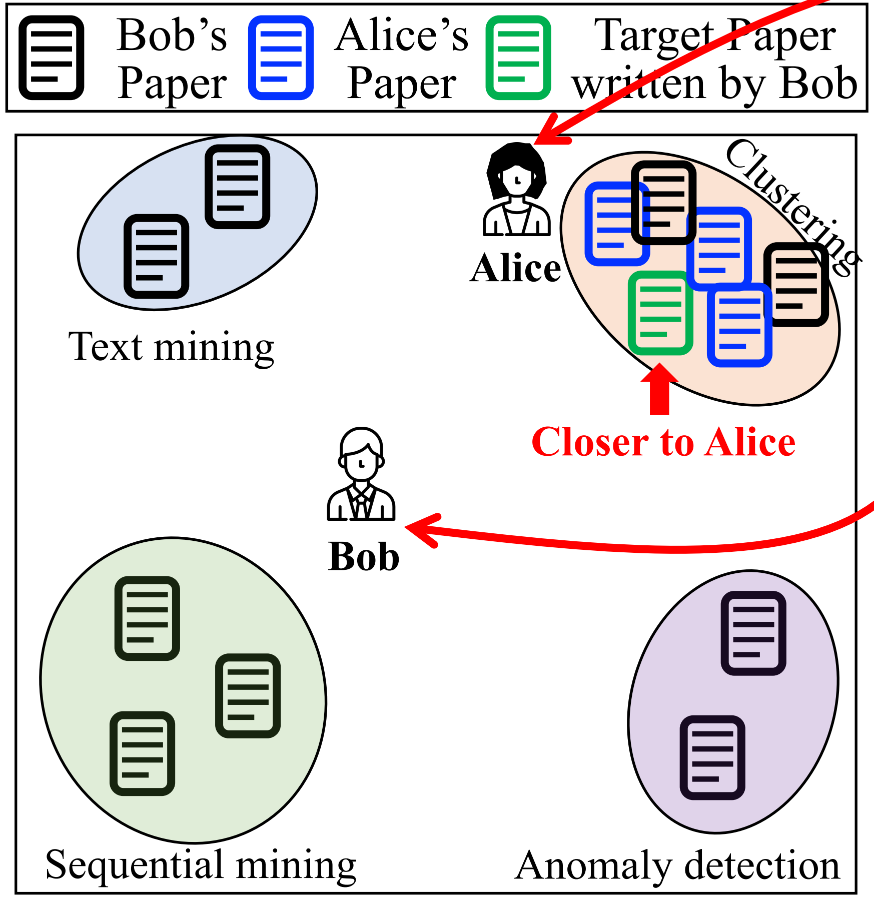
Only works on "Clustering" topic

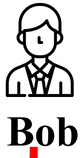
Where are the **optimal points for embedding**?

[CIKM19 Park et al] Task-Guided Pair Embedding in Heterogeneous Network

Why does it matter? ex) Author identification







Has **multiple** papers in various research areas

Only works on "Clustering" topic

Where are the **optimal points for embedding**?

Q. What will be the prediction for a new paper on "Clustering" written by Bob?

Ans. Alice's Paper ☹

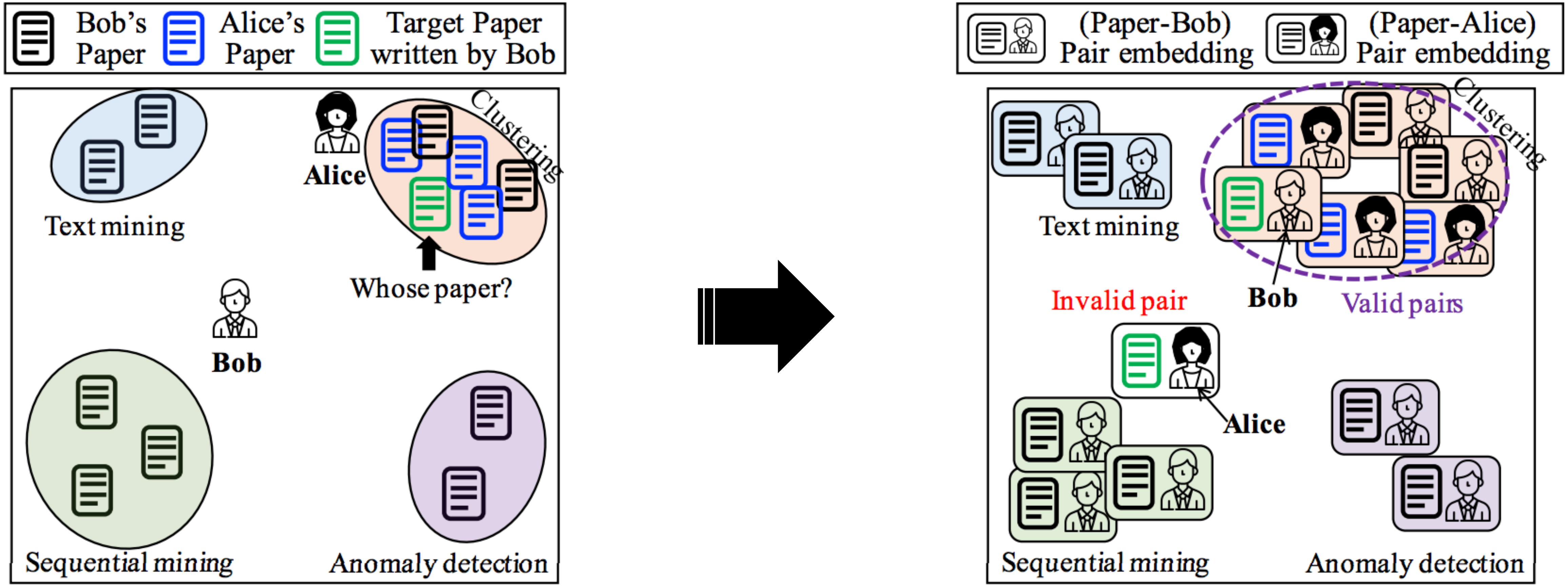
∵ Embedded together with "Clustering" papers

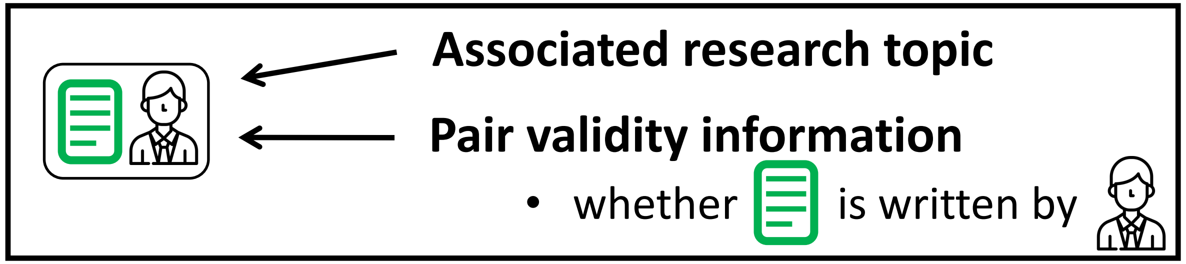
→ Closer to Alice

Bob has **multi-aspects** and should be considered

[CIKM19 Park et al] Task-Guided Pair Embedding in Heterogeneous Network

Pair Embedding Framework [CIKM'19]

