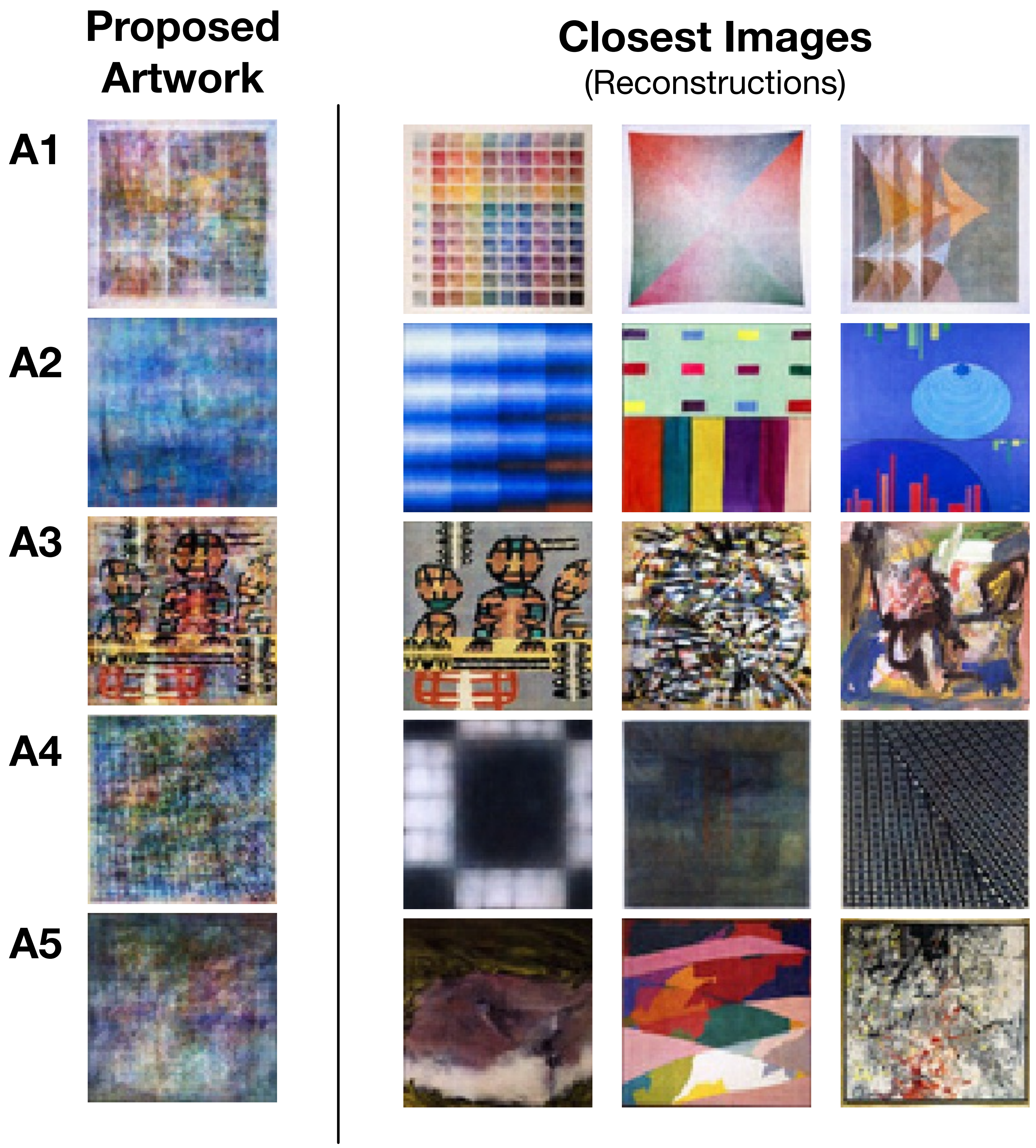


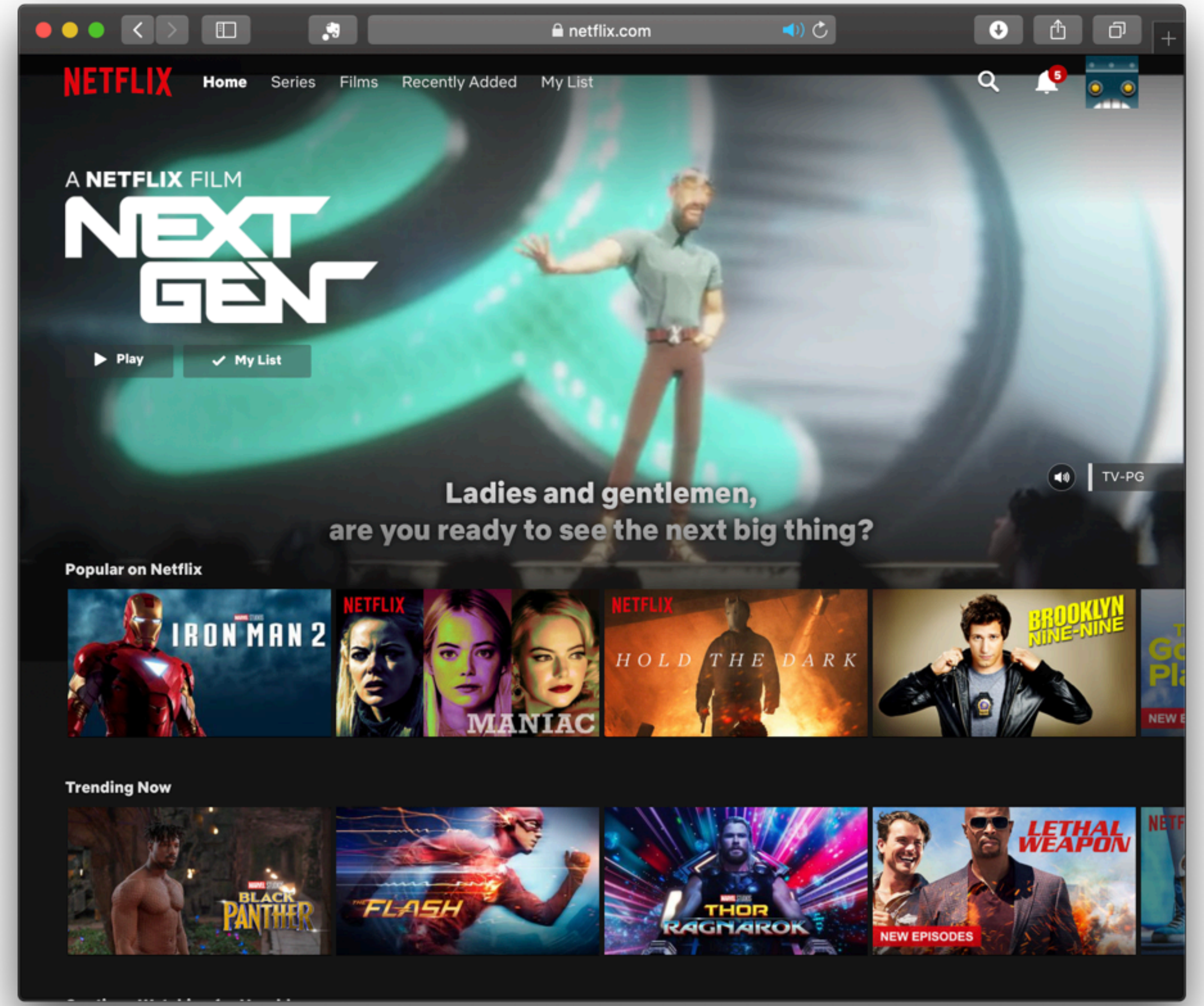
ACM RECSYS 2018

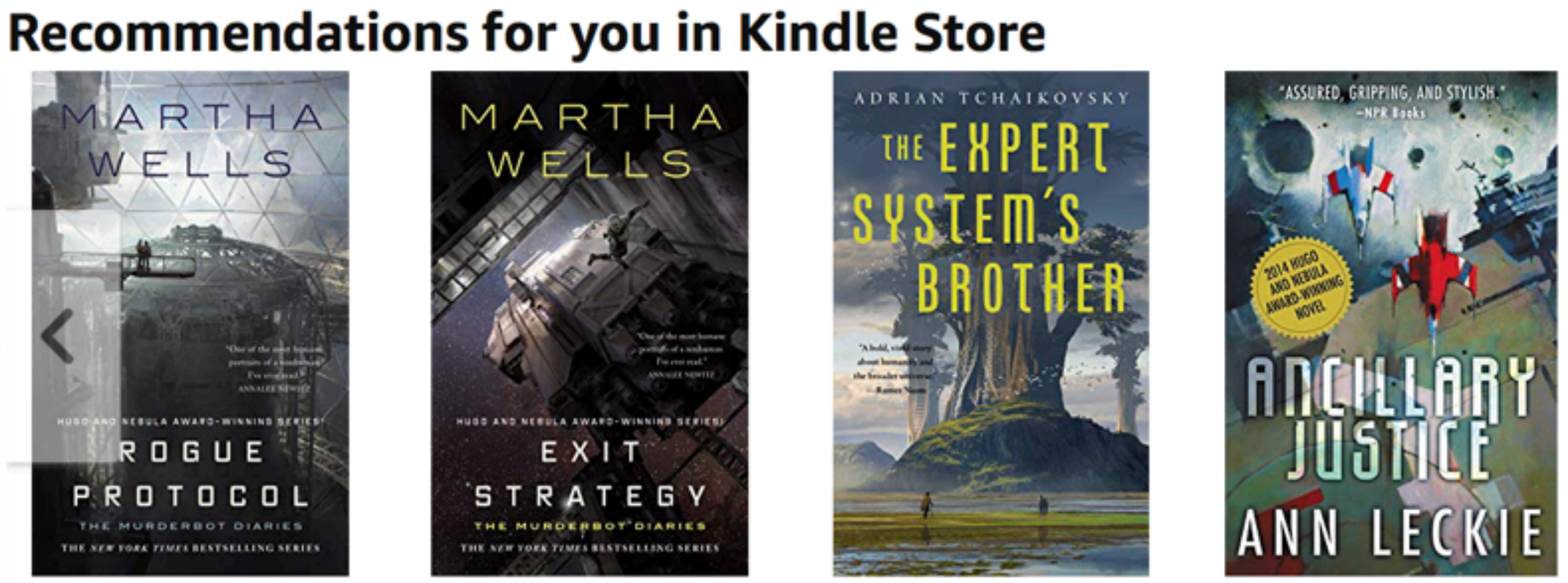
GENERATION MEETS RECOMMENDATION:

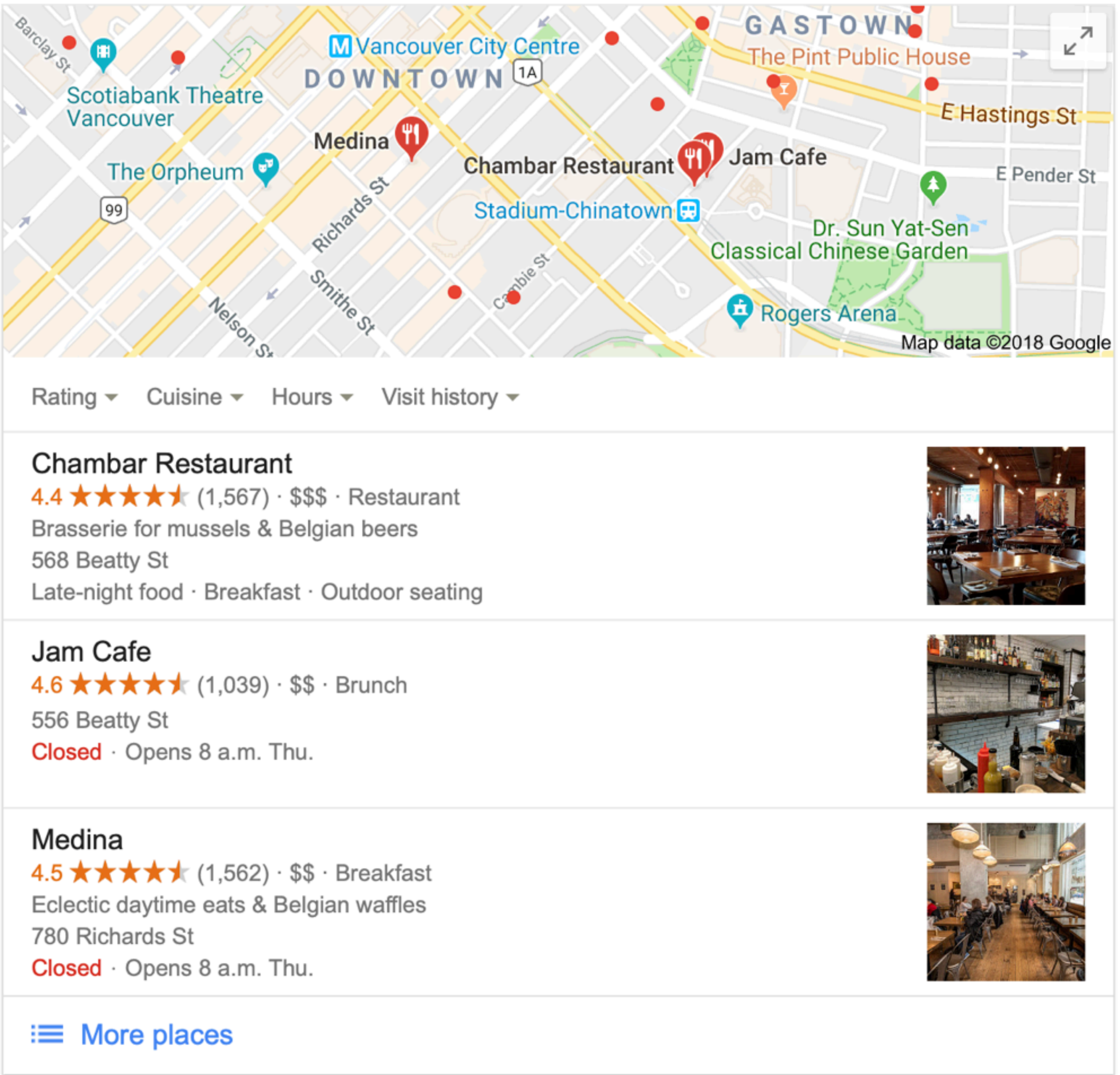
Proposing Novel Items for Groups of Users

THANH VINH VO & **HAROLD SOH**
DEPT. OF COMPUTER SCIENCE
SCHOOL OF COMPUTING
NATIONAL UNIVERSITY OF SINGAPORE
HAROLD@COMP.NUS.EDU.SG

Recommender Systems for Consumers







Recommender Systems for ~~Consumers~~ Makers

Recommender Systems for ~~Consumers~~ Makers

NUS National University of Singapore
School of Computing

RECSYS | Vancouver, BC | 2018


Generation Meets Recommendation: Proposing Novel Items for Groups of Users

Thanh Vinh Vo
National University of Singapore
vinhvt@comp.nus.edu.sg

Harold Soh
National University of Singapore
harold@comp.nus.edu.sg

ABSTRACT

Consider a movie studio aiming to produce a set of new movies for summer release: What types of movies it should produce? Who would the movies appeal to? How many movies should it make? Similar issues are encountered by a variety of organizations, e.g., mobile-phone manufacturers and online magazines, who have to create new (non-existent) items to satisfy groups of users with different preferences. In this paper, we present a joint problem formalization of these interrelated issues, and propose generative methods that address these questions simultaneously. Specifically, we leverage on the latent space obtained by training a deep generative model—the Variational Autoencoder (VAE)—via a loss function that incorporates both rating performance and item reconstruction terms. We use a greedy search algorithm that utilize this learned latent space to jointly obtain K plausible new items, and user groups that would find the items appealing. An evaluation of our methods on a synthetic dataset indicates that our approach is able to generate novel items similar to highly-desirable unobserved items. As case studies on real-world data, we applied our method on the MART abstract art and Movielens Tag Genome datasets, which resulted in promising results: small and diverse sets of novel items.