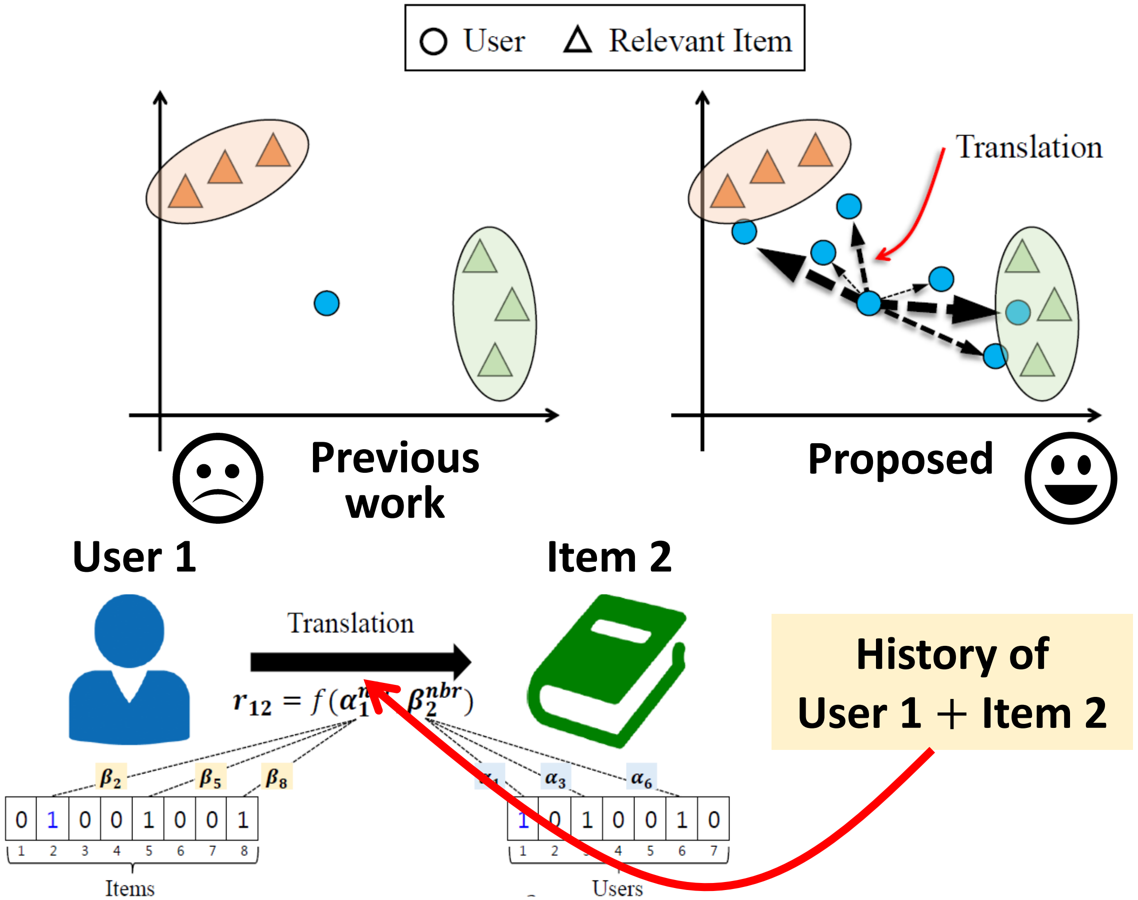


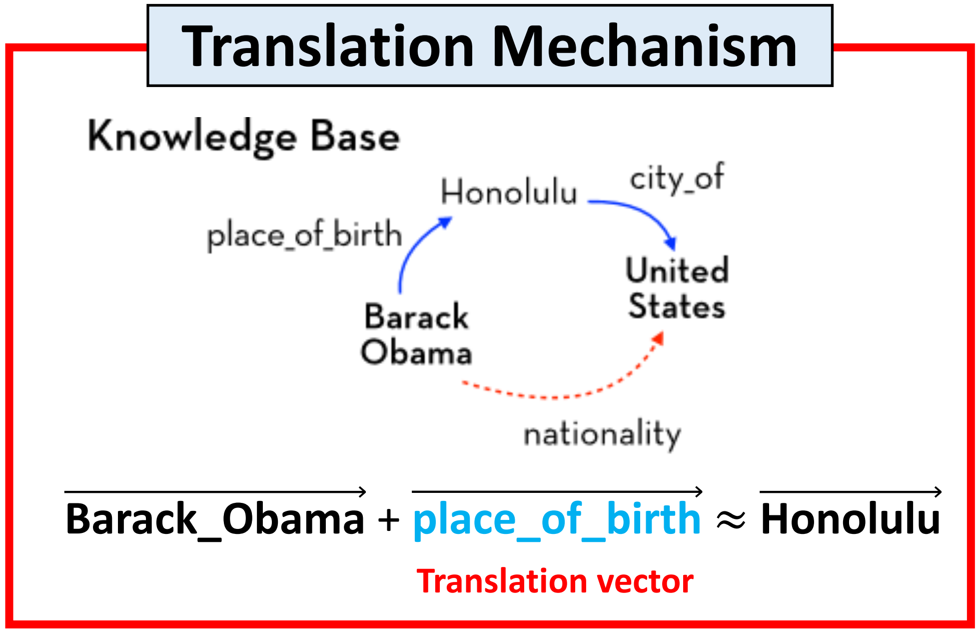


[CIKM19 Park et al] Task-Guided Pair Embedding in Heterogeneous Network

Translation-based Metric learning Approach [ICDM'18]

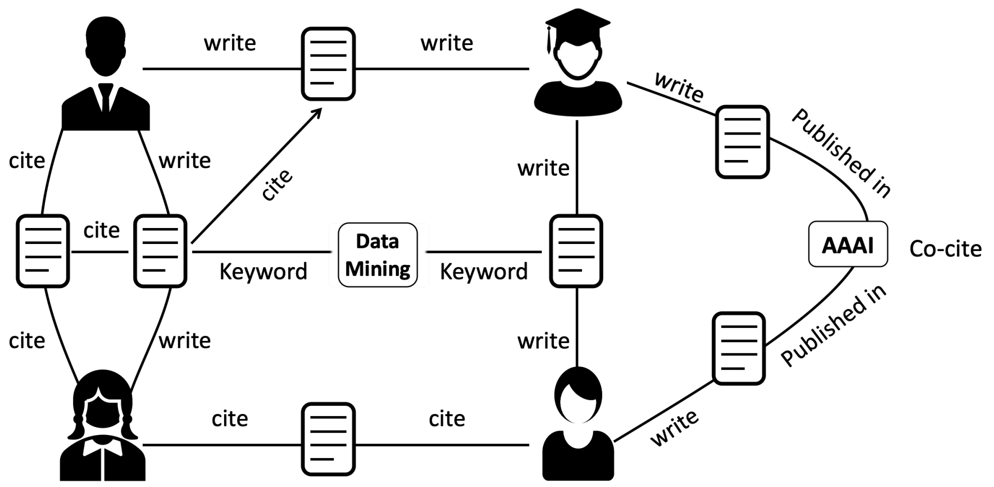
- Alternative way of pair embedding
- "Translate" each user to items considering the user's relation with items



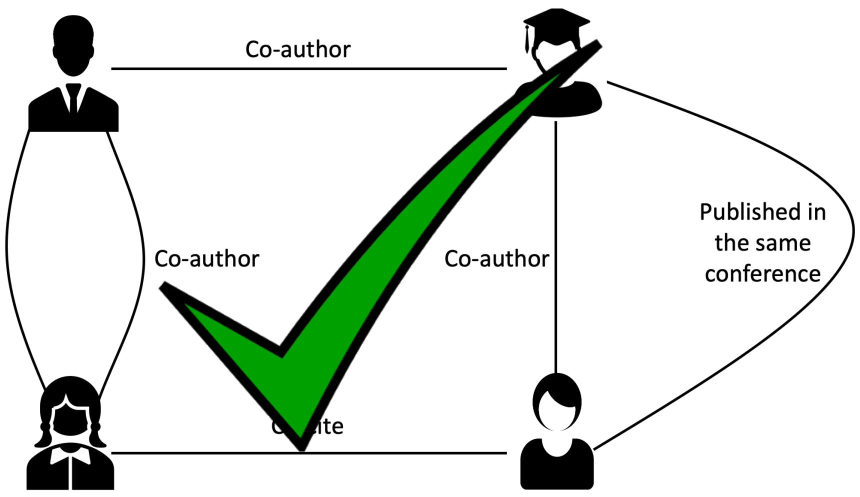


[ICDM18 Park et al] Collaborative Translational Metric Learning

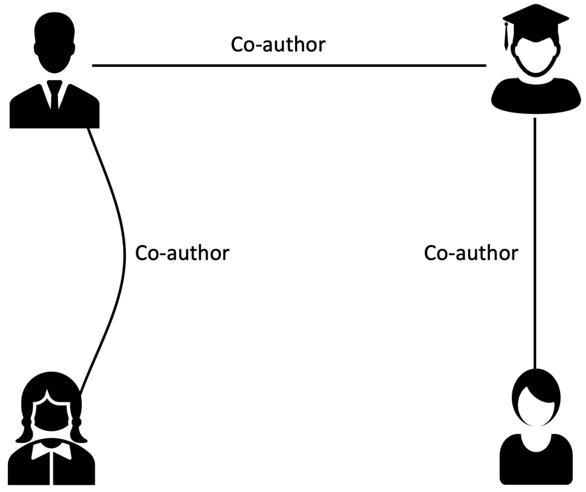
Recap: Multi-aspect User Behavior



Heterogeneous network
[ICDM18, CIKM19, sub_c]



Multiplex network
[AAAI20, KBS20]



Homogeneous network
[sub_a]

The amount of Information

Rich ← Sparse

Difficulty of understanding multi-aspect user behavior

Easy → Difficult

Why Consider Multiplex Network?

- **Example 1: Social network**
  - Relationship between users
- **Example 2: E-commerce**

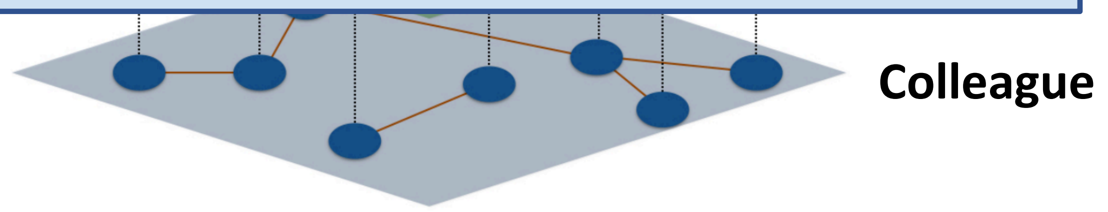
Num. node types $= 1$
Num. edge types > 1

Node: User

We know different relations exists between nodes.