CCS CONCEPTS

• Information systems → Recommender systems; • Computing methodologies → Learning latent representations;

KEYWORDS

Novel Item Recommendation; Deep Generative Models; Group Recommendation; Group Formation

ACM Reference Format:
Thanh Vinh Vo and Harold Soh. 2018. Generation Meets Recommendation: Proposing Novel Items for Groups of Users. In Twelfth ACM Conference on Recommender Systems (RecSys '18), October 2–7, 2018, Vancouver, BC, Canada. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3240323.3240357

1 INTRODUCTION

Automated recommenders [31] have become a mainstay of modern information retrieval and e-commerce systems. By suggesting

Figure 1: An Overview of Generating New (Non-Existent) Items for Groups of Users. Our approach relies on two related components: (i) a shared latent representation space learnt from user-item rating data and (ii) a pair of functions: the encoder $f_e(x)$ that maps item features x to their respective latent representations/embeddings z, and a decoder $g_e(z)$ that performs the inverse operation, i.e., maps embeddings z to item features x. We exploit this space to obtain novel items with high predicted desirability through greedy weighted maximum coverage, and then generate item features via the decoder.

related items (e.g., news articles [24], music [5], movies [37]), recommender systems filter away irrelevant data, enabling users to efficiently locate items of interest. However, current work has focussed primarily on providing recommendations of *existing items* to consumers (either personalized to individuals or groups [29]). In this work, we take on the problem of recommending or proposing *new* (*non-existent*) items that could be *created* to satisfy users.

New item recommendation would be beneficial for item *makers*—designers, producers, and manufacturers—who generate novel items for the marketplace. For example, mobile phone makers often produce a set of phone models because different users have different preferences; one group of users may prefer a phone with a large touch-screen, whilst another group may prefer a smaller, lighter, device. For item makers, important questions include: Is it possible to craft a set of new phone models that appeal to the widest possible number of customers? Which users would each phone design attract? And how large is each of these user groups? These questions are currently *not* answered by existing recommender systems research and may seem intractable at first glance.

In this paper, we take a first step towards a new class of recommender systems: item recommendation for item makers. We propose a principled approach for addressing the aforementioned questions: first, we formalize the problem and show it to be a variant of maximum coverage, which is NP-hard in general. We then

Recommender Systems for ~~Consumers~~ Makers

Generation Meets Recommendation: Proposing Novel Items for Groups of Users

Thanh Vinh Vo
National University of Singapore
vinhvt@comp.nus.edu.sg

Harold Soh
National University of Singapore
harold@comp.nus.edu.sg

ABSTRACT

Consider a movie studio aiming to produce a set of *new* movies for summer release: What types of movies it should produce? Who would the movies appeal to? How many movies should it make? Similar issues are encountered by a variety of organizations, e.g., mobile-phone manufacturers and online magazines, who have to create new (non-existent) items to satisfy groups of users with different preferences. In this paper, we present a joint problem formalization of these interrelated issues, and propose generative methods that address these questions simultaneously. Specifically, we leverage on the latent space obtained by training a deep generative model—the Variational Autoencoder (VAE)—via a loss function that incorporates both rating performance and item reconstruction terms. We use a greedy search algorithm that utilize this learned latent space to jointly obtain K plausible new items, and user groups that would find the items appealing. An evaluation of our methods on a synthetic dataset indicates that our approach is able to generate novel items similar to highly-desirable unobserved items. As case studies on real-world data, we applied our method on the MART abstract art and Movielens Tag Genome datasets, which resulted in promising results: small and diverse sets of novel items.

CCS CONCEPTS

• **Information systems → Recommender systems**; • **Computing methodologies → Learning latent representations**;

KEYWORDS

Novel Item Recommendation; Deep Generative Models; Group Recommendation; Group Formation

ACM Reference Format:

Thanh Vinh Vo and Harold Soh. 2018. Generation Meets Recommendation: Proposing Novel Items for Groups of Users. In *Twelfth ACM Conference on Recommender Systems (RecSys '18), October 2–7, 2018, Vancouver, BC, Canada.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3240323.3240357

1 INTRODUCTION

Automated recommenders [31] have become a mainstay of modern information retrieval and e-commerce systems. By suggesting

ACM ISBN 978-1-4503-5901-6/18/10...$15.00
https://doi.org/10.1145/3240323.3240357

Figure 1: An Overview of Generating New (Non-Existent) Items for Groups of Users. Our approach relies on two related components: (i) a shared latent representation space learnt from user-item rating data and (ii) a pair of functions: the encoder $f_s(\mathbf{x})$ that maps item features $\mathbf{x}$ to their respective latent representations/embeddings $\mathbf{z}$, and a decoder $g_s(\mathbf{z})$ that performs the inverse operation, i.e., maps embeddings $\mathbf{z}$ to item features $\mathbf{x}$. We exploit this space to obtain novel items with high predicted desirability through greedy weighted maximum coverage, and then generate item features via the decoder.

related items (e.g., news articles [24], music [5], movies [37]), recommender systems filter away irrelevant data, enabling users to efficiently locate items of interest. However, current work has focussed primarily on providing recommendations of *existing items* to consumers (either personalized to individuals or groups [29]). In this work, we take on the problem of recommending or proposing *new* (*non-existent*) items that could be *created* to satisfy users.

New item recommendation would be beneficial for item *makers*—designers, producers, and manufacturers—who generate novel items for the marketplace. For example, mobile phone makers often produce a set of phone models because different users have different preferences; one group of users may prefer a phone with a large touch-screen, whilst another group may prefer a smaller, lighter, device. For item makers, important questions include: Is it possible to craft a set of new phone models that appeal to the widest possible number of customers? Which users would each phone design attract? And how large is each of these user groups? These questions are currently *not* answered by existing recommender systems research and may seem intractable at first glance.

In this paper, we take a first step towards a new class of recommender systems: item recommendation for item makers. We propose a principled approach for addressing the aforementioned questions: first, we formalize the problem and show it to be a variant of maximum coverage, which is NP-hard in general. We then

Recommender Systems for ~~Consumers~~ Makers

- **What set of products** would appeal to a wide range of people?
- **Who** would like those products?
- **How many** products should you make?

Formalized problem & developed methods to *jointly* address these Questions

Formalizing the Problem

u_1 u_2 u_3

$\ldots \in \mathcal{U}$ (Set of Users)

Formalizing the Problem

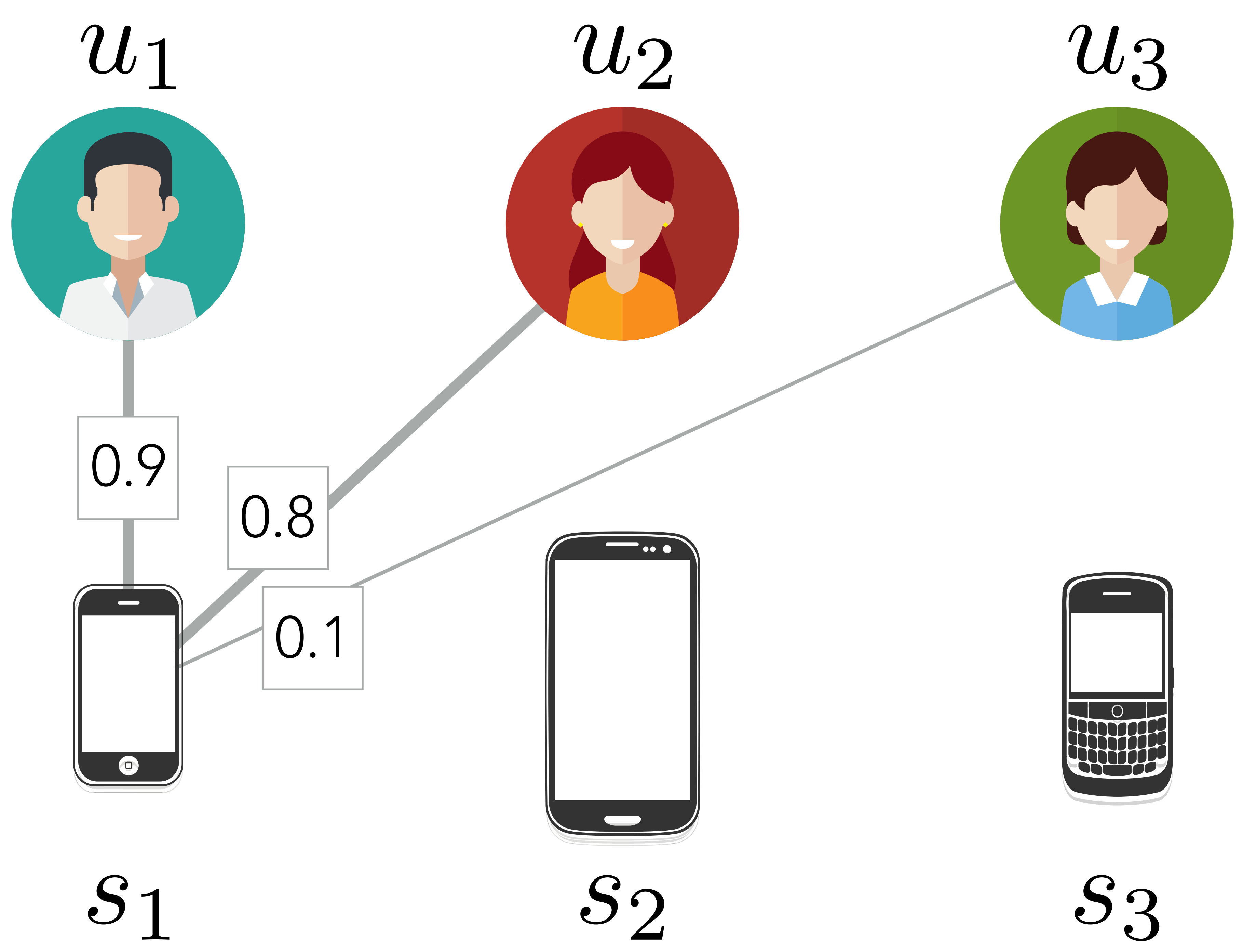
$u_1 \quad u_2 \quad u_3$

$\ldots \in \mathcal{U}$ (Set of Users)

$\ldots \in \mathcal{S}$ (Set of Items)

$s_1 \quad s_2 \quad s_3$

Formalizing the Problem

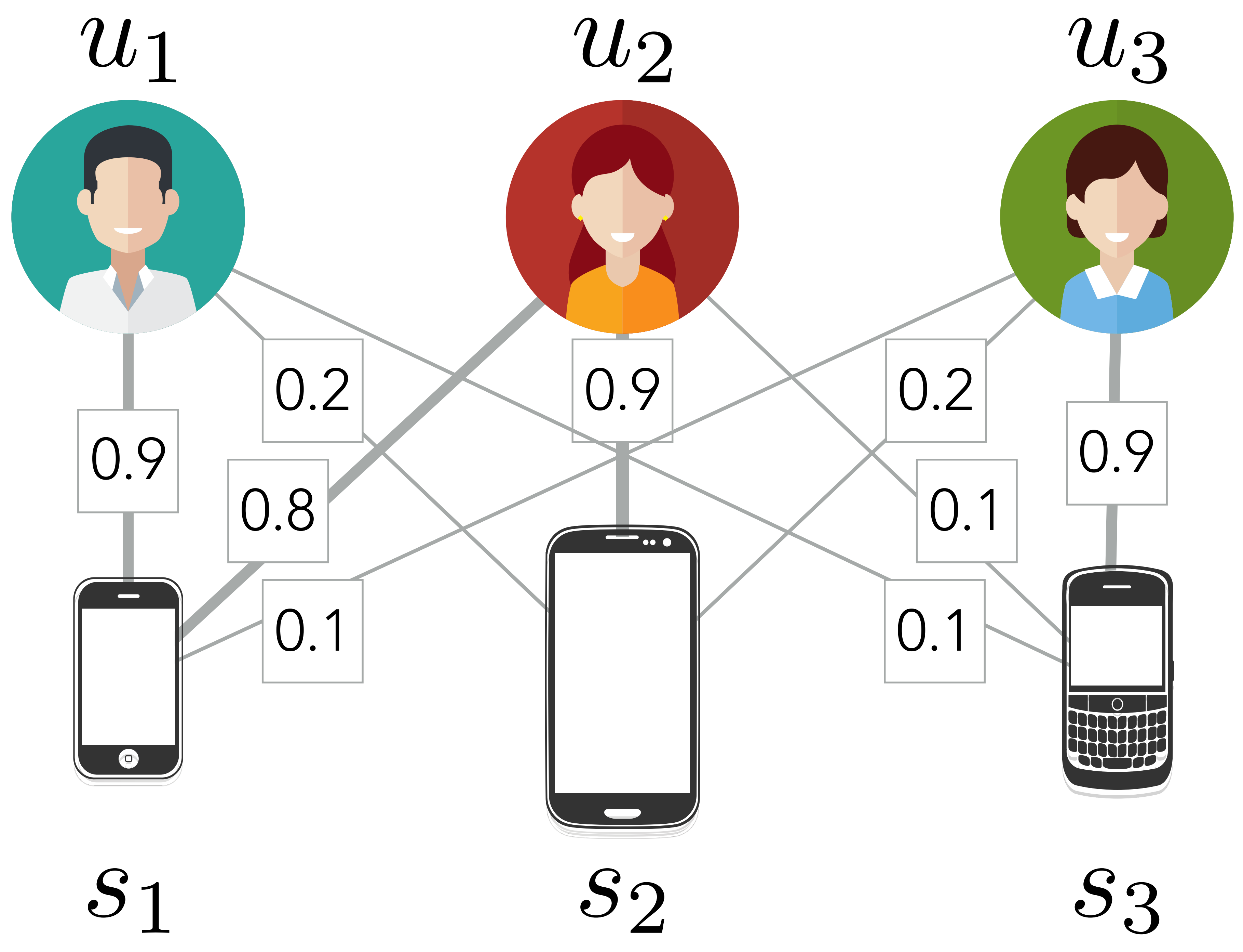


$\ldots \in \mathcal{U}$ (Set of Users)

$r(u, s)$ (Rating Function)

$\ldots \in \mathcal{S}$ (Set of Items)

Formalizing the Problem

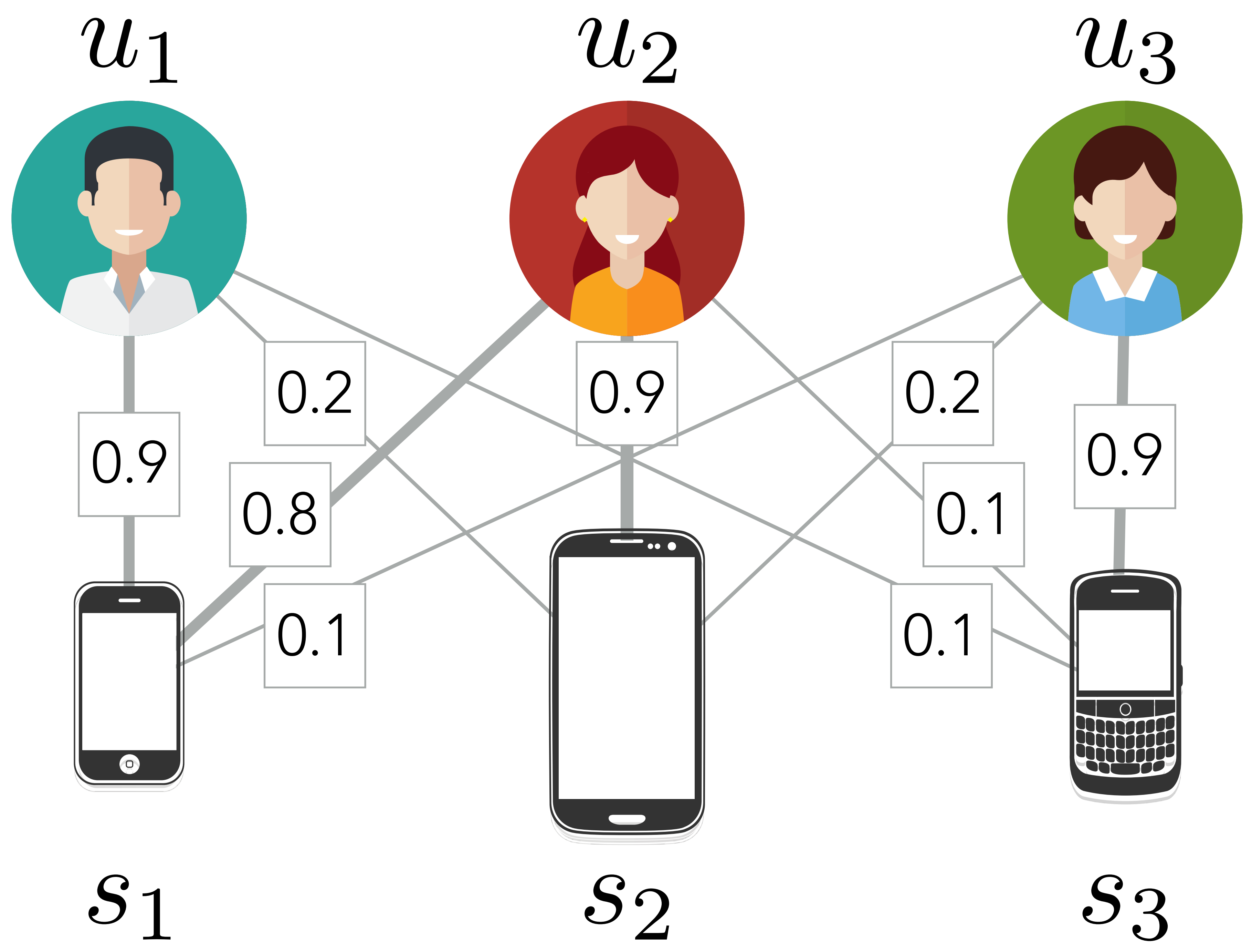


$\ldots \in \mathcal{U}$ (Set of Users)

$r(u, s)$ (Rating Function)

$\ldots \in \mathcal{S}$ (Set of Items)

Formalizing the Problem



$\ldots \quad \in \mathcal{U}$ (Set of Users)

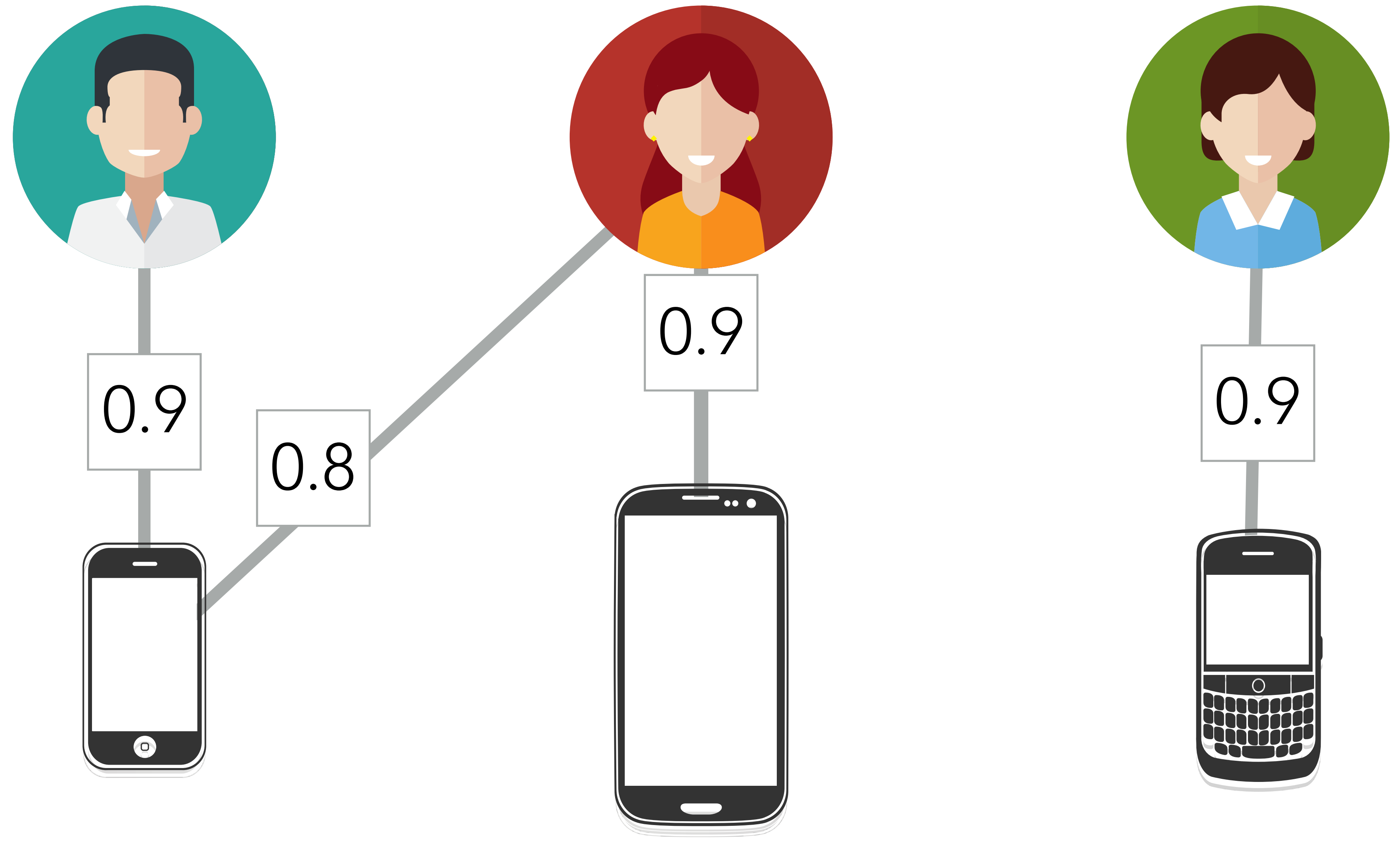
$r(u, s)$ (Rating Function)

$\ldots \quad \in \mathcal{S}$ (Set of Items)

$$\text{"Cover"} \quad c(u, s) = \begin{cases} 1 & \text{if } \; r(u, s) > \tau \quad \text{(above threshold)} \\ 0 & \text{otherwise} \end{cases}$$

Formalizing the Problem

u_1 u_2 u_3 $\dots \in \mathcal{U}$ (Set of Users)



$\tau = 0.7$

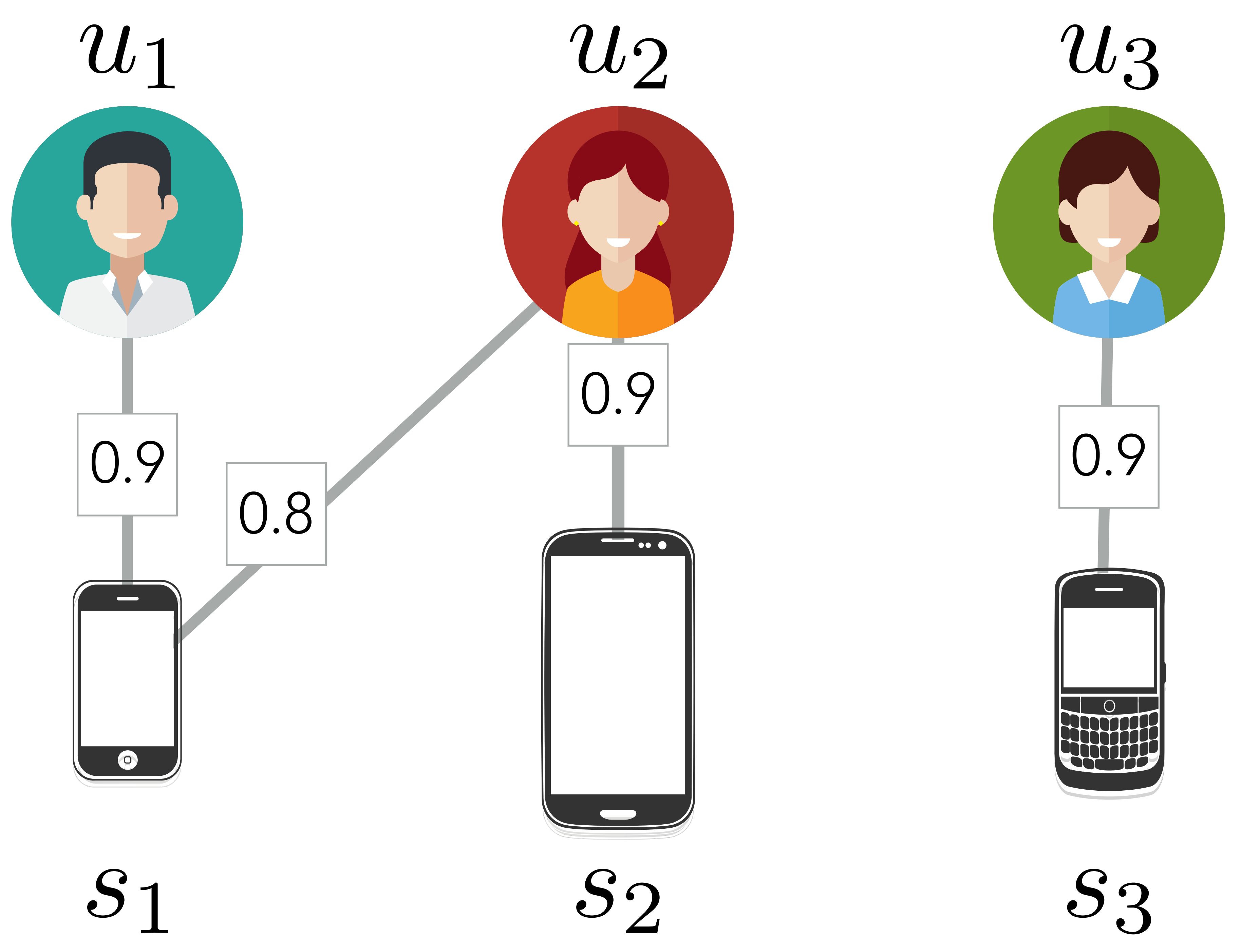
$r(u, s)$ (Rating Function)

s_1 s_2 s_3 $\dots \in \mathcal{S}$ (Set of Items)

$$\text{"Cover"} \quad c(u, s) = \begin{cases} 1 & \text{if } \; r(u, s) > \tau \quad \text{(above threshold)} \\ 0 & \text{otherwise} \end{cases}$$

Formalizing the Problem

$\tau = 0.7$



$\dots \in \mathcal{U}$ (Set of Users)