Then, how can we use them to model multi-aspect user behavior?

  - Relationship between movies
    - Common director, common actor
- **Example 5: Transportation network in a city**
  - Relation between locations in a city
    - Bus, train, car, taxi

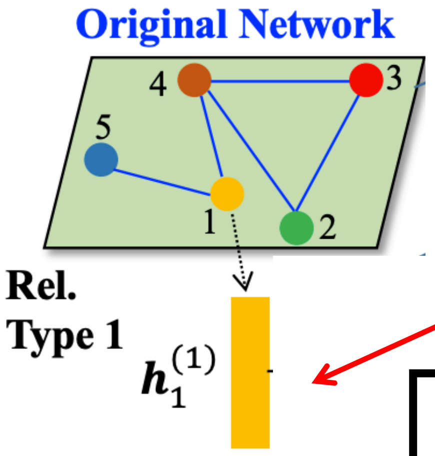


Social Network

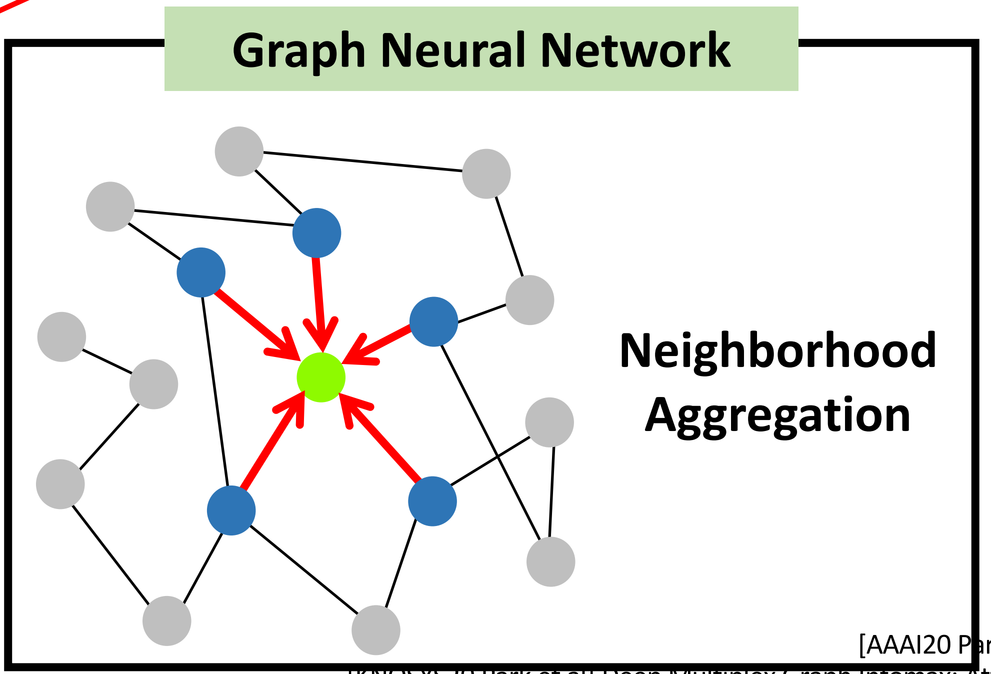
Deep Multiplex Graph Infomax [AAAI’20]

Social Network

Schoolmate



Obtained by Graph Neural Network



[AAAI20 Park et al] Unsupervised Attributed Multiplex Network Embedding
[KNOSYS20 Park et al] Deep Multiplex Graph Infomax: Attentive Multiplex Network Embedding using Global Information

Deep Multiplex Graph Infomax [AAAI'20]

This is required for unsupervised training

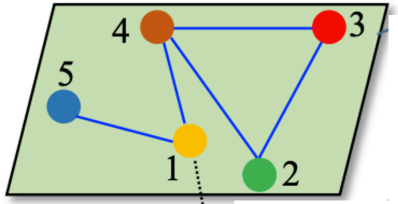
Social Network

Shuffle feature matrix

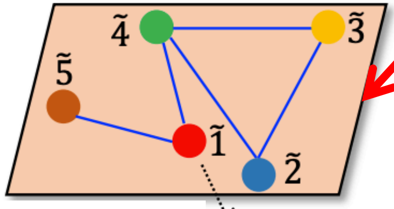
Schoolmate

$\boldsymbol{X}$ Corrupt $\tilde{\boldsymbol{X}}$

Original Network



Corrupted Network



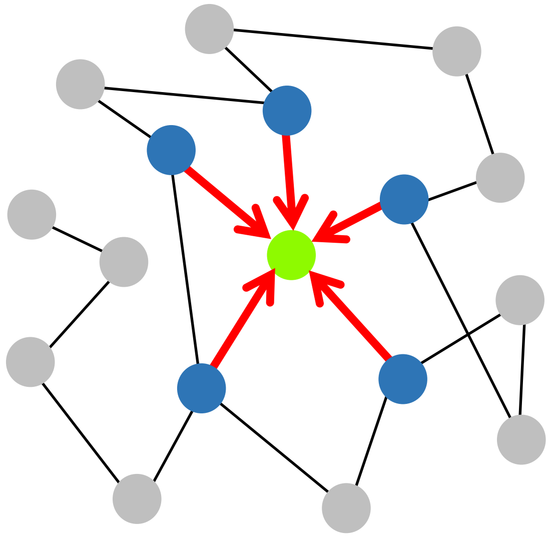
Obtained by Graph Neural Network

Rel. Type 1 $\boldsymbol{h}_1^{(1)}$

$\tilde{\boldsymbol{h}}_1^{(1)}$

Local information

Graph Neural Network



Neighborhood Aggregation

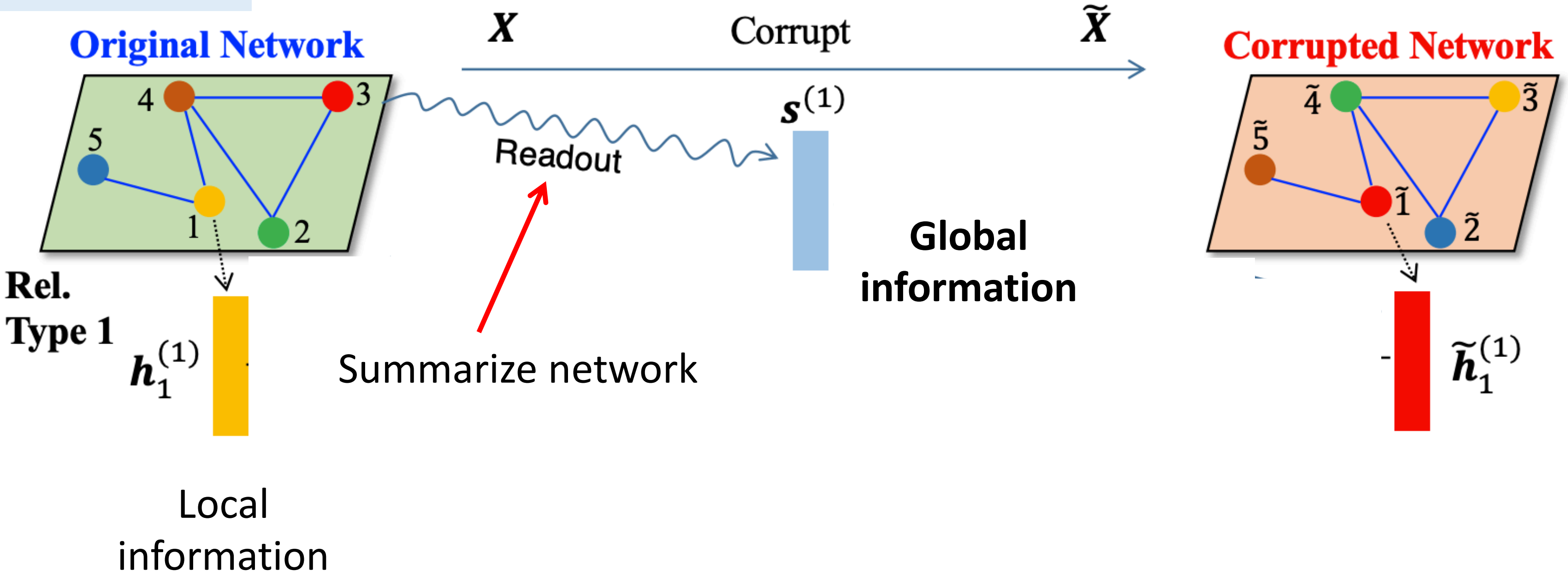
Local information

[AAAI20 Park et al] Unsupervised Attributed Multiplex Network Embedding
[KNOSYS20 Park et al] Deep Multiplex Graph Infomax: Attentive Multiplex Network Embedding using Global Information

Deep Multiplex Graph Infomax [AAAI'20]