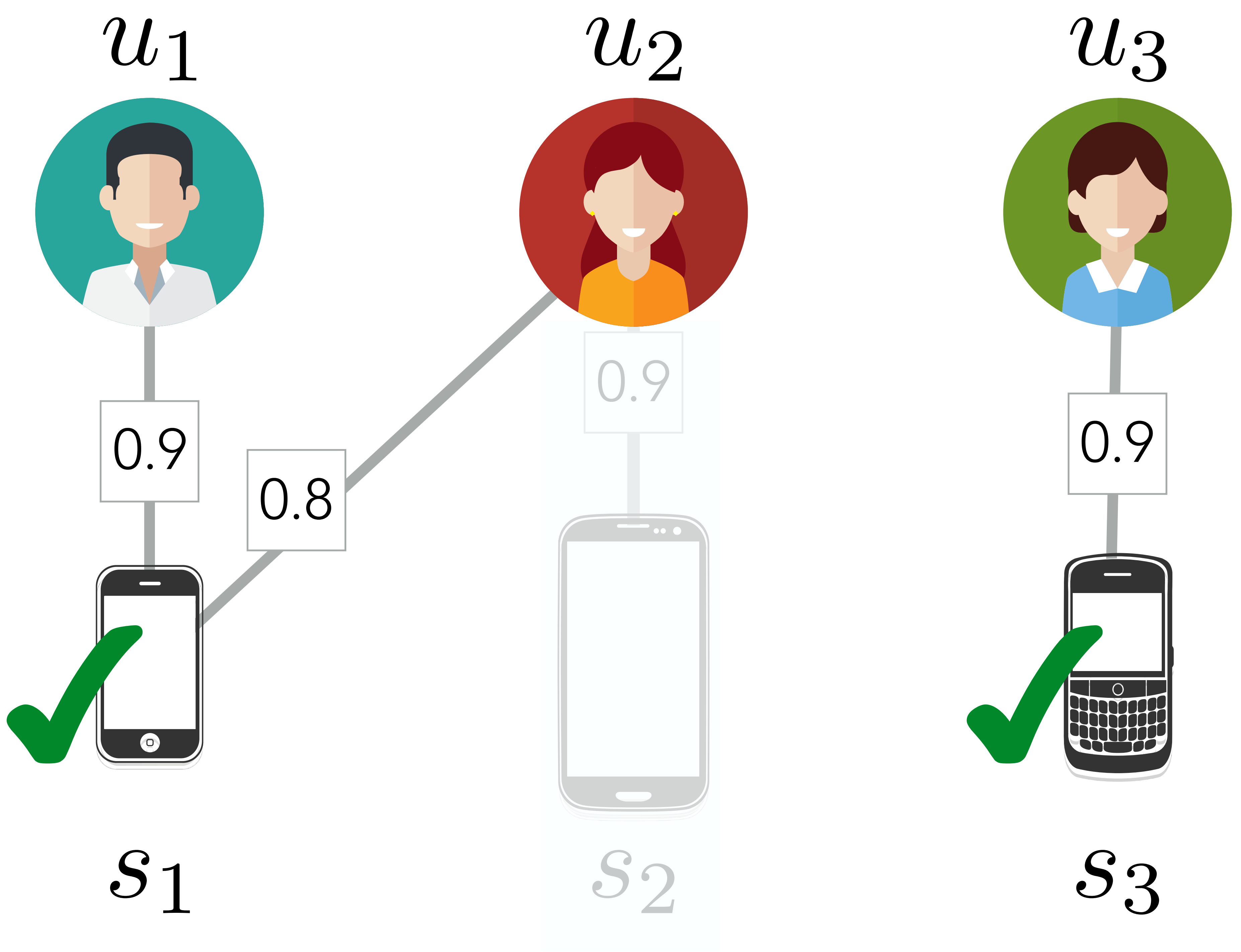
$r(u, s)$ (Rating Function)

$\dots \in \mathcal{S}$ (Set of Items)

if $k = 2$ (can only pick 2 items), which would you pick?

Formalizing the Problem

$\tau = 0.7$



$\ldots \in \mathcal{U}$ (Set of Users)

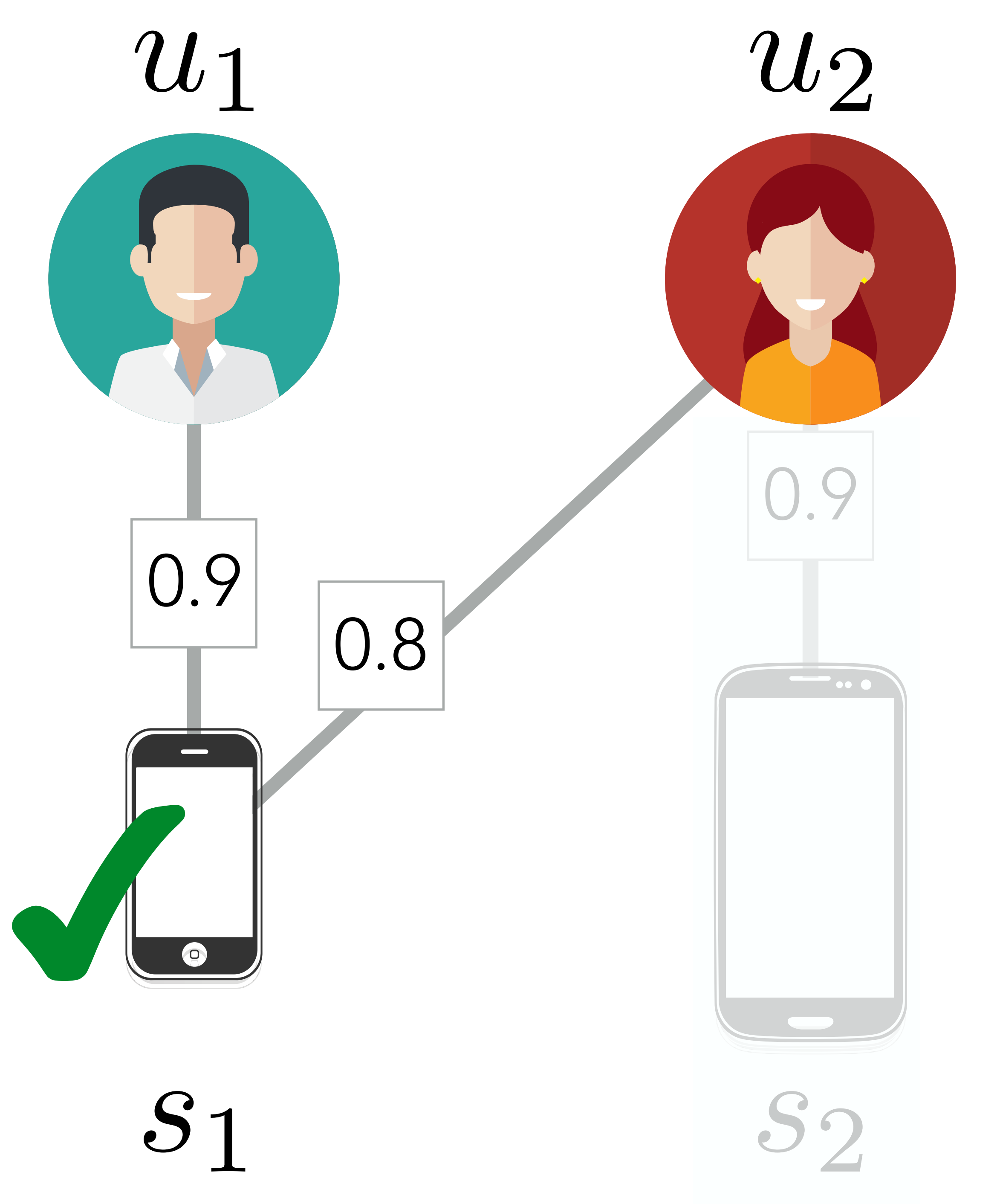
$r(u, s)$ (Rating Function)

$\ldots \in \mathcal{S}$ (Set of Items)

if $k = 2$ (can only pick 2 items), which would you pick?

Formalizing the Problem

$\tau = 0.7$



$\dots \in \mathcal{U}$ (Set of Users)

$r(u, s)$ (Rating Function)

$\dots \in \mathcal{S}$ (Set of Items)

$$\boxed{\text{argmax}_{\mathcal{V} \subseteq \mathcal{S}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} r(u, v) c_{uv}}$$

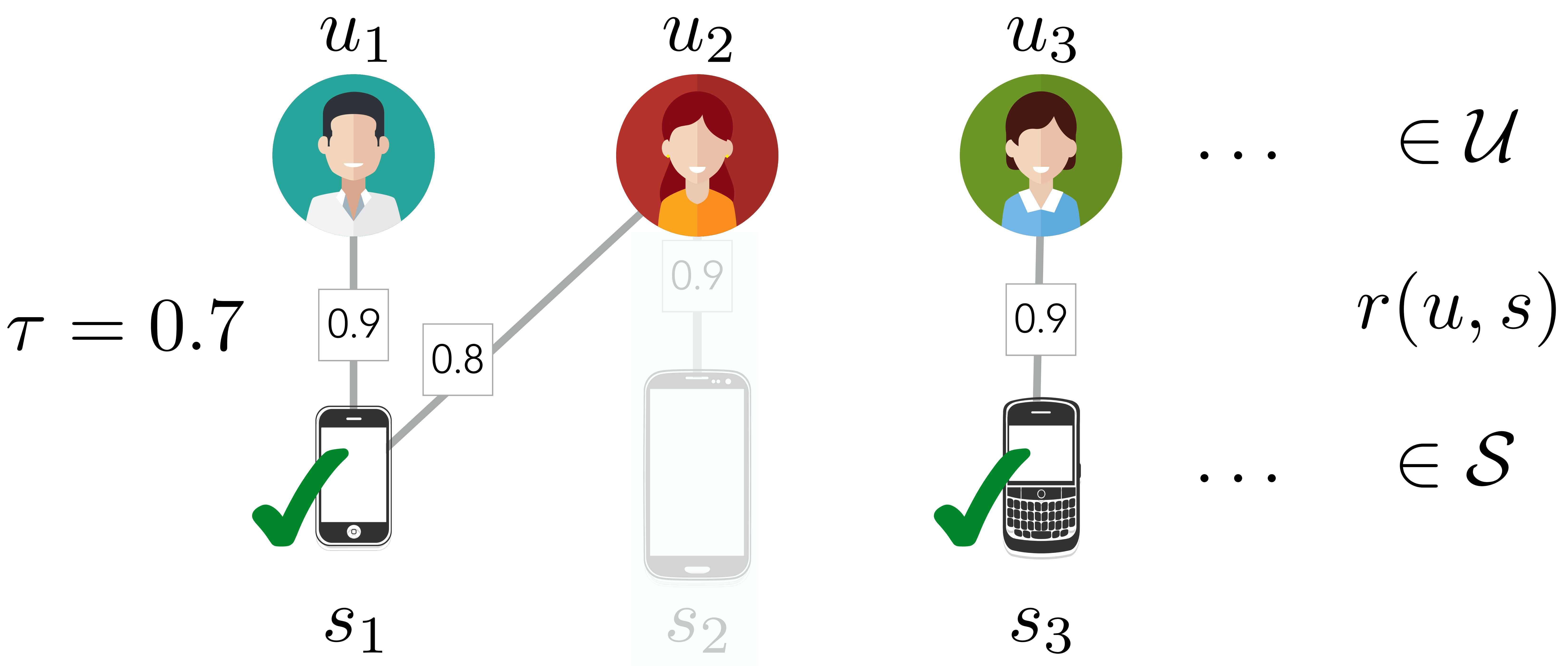
such that:

a user is only covered by one item

$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

k items selected

Formalizing the Problem



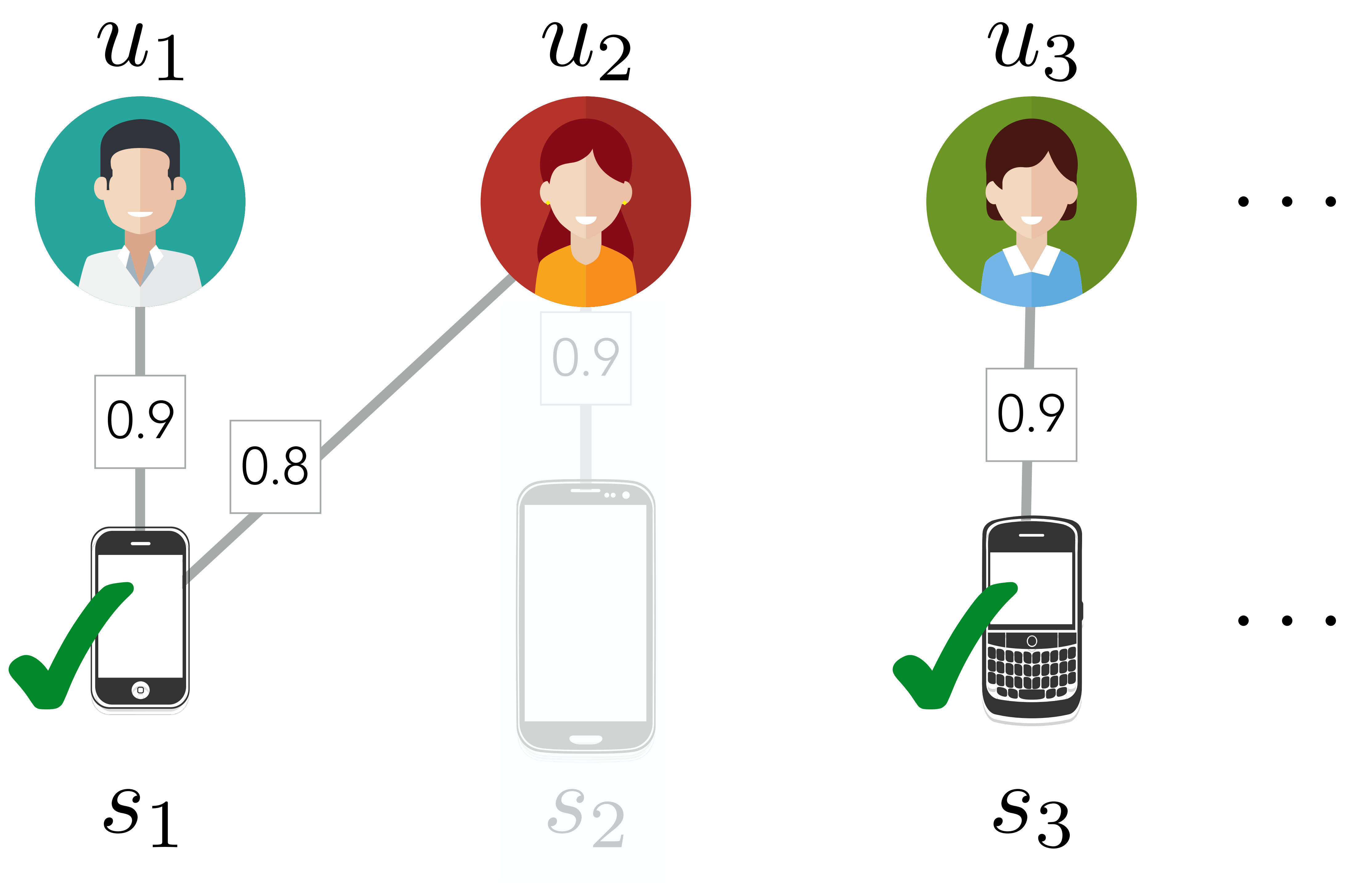
$$\operatorname{argmax}_{\mathcal{V} \subseteq \mathcal{S}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} r(u, v) c_{uv}$$

such that:

$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Formalizing the Problem

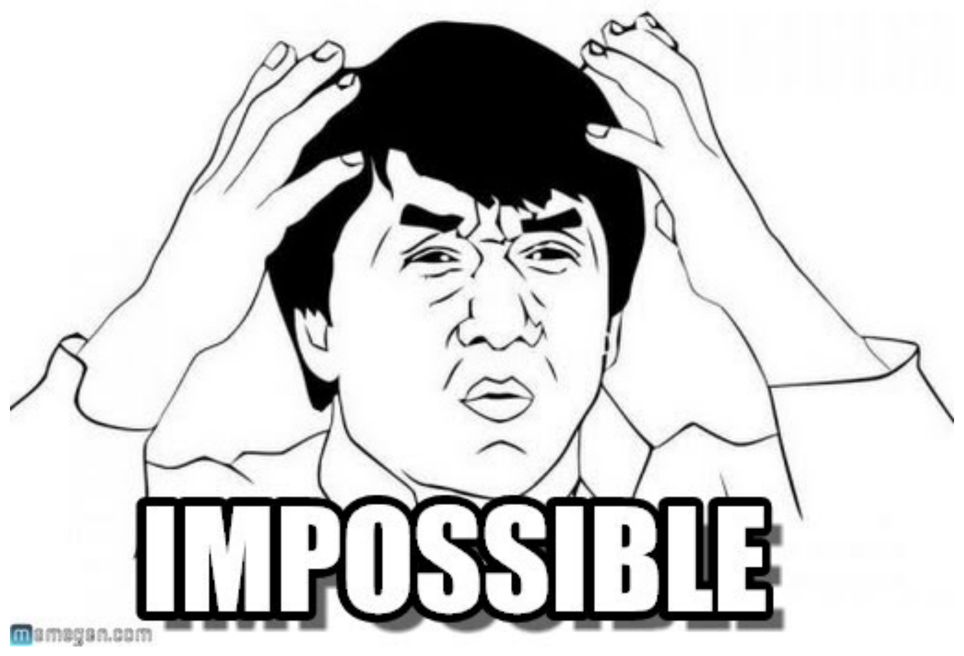
$\tau = 0.7$



$\dots \in \mathcal{U}$

$r(u, s)$

$\dots \in \mathcal{S}$

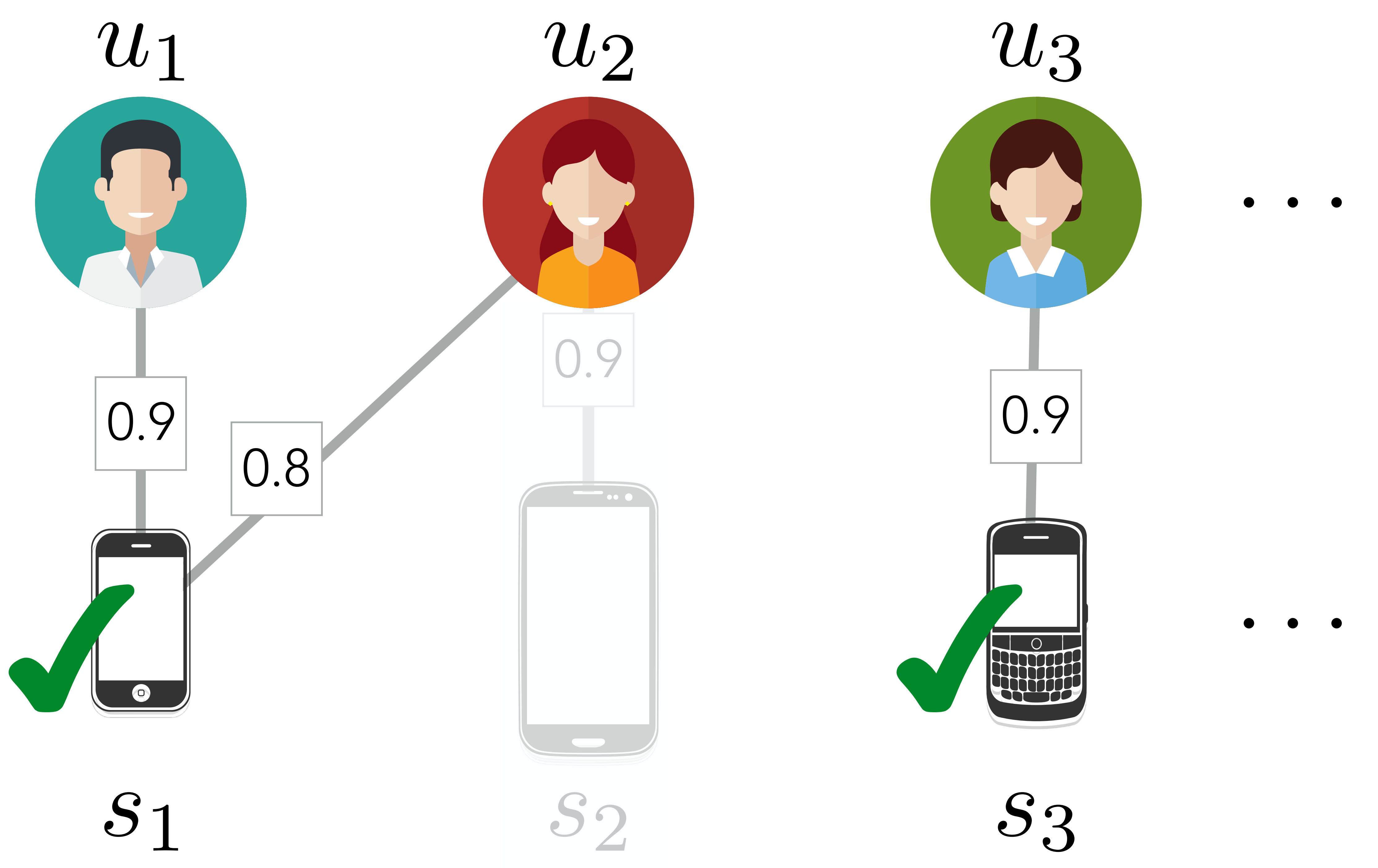


$$\boxed{\operatorname{argmax}_{\mathcal{V} \subseteq \mathcal{S}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} r(u, v) c_{uv}}$$

such that:

$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Formalizing the Problem

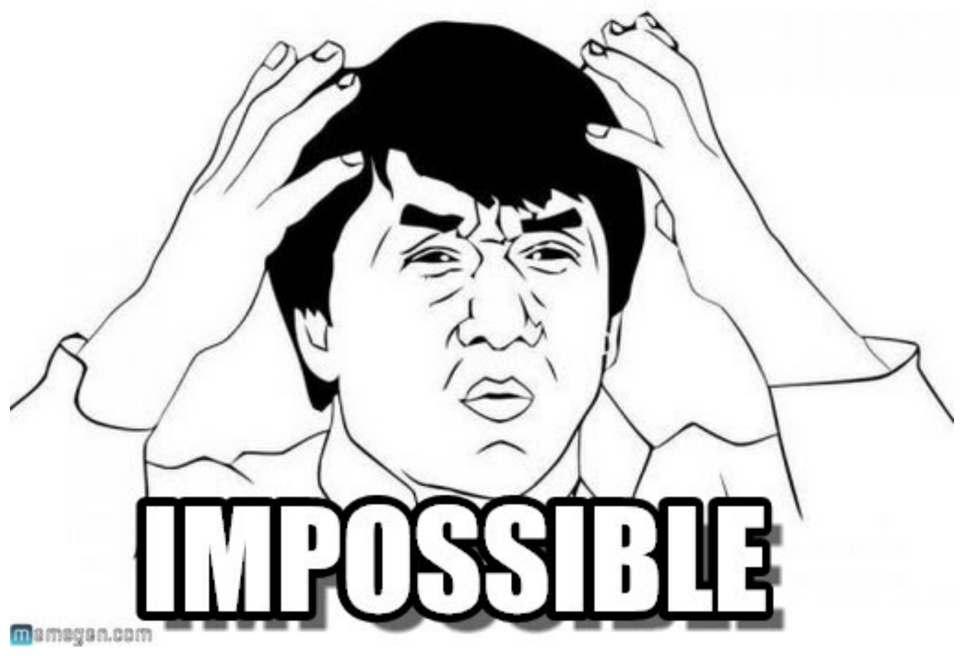


$\tau = 0.7$

$\in \mathcal{U}$

$r(u, s)$

$\in \mathcal{S}$ ← Don't know this.



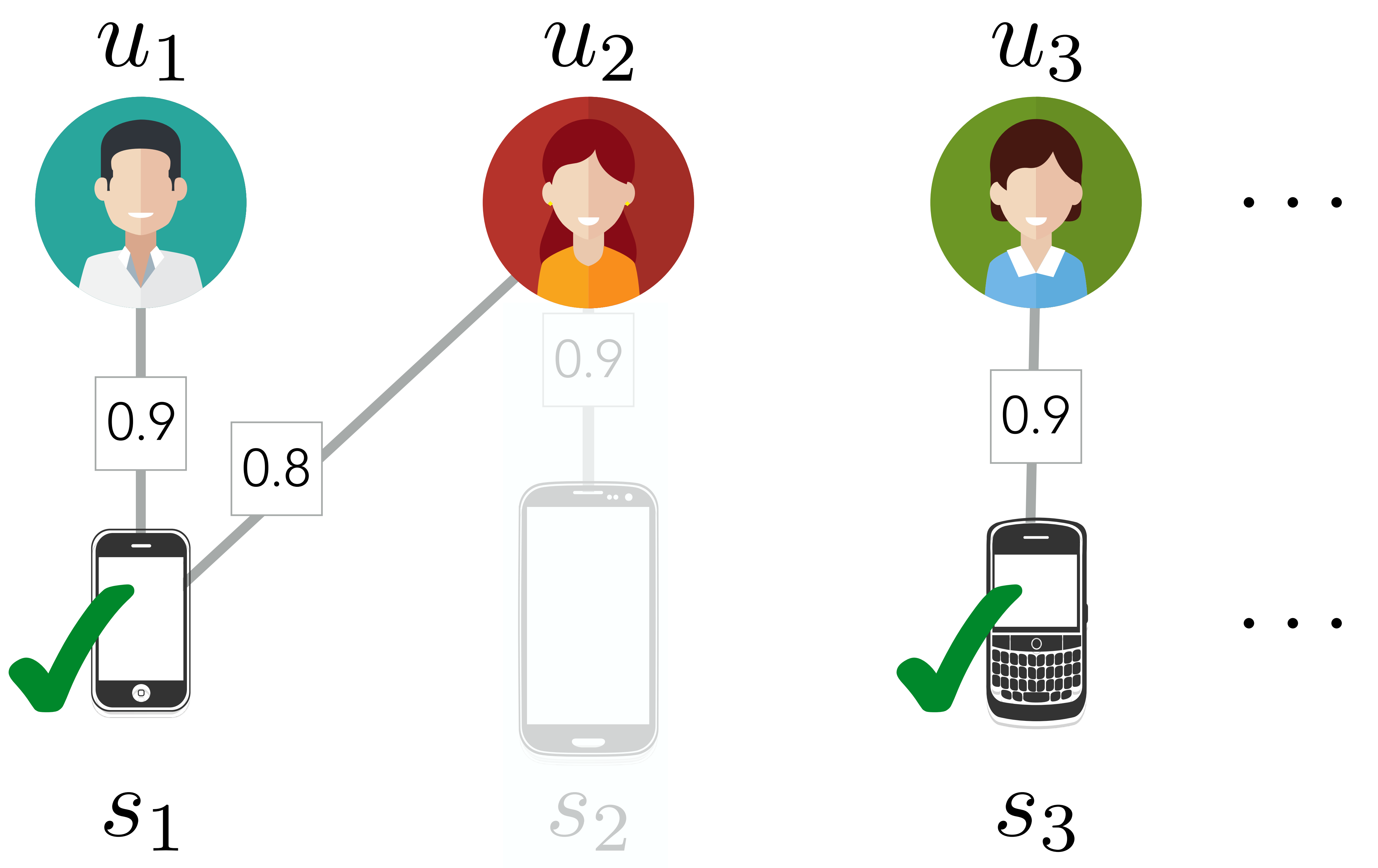
$$\text{argmax}_{\mathcal{V} \subseteq \mathcal{S}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} r(u, v) c_{uv}$$

such that:

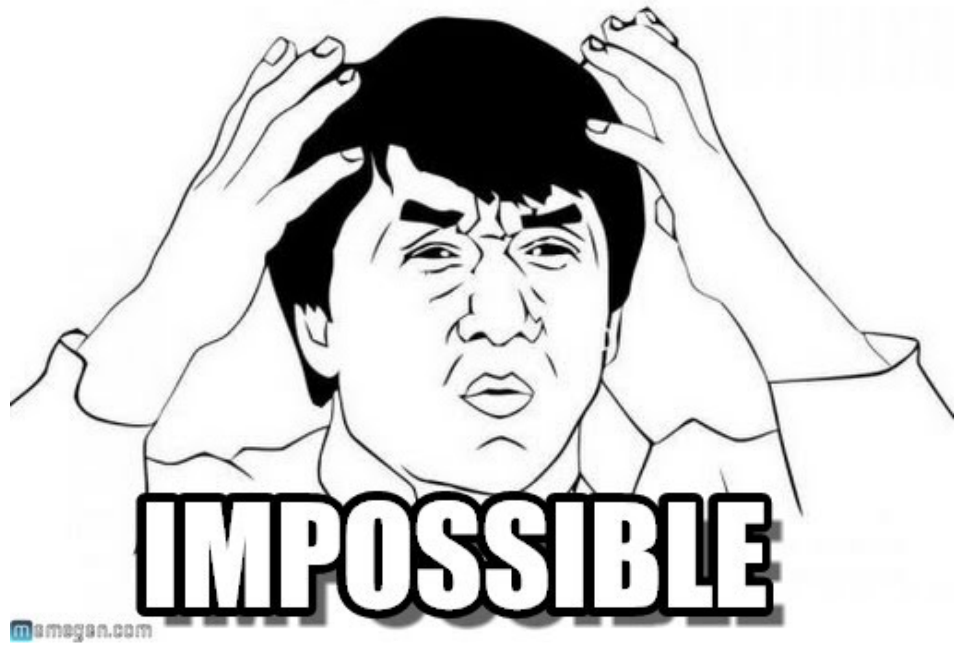
$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Formalizing the Problem

$\tau = 0.7$



$\in \mathcal{U}$



$r(u, s)$ Don't have this either.