Social Network

Schoolmate

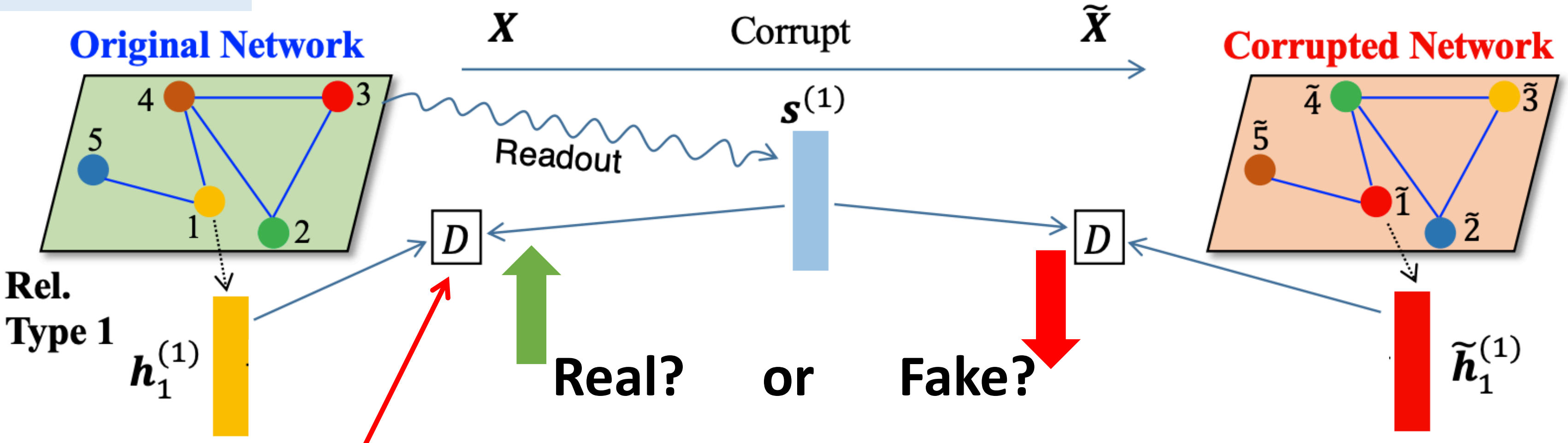


[AAAI20 Park et al] Unsupervised Attributed Multiplex Network Embedding
[KNOSYS20 Park et al] Deep Multiplex Graph Infomax: Attentive Multiplex Network Embedding using Global Information

Deep Multiplex Graph Infomax [AAAI'20]

Social Network

Schoolmate



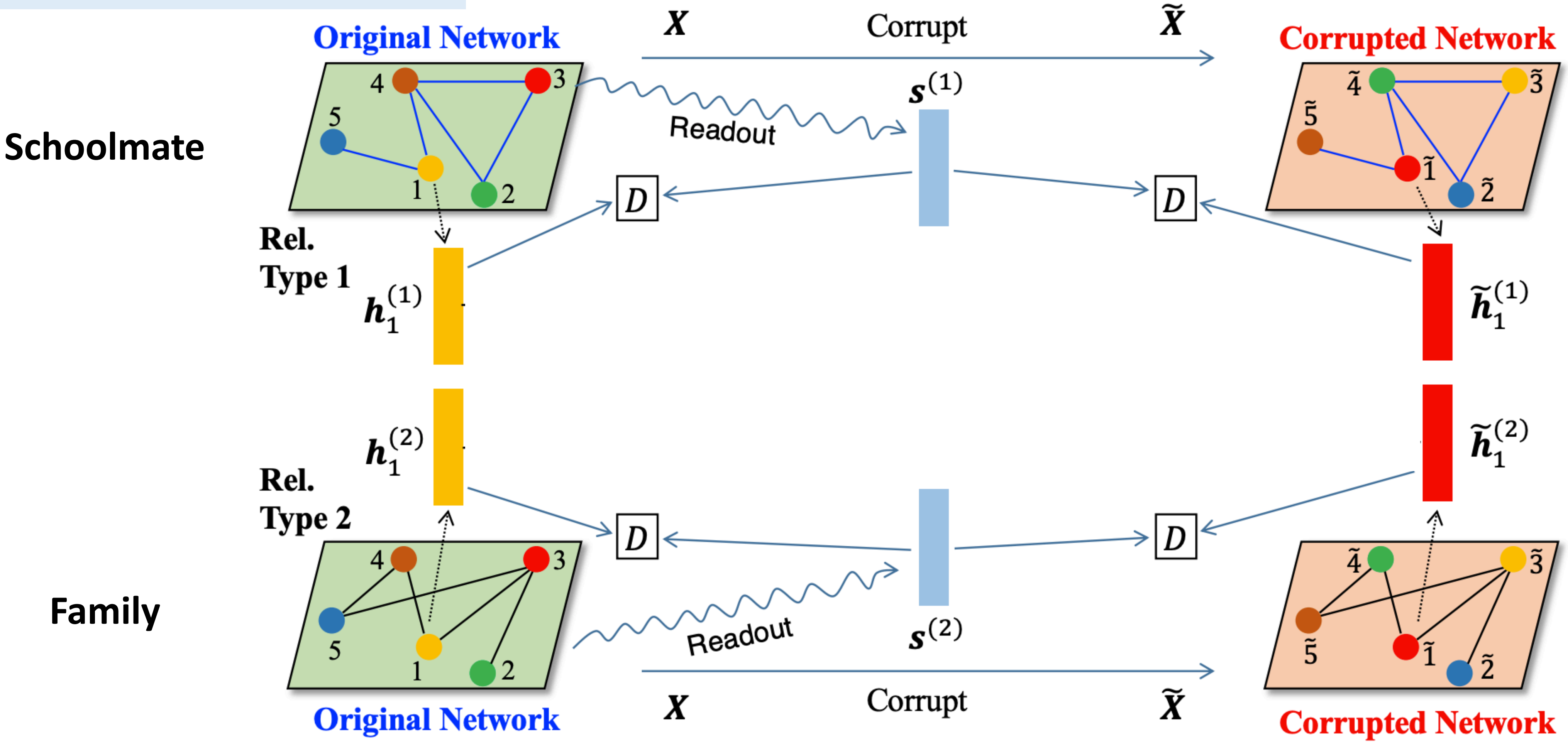
$$\mathcal{D}\left(\mathbf{h}_i^{(r)}, \mathbf{s}^{(r)}\right) = \sigma(\mathbf{h}_i^{(r)T}\mathbf{M}^{(r)}\mathbf{s}^{(r)})$$

Score matrix w.r.t. relation r

[AAAI20 Park et al] Unsupervised Attributed Multiplex Network Embedding
[KNOSYS20 Park et al] Deep Multiplex Graph Infomax: Attentive Multiplex Network Embedding using Global Information

Deep Multiplex Graph Infomax [AAAI'20]

Social Network



[AAAI20 Park et al] Unsupervised Attributed Multiplex Network Embedding
[KNOSYS20 Park et al] Deep Multiplex Graph Infomax: Attentive Multiplex Network Embedding using Global Information

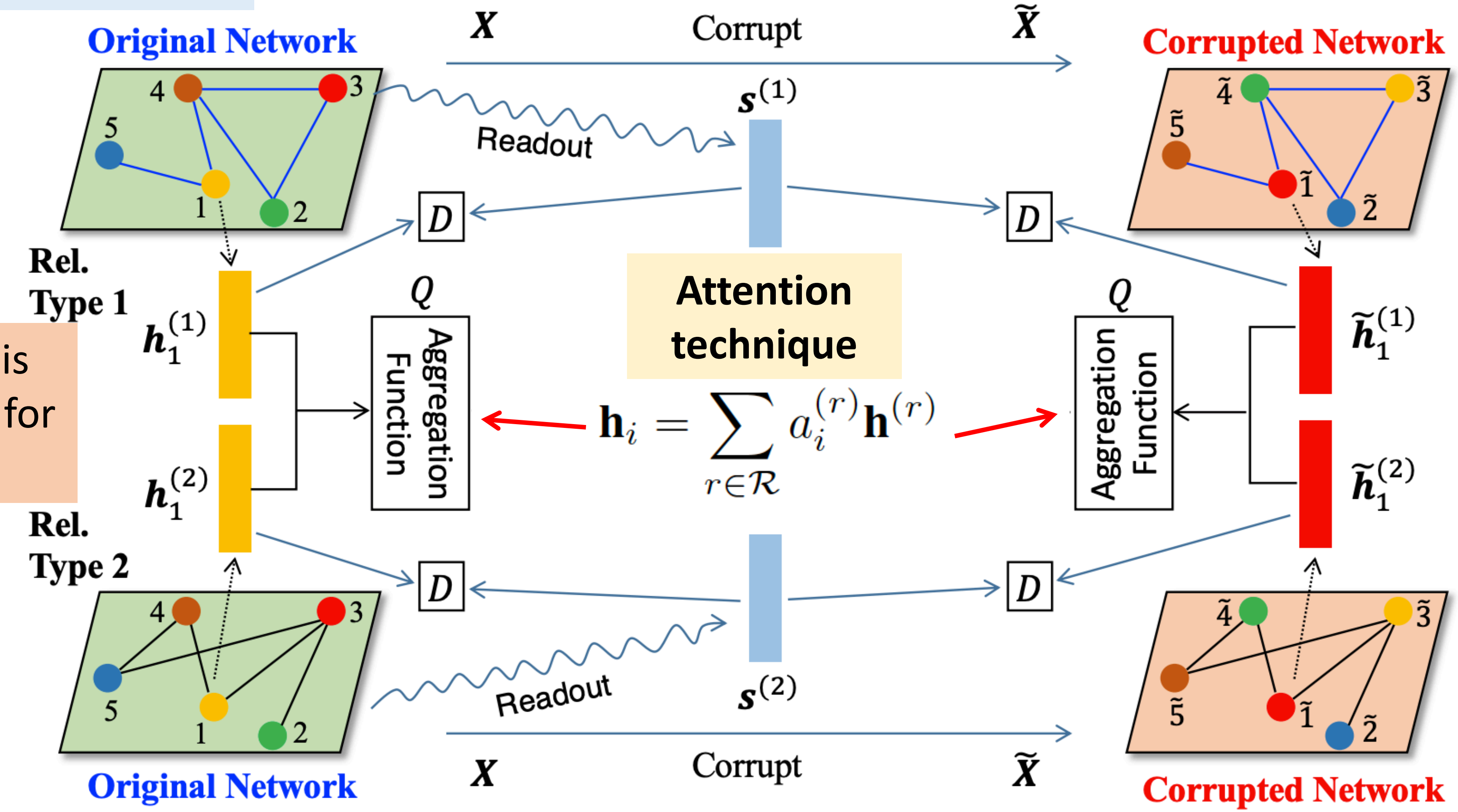
Deep Multiplex Graph Infomax [AAAI'20]

Social Network

Schoolmate

Which relation is **more important** for each node?

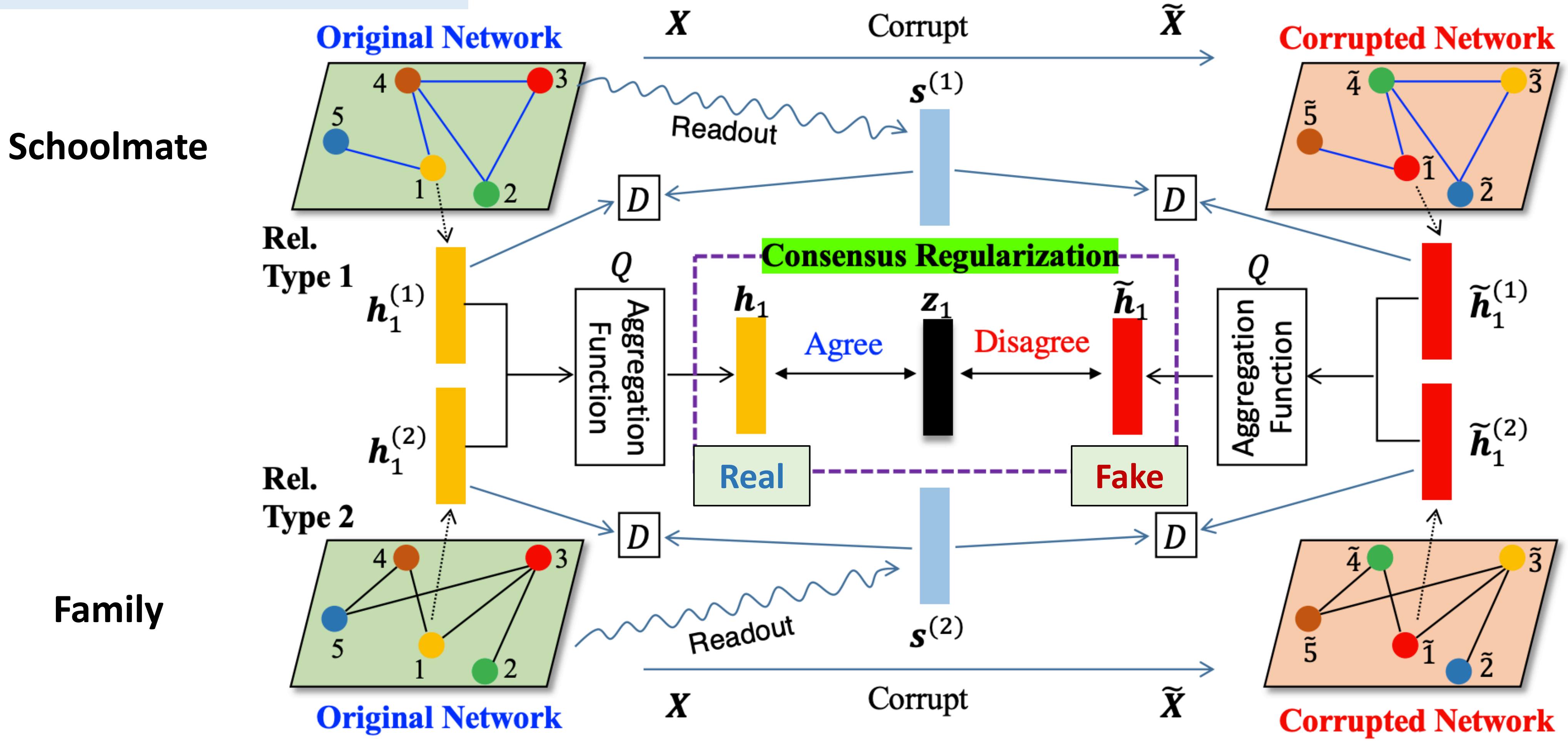
Family



[AAAI20 Park et al] Unsupervised Attributed Multiplex Network Embedding
[KNOSYS20 Park et al] Deep Multiplex Graph Infomax: Attentive Multiplex Network Embedding using Global Information

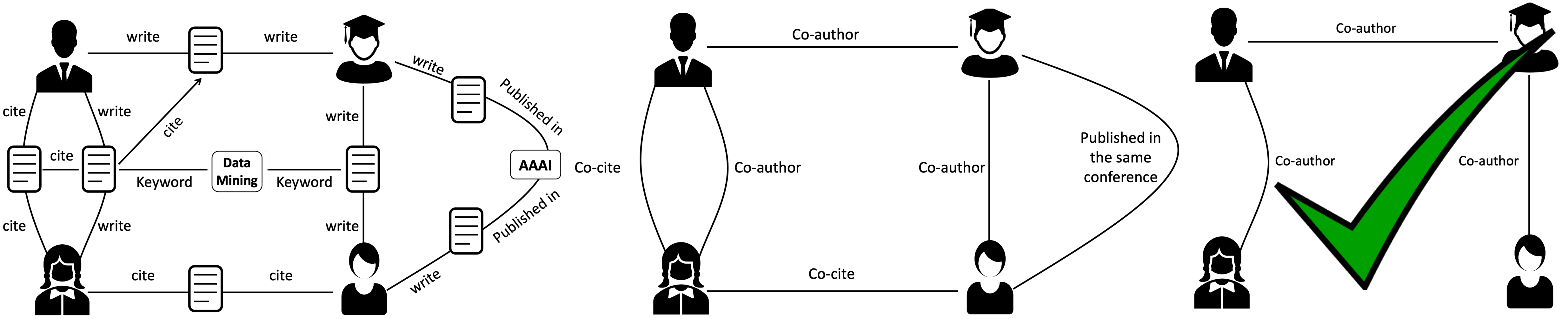
Deep Multiplex Graph Infomax [AAAI'20]

Social Network



[AAAI20 Park et al] Unsupervised Attributed Multiplex Network Embedding
[KNOSYS20 Park et al] Deep Multiplex Graph Infomax: Attentive Multiplex Network Embedding using Global Information

Recap: Multi-aspect User Behavior



Heterogeneous network
[ICDM18, CIKM19, KDD20, sub_c]

Multiplex network
[AAAI20, KNOSYS20]

Homogeneous network
[KDD20]

The amount of Information

Rich ← → Sparse

Difficulty of understanding multi-aspect user behavior

Easy → Difficult

Why Homogeneous Network?

In reality,

Node features (types) or labels are not always given

☹

Can we capture the multi-aspect user behavior **solely based on the network structure**?

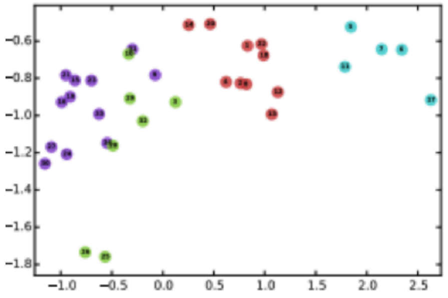
The most challenging case!

[KDD20 Park et al] Unsupervised Differentiable Multi-aspect Node Representation Learning

Clustering-based aspect assignment

1. Each node always has the same **fixed aspect** regardless of its current context
2. Final network embedding **quality depends on the performance of clustering**
   - Training **cannot be done end-to-end**

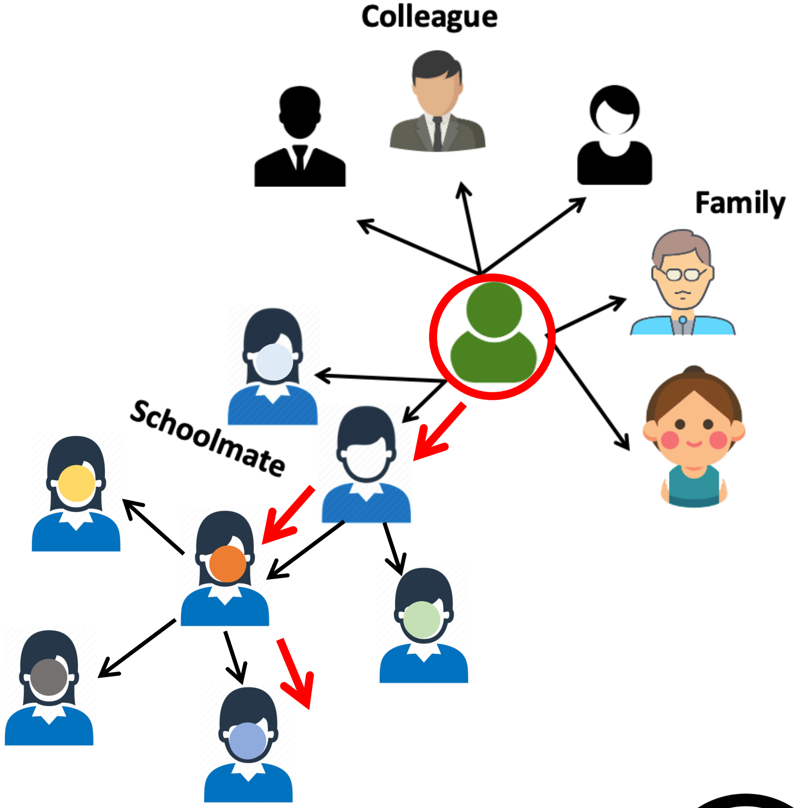
Start network embedding



Done!

[WWW19 Epasto et al]
[KDD19 Liu et al]

Context-based aspect assignment [KDD'20]

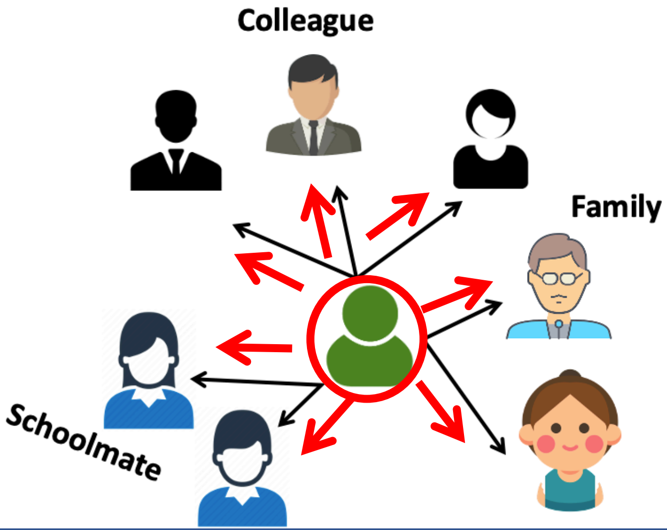


Considers
multi-hop neighbors

Context: Schoolmate

Assign "Schoolmate" aspect

Previous clustering-based method

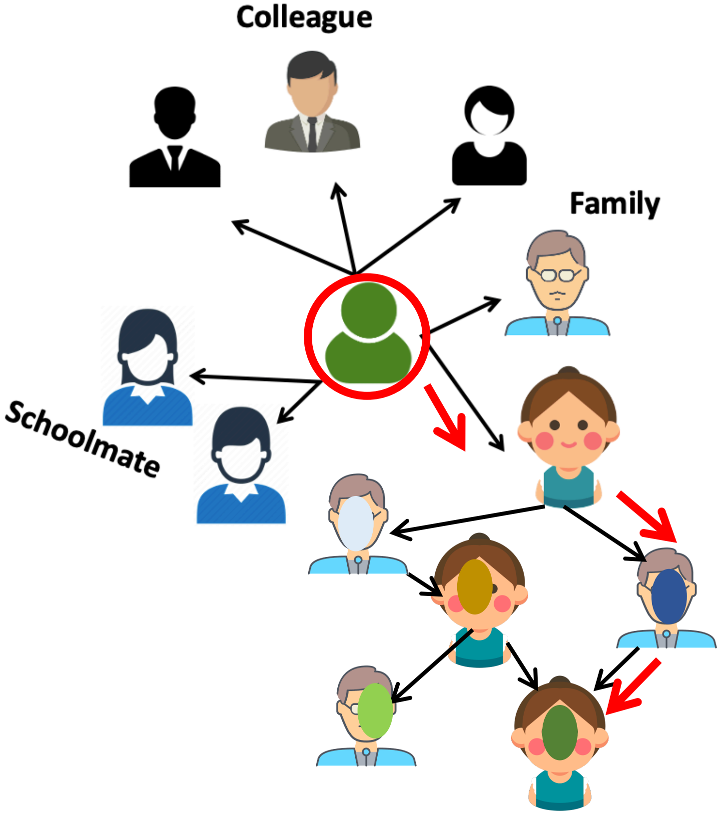


Only considers one-hop neighbors

More effective for **capturing multi-aspect** user behavior

[KDD20 Park et al] Unsupervised Differentiable Multi-aspect Node Representation Learning

Context-based aspect assignment [KDD'20]



Context: Family

Assign "Family" aspect

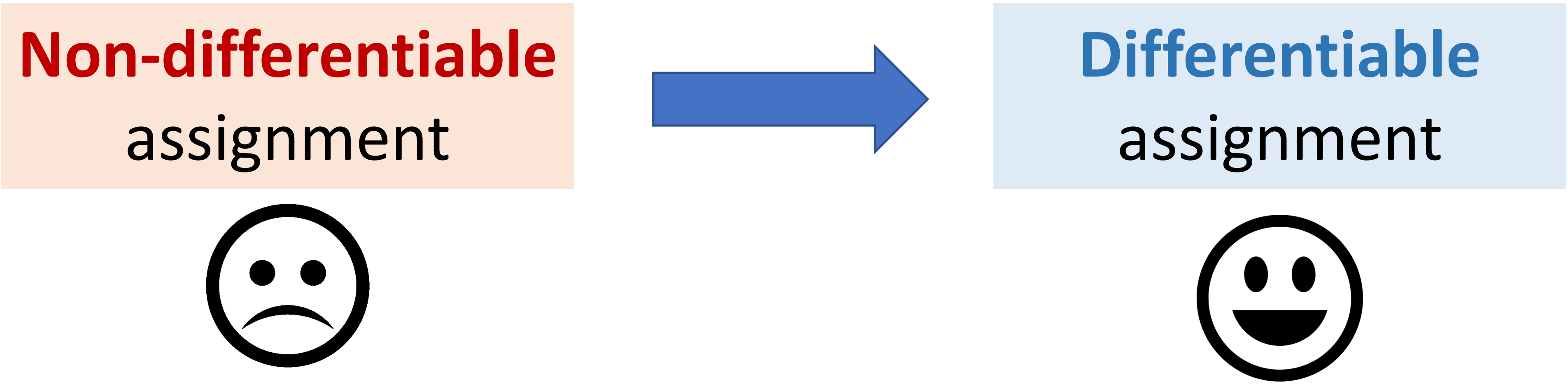
Assign a single aspect for each node based on the context

This assignment process is non-differentiable

[KDD20 Park et al] Unsupervised Differentiable Multi-aspect Node Representation Learning

Gumbel-Softmax based Aspect Selection [KDD'20]

- Adopt the **Gumbel-softmax trick** to **dynamically sample aspects based on the context**



[KDD20 Park et al] Unsupervised Differentiable Multi-aspect Node Representation Learning

Gumbel-Softmax Trick (Jang et al, 2017)

- A simple way to draw a one-hot sample $\boldsymbol{z}$ from the **categorical distribution**
- **Given**: A K-dimensional **categorical distribution** with class probability $\pi_1, \pi_2, \ldots, \pi_K$

Non-differentiable

Gumbel noise drawn from Gumbel(0,1)

$$g_i = -\log(-\log(u_i))$$
$$u_i \sim Uniform(0,1)$$

☹ Gumbel-max

$$z = \text{one-hot}\left(\text{argmax}_i \left[\log \pi_i + g_i\right]\right)$$

Differentiable

☺ Gumbel-softmax

$$z_i = \text{softmax}\left[\log \pi_i + g_i\right] = \frac{\exp\left(\left(\log \pi_i + g_i\right)/\tau\right)}{\sum_{j=1}^{K} \exp\left(\left(\log \pi_j + g_j\right)/\tau\right)}$$

for $i = 1, \ldots, K$

Temperature parameter

As $\tau \to 0$, samples from the Gumbel-Softmax distribution become one-hot

Continuous relaxation of discrete random variable

Gumbel-Softmax based Aspect Selection [KDD'20]

- Adopt the **Gumbel-softmax trick** to **dynamically sample aspects based on the context**

Gumbel-softmax

Embedding of v_i

Embedding of $N(v_i)$ regarding aspect s

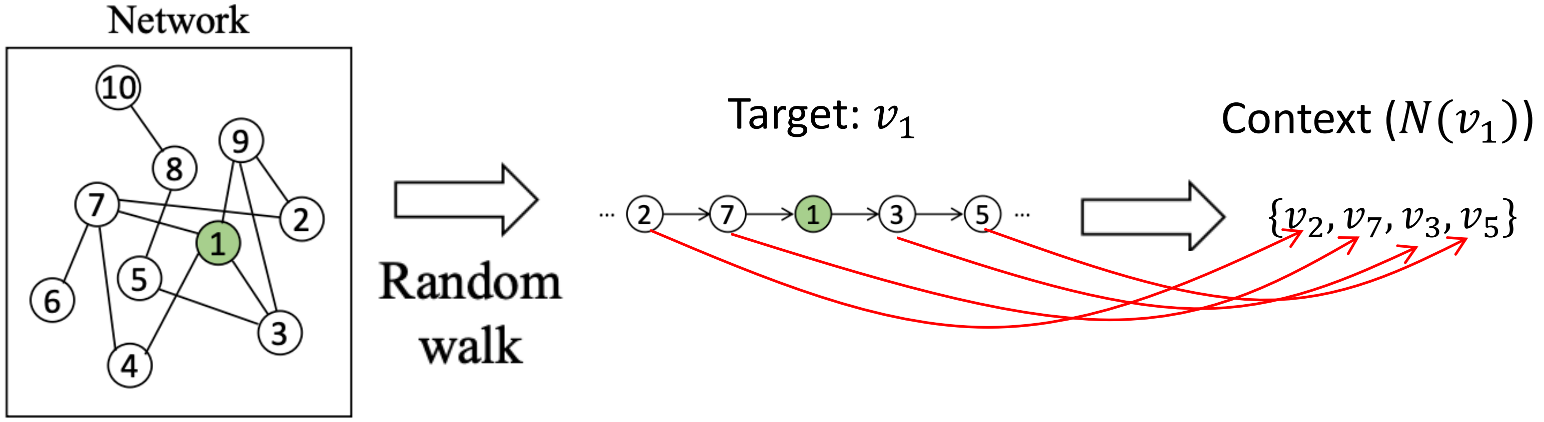
Aspect of node v_i

$$p(\quad | \mathcal{N}(v_i)) = \frac{\exp[\langle \mathbf{P}_i, \mathrm{Readout}^{(s)}(\mathcal{N}(v_i))\rangle + g_s]/\tau}{\sum_{s'=1}^{K} \exp[\langle \mathbf{P}_i, \mathrm{Readout}^{(s')}(\mathcal{N}(v_i))\rangle + g_{s'}]/\tau}$$

Local context of v_i

Sample the aspect that gives the highest value

Probability of v_i being selected as aspect s given its context $N(v_i)$



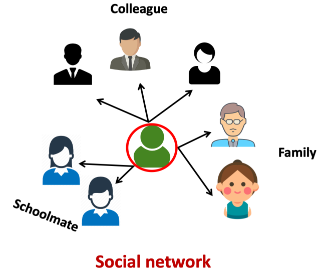
[KDD20 Park et al] Unsupervised Differentiable Multi-aspect Node Representation Learning

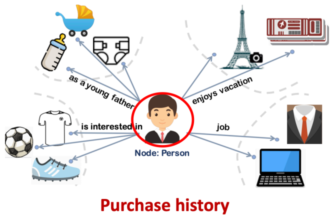
Summary of Contributions

- Developed tools for **mining meaningful knowledge** from **multi-modal** and **multi-aspect** user behavior data

Knowledge representation

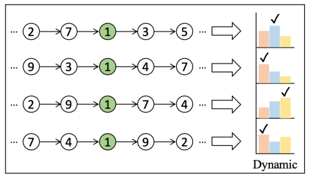
Represent user behavior based on network structures

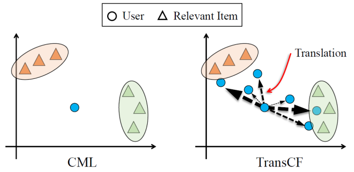


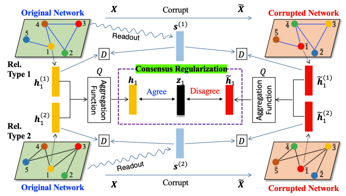


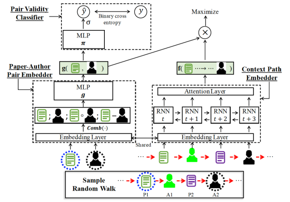
Information Extraction

Develop network mining techniques for user behavior understanding









Outline

Part 1: Research Motivation & Background

Part 2: **Multi-modal** User Behavior Analysis

Part 3: **Multi-aspect** User Behavior Analysis

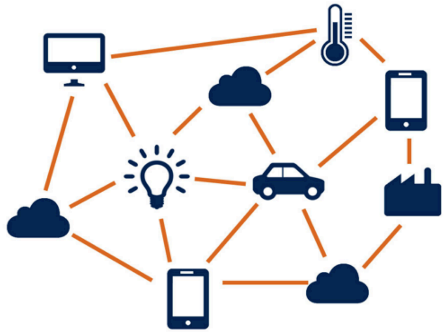
Part 4: Vision for the future ✓

Research Agenda

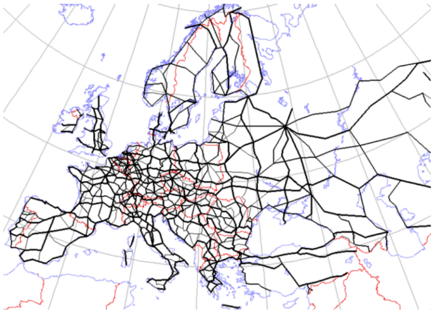
- **Research Philosophy**
  - Research in data mining should be **driven by the real-world needs**
- **Research Goal**
  - Building **practical** artificial intelligent solutions with potential for **impact** in the real-world

Network as a General Framework

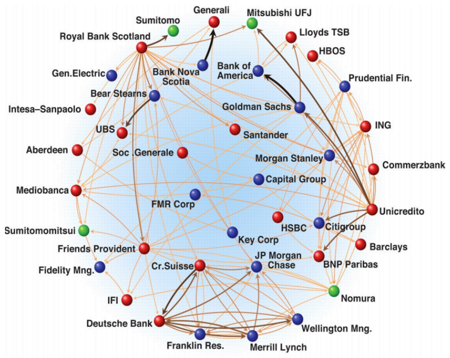
- **Current research**: **User behavior** as a network
- **Future direction:** Network as a **General Framework**
- Our world is more closely connected than we think
- Network is **a general yet powerful framework to represent complex relationships** in reality
  - **Any type of relations** between **any type of entities (+ optionally features)**



Internet-of-Things

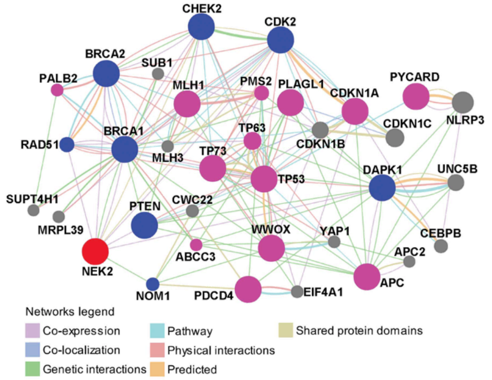


Road network



Financial network

Logistic network

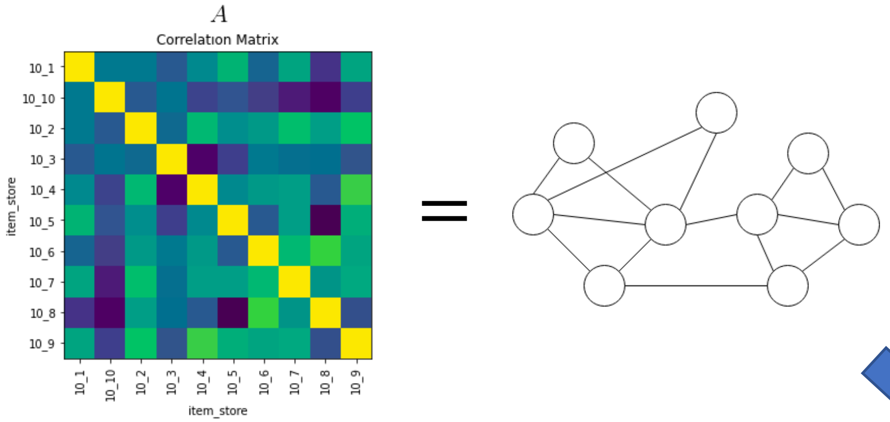


Gene network

Many problems in our real-life can be modeled as machine learning tasks **over large networks**

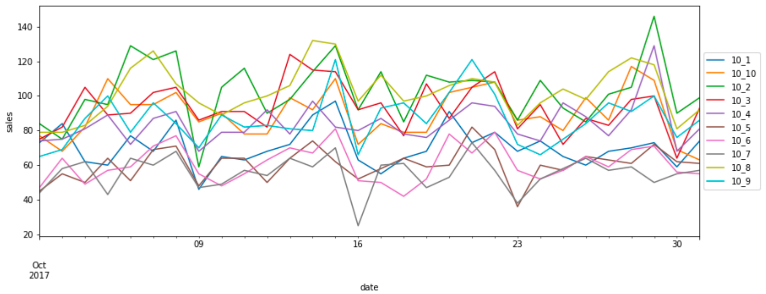
Network mining in **Retail / Manufacturing Industry**

- Demand forecasting
- Sales forecasting
- Anomaly detection in sensor stream



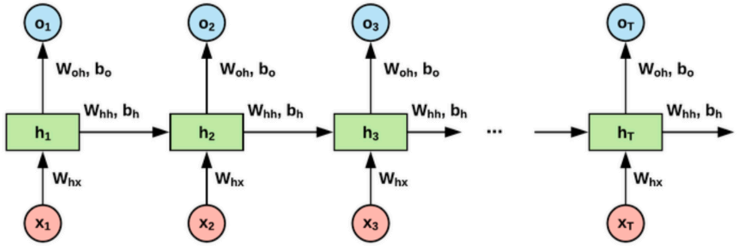
Captures relationship between stores
(Global view)

GNN



Sales of an item for multiple stores

Predict future sales



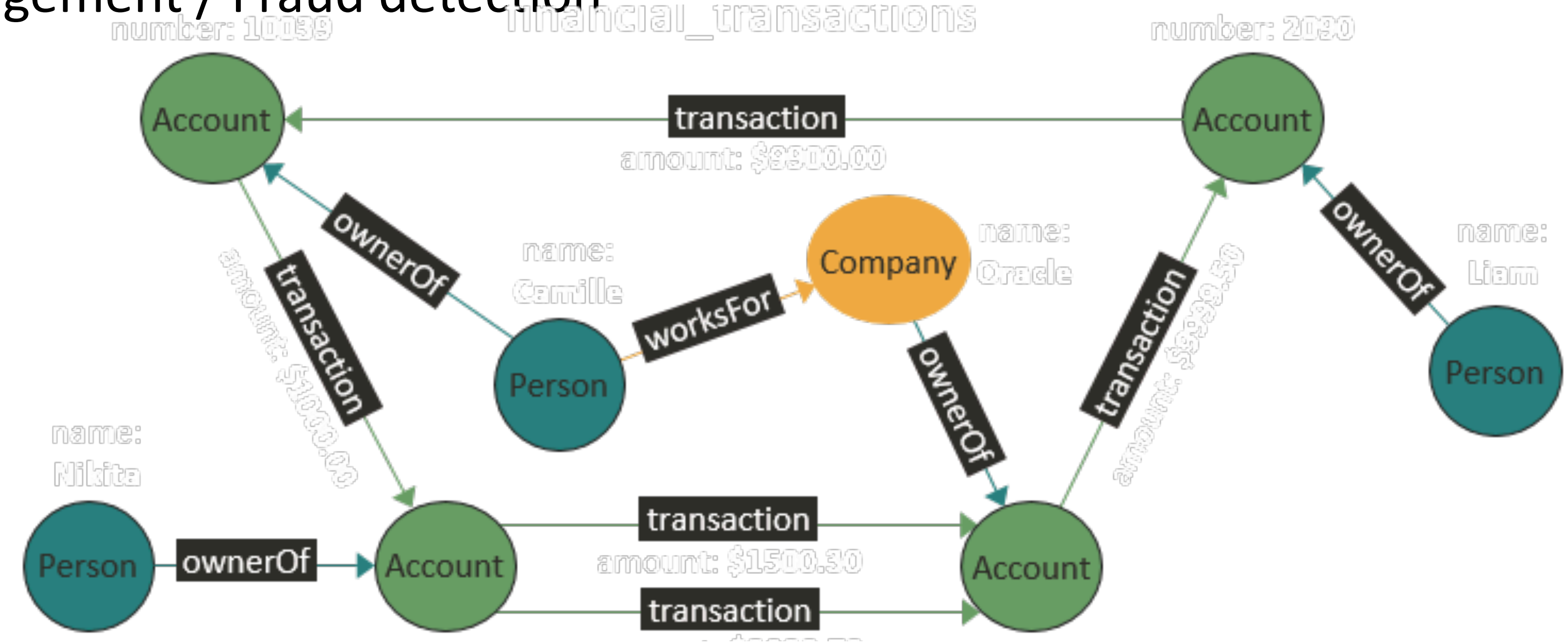
RNN

Captures sequential pattern
(Local view)

Use a **graphical representation of time series** to predict future

Network mining in **Finance**

- Stock market prediction
- Risk management / Fraud detection

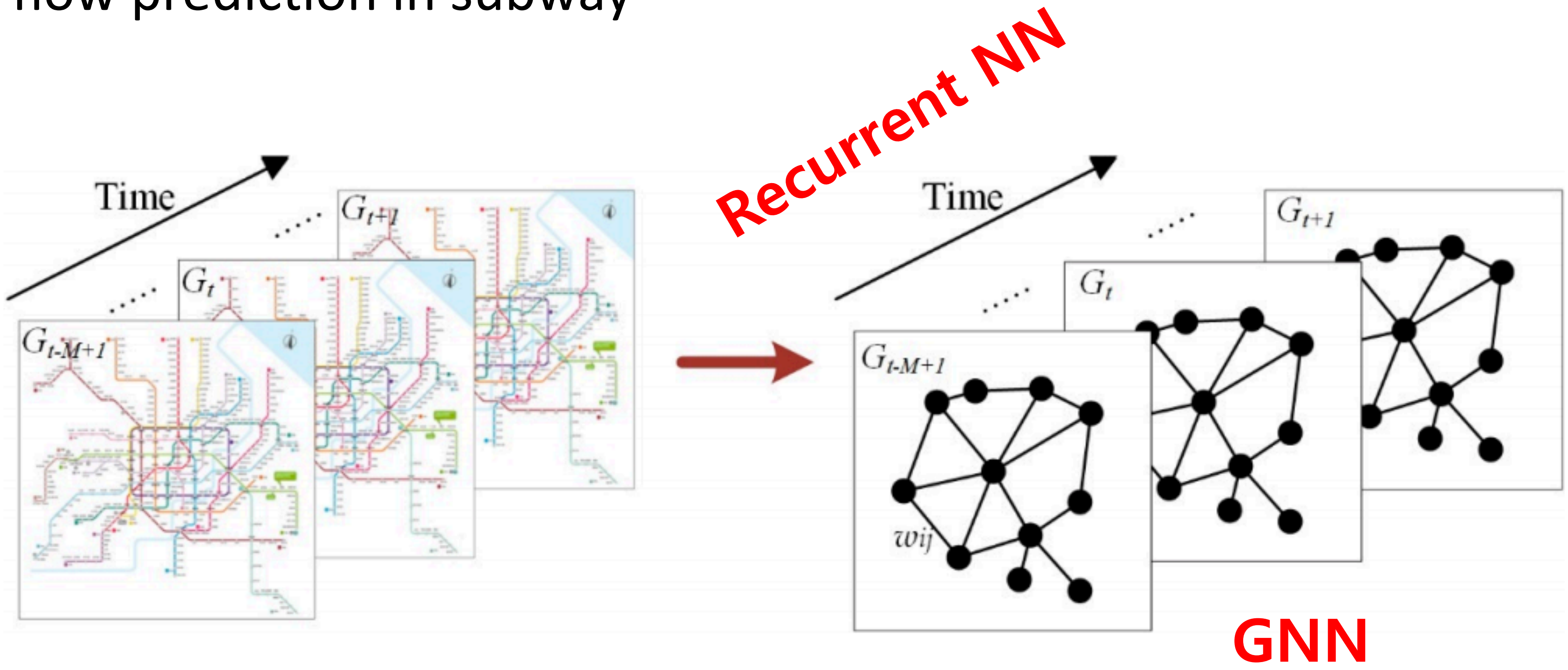


Fraud detection in heterogeneous information network

https://pgql-lang.org/spec/1.2/

Network mining in **Urban Computing**

- Bike flow prediction
- Taxi demand prediction / Ride-hailing demand forecasting (Uber, Lyft)
- Passenger flow prediction in subway



Spatio-temporal GNN

Vision for the Future: Overview

Short Term

In-depth User Behavior Analysis

- Malicious user behavior
- Evolving user behavior
- POI recommendation
- Influence maximization
- ⋮

Broader Applications in Various Domains

- Retail
- Manufacturing
- Finance
- Urban Computing
- ⋮

Long Term

Advancing Fundamentals on Network Algorithms

Robustness	Flexibility
Data Sparsity	Interpretability
Fairness	Security & Privacy
Scalability	Visualization
⋮	⋮

Expertise in multi-modal data mining using network–based technology

→ Solid foundation for building practical solutions **across various disciplines**

Thank you!
Questions?