$\in \mathcal{S}$ Don't know this.

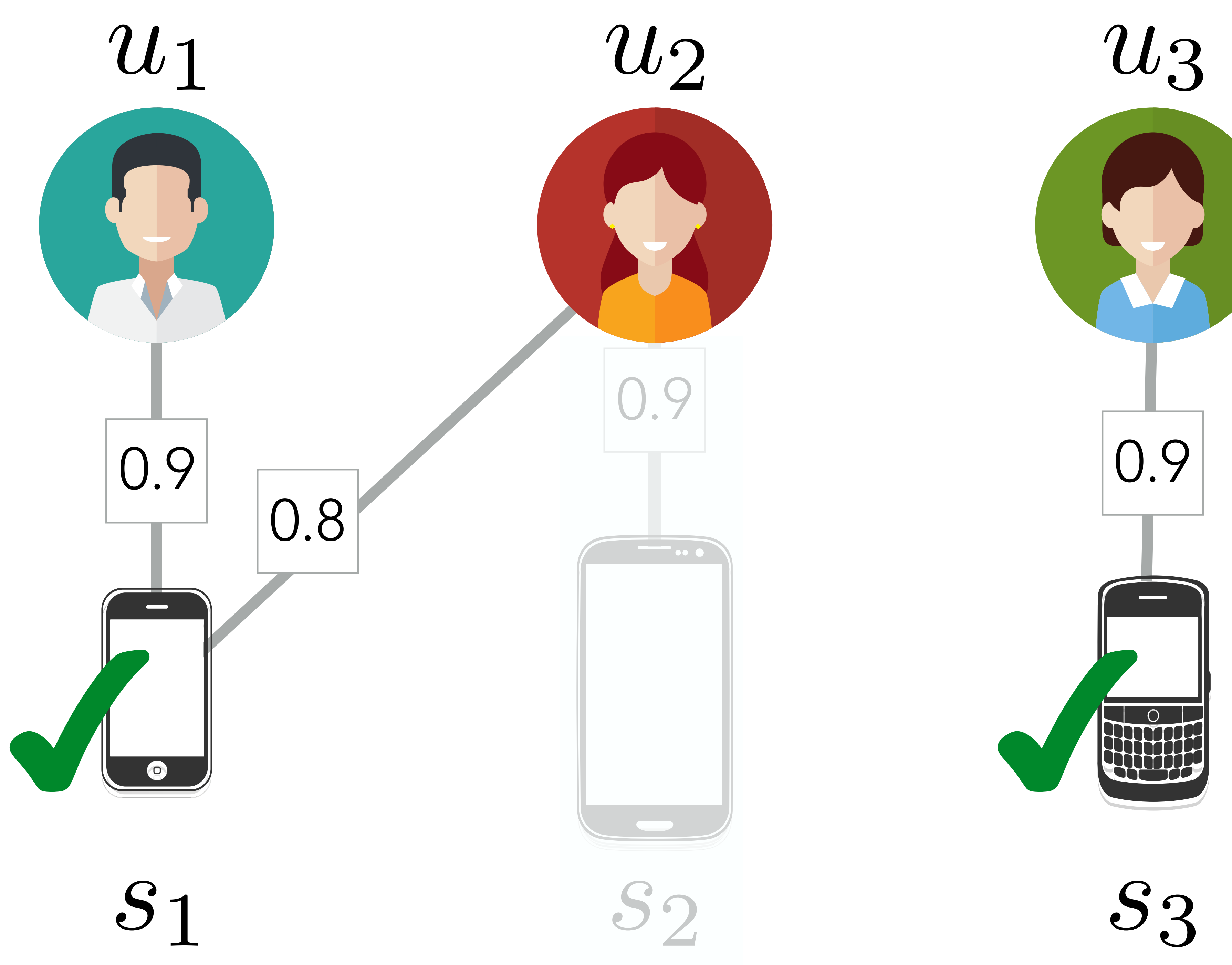
$$\text{argmax}_{\mathcal{V} \subseteq \mathcal{S}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} r(u, v) c_{uv}$$

such that:

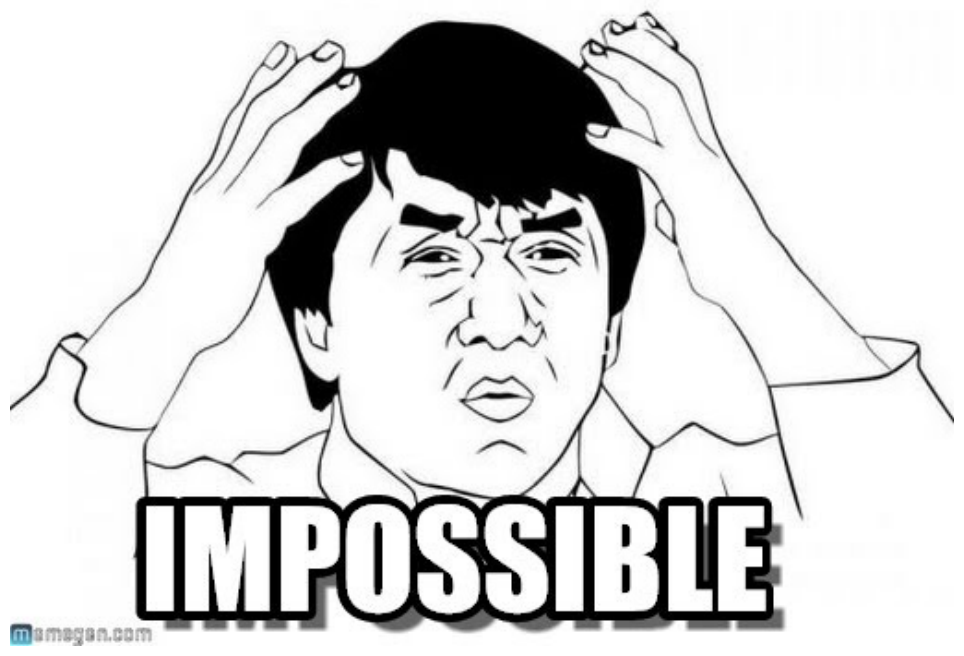
$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Formalizing the Problem

$\tau = 0.7$



$\dots \in \mathcal{U}$



$r(u, s)$ Don't have this either.

$\dots \in \mathcal{S}$ Don't know this.

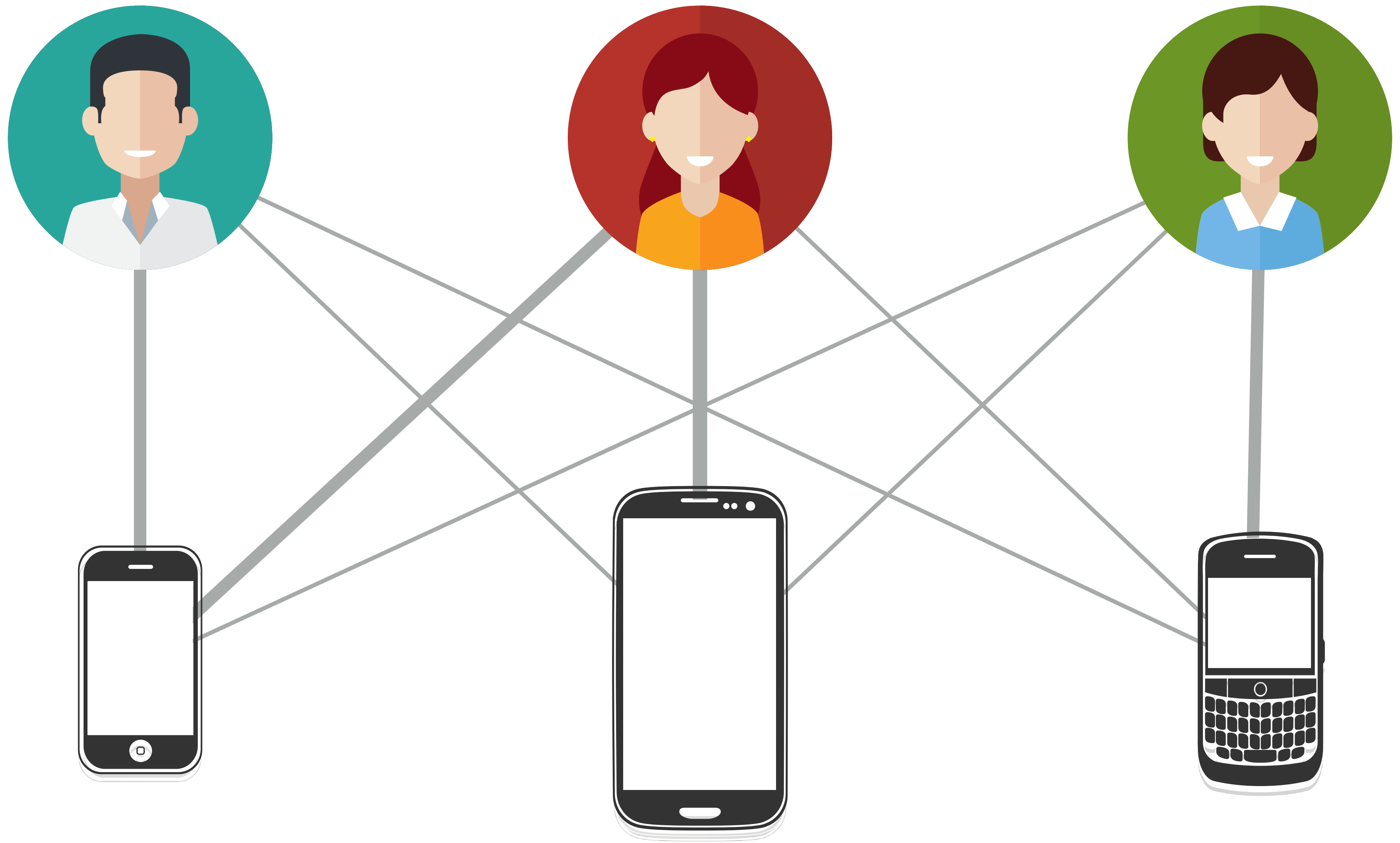
Maximum coverage over a unknown set?!

$$\boxed{\text{argmax}_{\mathcal{V} \subseteq \mathcal{S}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} r(u, v) c_{uv}}$$

such that:

$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Re-Formalizing with a Latent Space



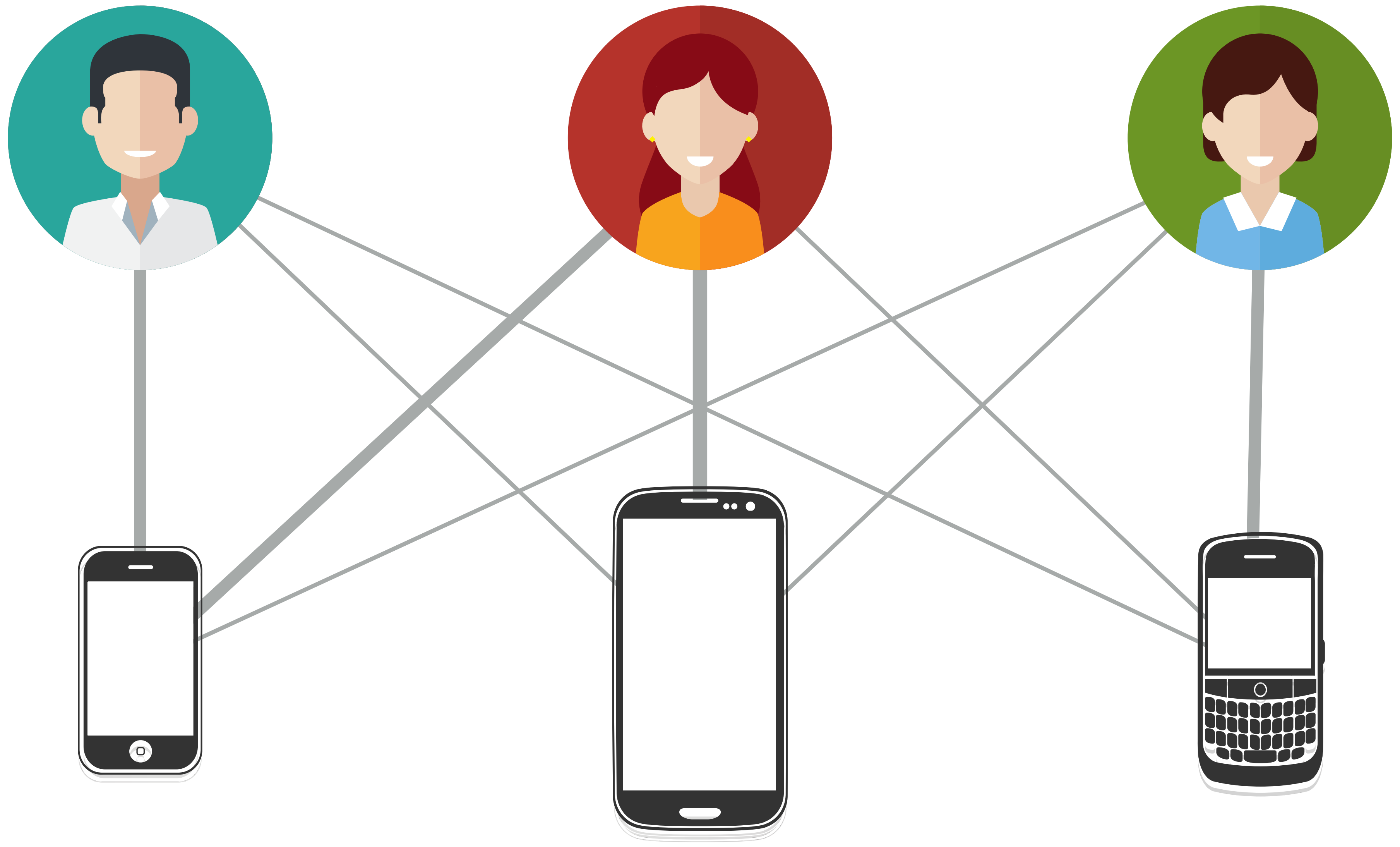
$$\text{argmax}_{\mathcal{V} \subseteq \mathcal{S}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} r(u, v) c_{uv}$$

such that:

$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Re-Formalizing with a Latent Space

$$\mathbf{z}_{u_1} \quad \mathbf{z}_{u_2} \quad \mathbf{z}_{u_3} \quad \cdots \quad \in \mathcal{Z} \subseteq \mathbb{R}^{d_z}$$



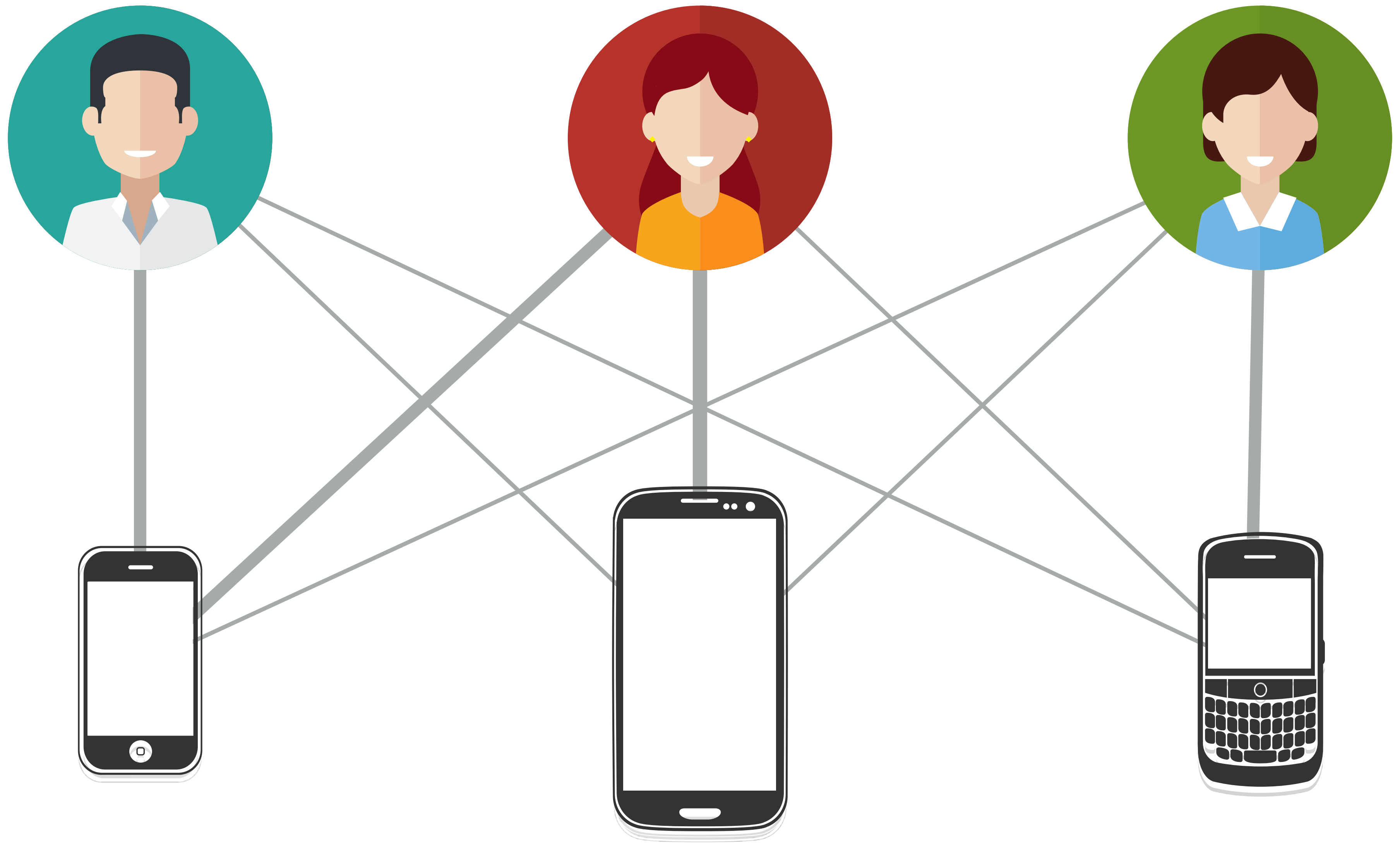
$$\boxed{\operatorname{argmax}_{\mathcal{V} \subseteq \mathcal{S}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} r(u, v) c_{uv}}$$

such that:

$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Re-Formalizing with a Latent Space

$$\mathbf{z}_{u_1} \quad \mathbf{z}_{u_2} \quad \mathbf{z}_{u_3} \quad \cdots \quad \in \mathcal{Z} \subseteq \mathbb{R}^{d_z}$$



$$\mathbf{z}_{s_1} \quad \mathbf{z}_{s_2} \quad \mathbf{z}_{s_3} \quad \cdots \quad \in \mathcal{Z} \subseteq \mathbb{R}^{d_z}$$

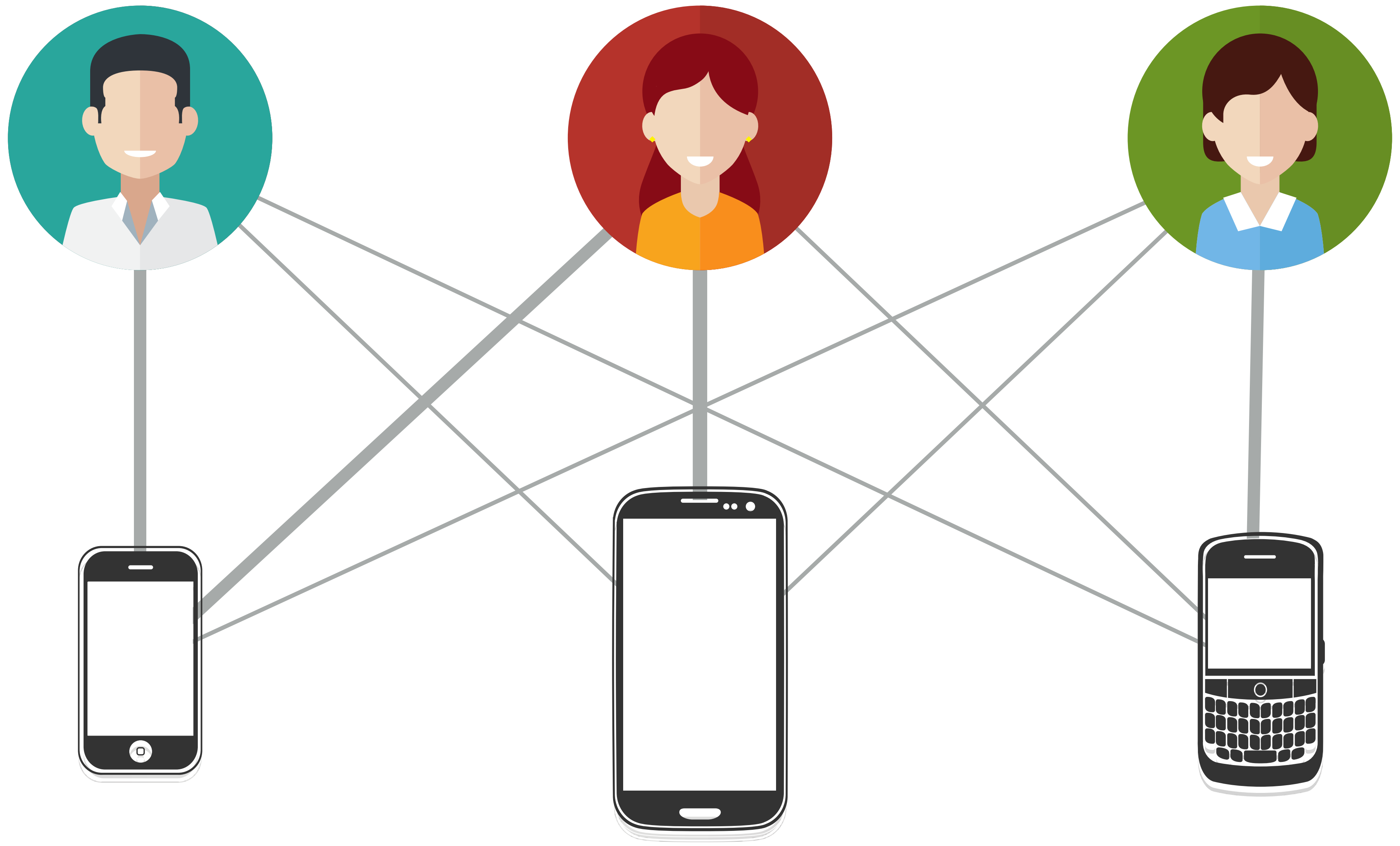
$$\boxed{\operatorname{argmax}_{\mathcal{V} \subseteq \mathcal{S}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} r(u, v) c_{uv}}$$

such that:

$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Re-Formalizing with a Latent Space

$$\mathbf{z}_{u_1} \quad \mathbf{z}_{u_2} \quad \mathbf{z}_{u_3} \quad \cdots \quad \in \mathcal{Z} \subseteq \mathbb{R}^{d_z}$$



$$\mathbf{z}_{s_1} \quad \mathbf{z}_{s_2} \quad \mathbf{z}_{s_3} \quad \cdots \quad \in \mathcal{Z} \subseteq \mathbb{R}^{d_z}$$

Choose a convenient parametric function:

$$\hat{r}_{\text{dot}}(\mathbf{z}_u, \mathbf{z}_s) = \mathbf{z}_u^\top \mathbf{z}_s$$

$$\hat{r}_{\text{euc}}(\mathbf{z}_u, \mathbf{z}_s) = \alpha - \|\mathbf{z}_u - \mathbf{z}_s\|_2$$

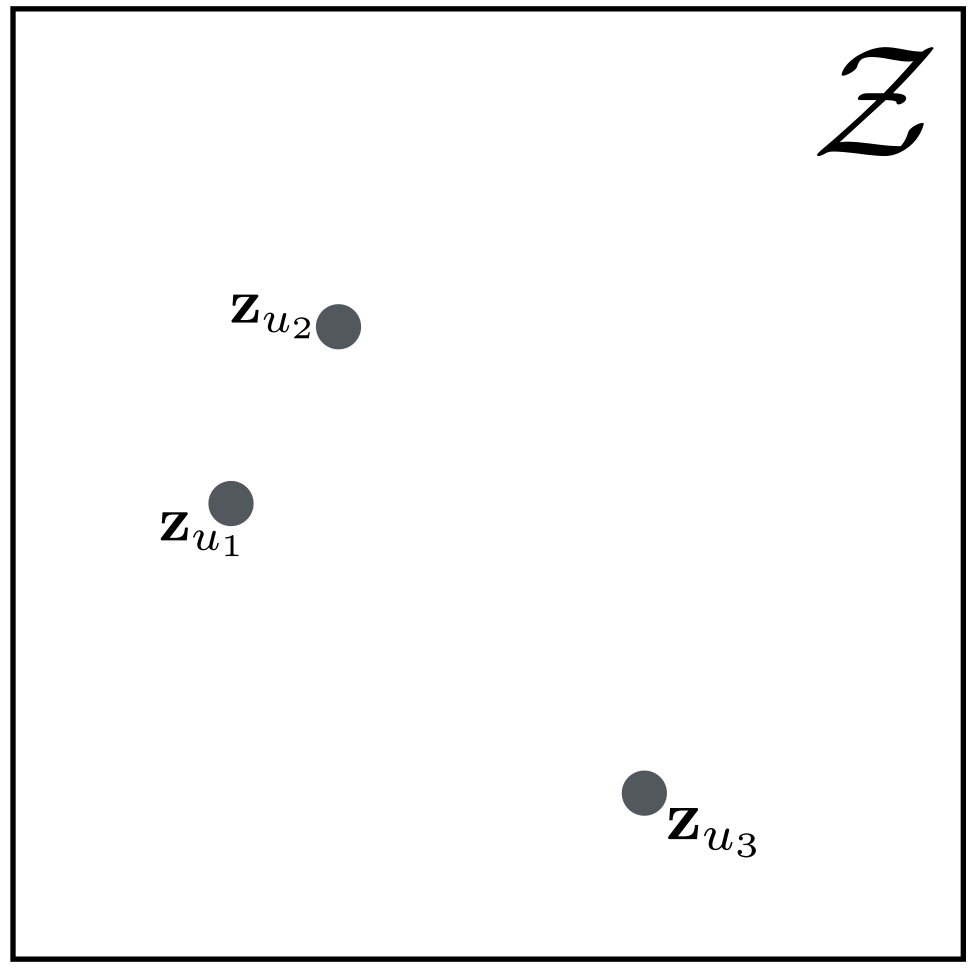
$$\boxed{\operatorname{argmax}_{\mathcal{V} \subseteq \mathcal{S}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} r(u, v) c_{uv}}$$

such that:

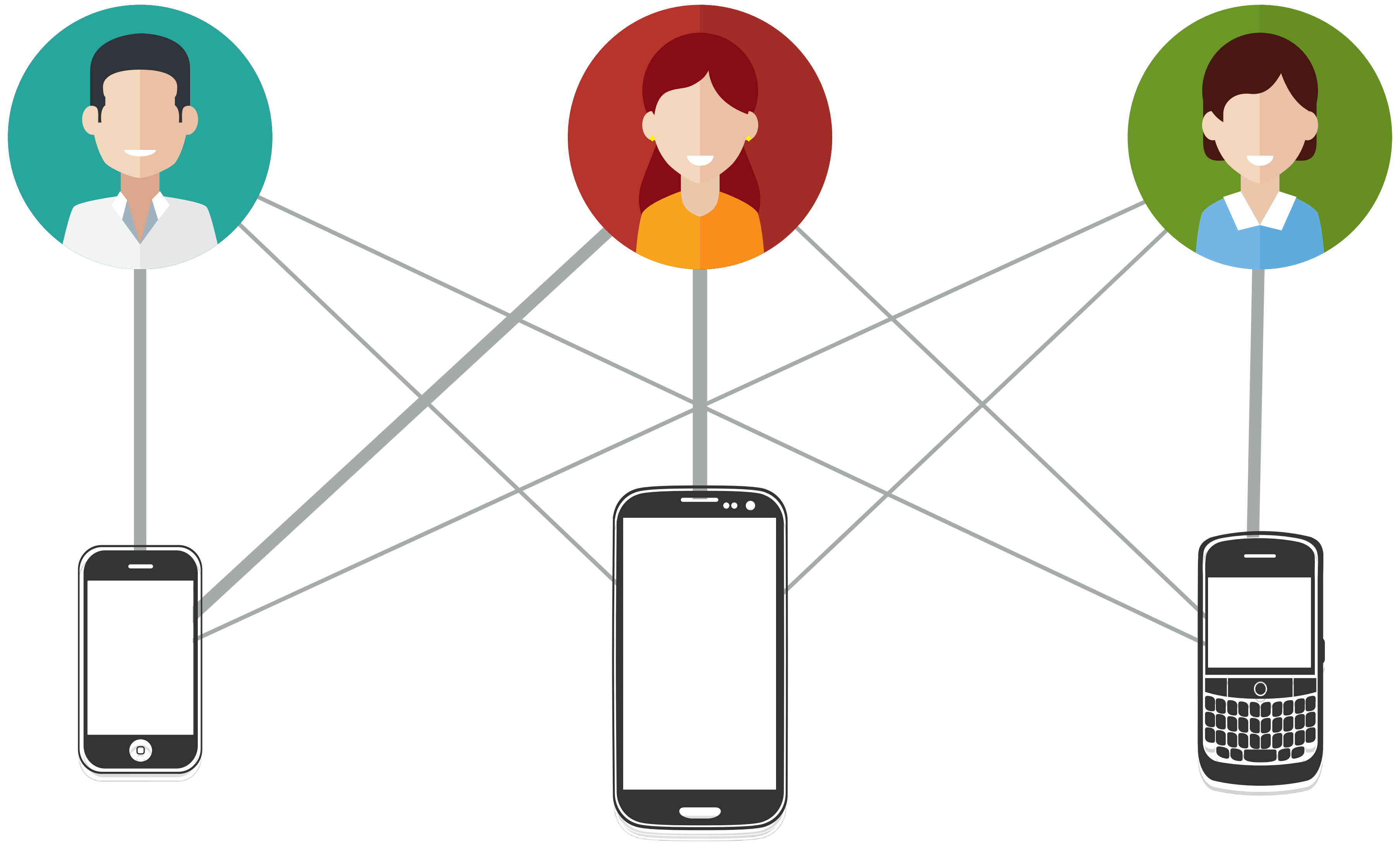
$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Re-Formalizing with a Latent Space

(Real) Latent Space



$\mathbf{z}_{u_1}$ $\quad$ $\mathbf{z}_{u_2}$ $\quad$ $\mathbf{z}_{u_3}$ $\quad \cdots \quad \in \mathcal{Z} \subseteq \mathbb{R}^{d_z}$



$\mathbf{z}_{s_1}$ $\quad$ $\mathbf{z}_{s_2}$ $\quad$ $\mathbf{z}_{s_3}$ $\quad \cdots \quad \in \mathcal{Z} \subseteq \mathbb{R}^{d_z}$

Choose a convenient parametric function:

$$\hat{r}_{\text{dot}}(\mathbf{z}_u, \mathbf{z}_s) = \mathbf{z}_u^\top \mathbf{z}_s$$

$$\hat{r}_{\text{euc}}(\mathbf{z}_u, \mathbf{z}_s) = \alpha - \|\mathbf{z}_u - \mathbf{z}_s\|_2$$

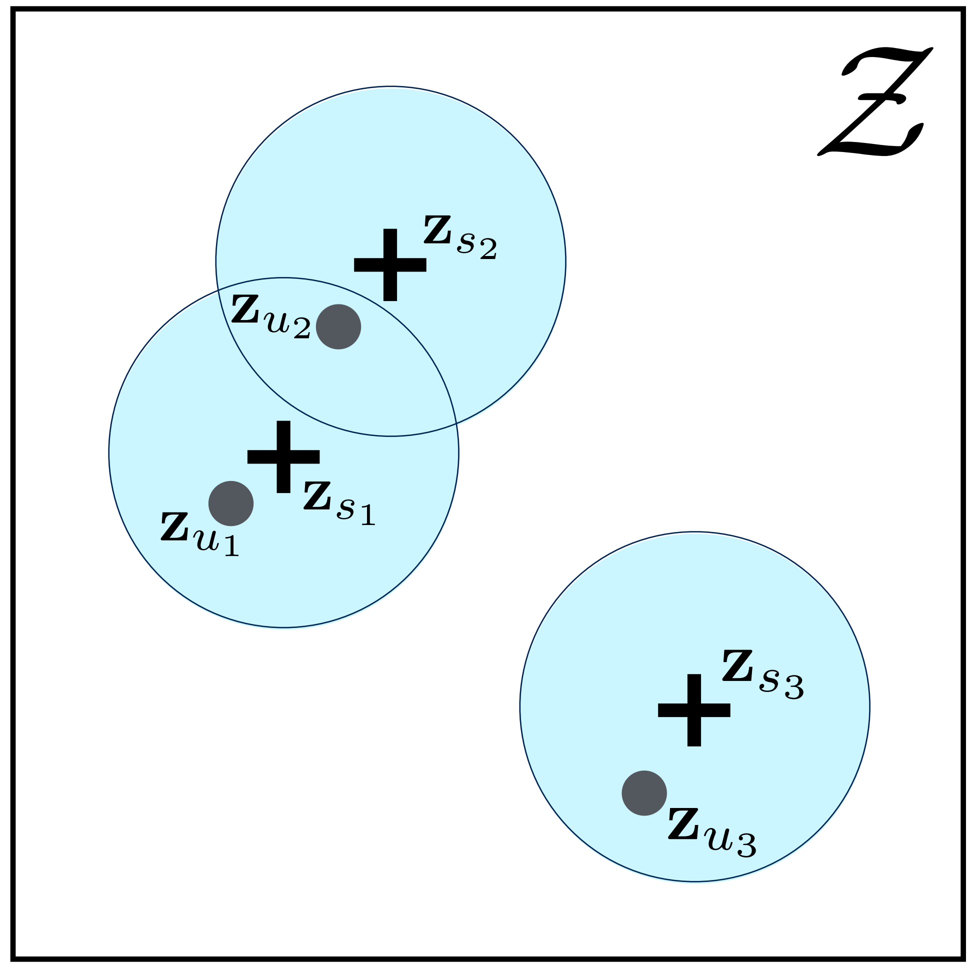
$$\boxed{\operatorname{argmax}_{\mathcal{V} \subseteq \mathcal{S}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} r(u, v) c_{uv}}$$

such that:

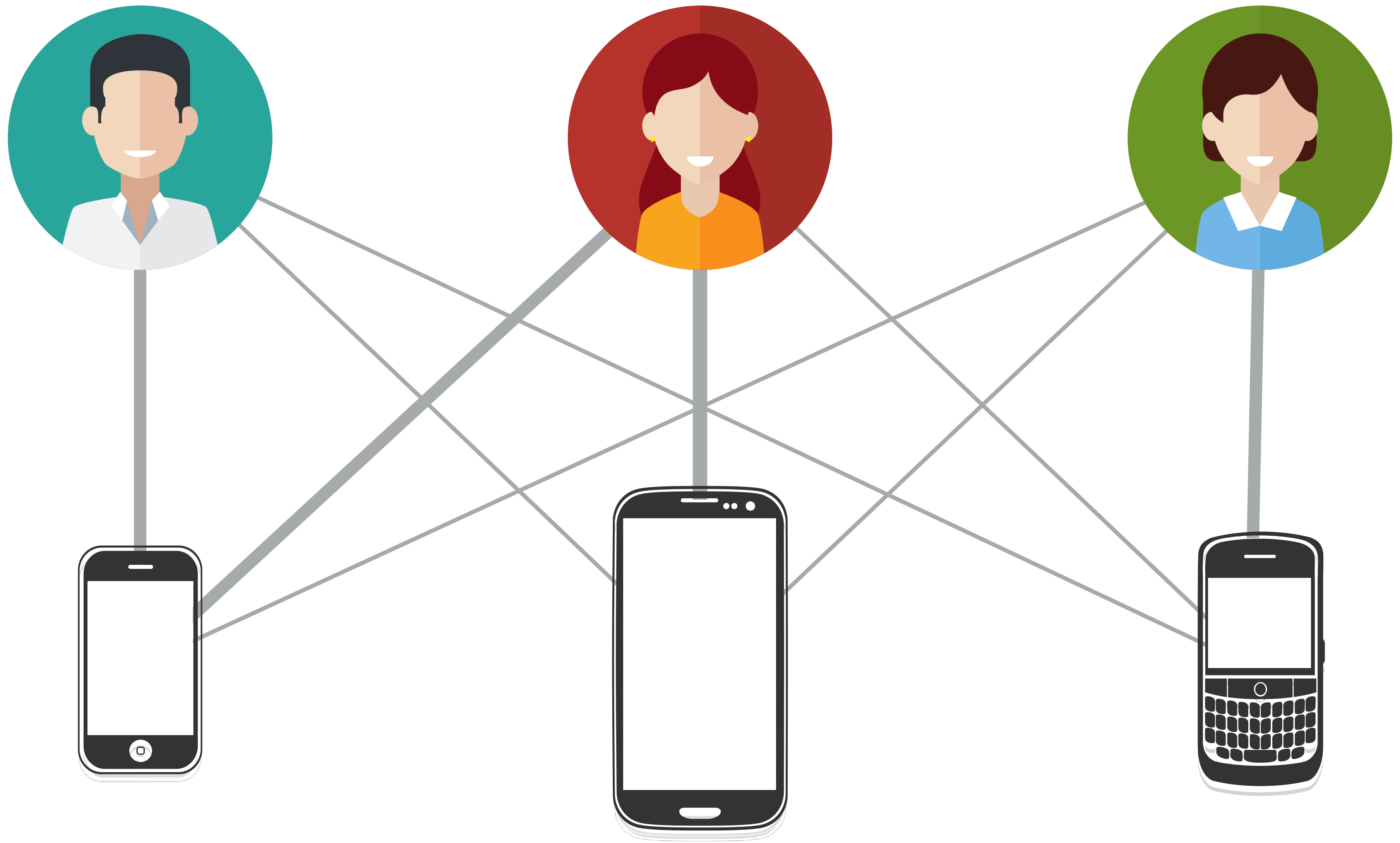
$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Re-Formalizing with a Latent Space

(Real) Latent Space



$$\mathbf{z}_{u_1} \quad \mathbf{z}_{u_2} \quad \mathbf{z}_{u_3} \quad \cdots \quad \in \mathcal{Z} \subseteq \mathbb{R}^{d_z}$$



$$\mathbf{z}_{s_1} \quad \mathbf{z}_{s_2} \quad \mathbf{z}_{s_3} \quad \cdots \quad \in \mathcal{Z} \subseteq \mathbb{R}^{d_z}$$

Choose a convenient parametric function:

$$\hat{r}_{\text{dot}}(\mathbf{z}_u, \mathbf{z}_s) = \mathbf{z}_u^\top \mathbf{z}_s$$

$$\hat{r}_{\text{euc}}(\mathbf{z}_u, \mathbf{z}_s) = \alpha - \|\mathbf{z}_u - \mathbf{z}_s\|_2$$

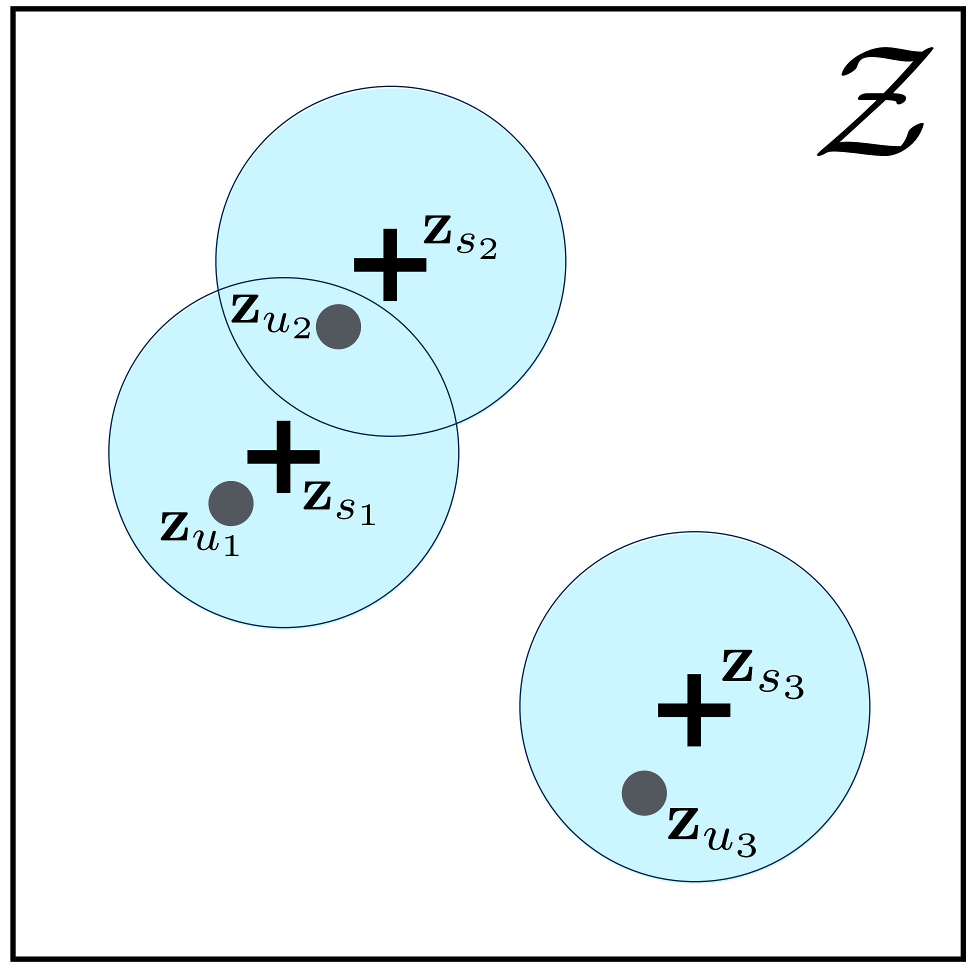
$$\boxed{\operatorname{argmax}_{\mathcal{V}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} \hat{r}(\mathbf{z}_u, \mathbf{z}_v) c_{uv}}$$

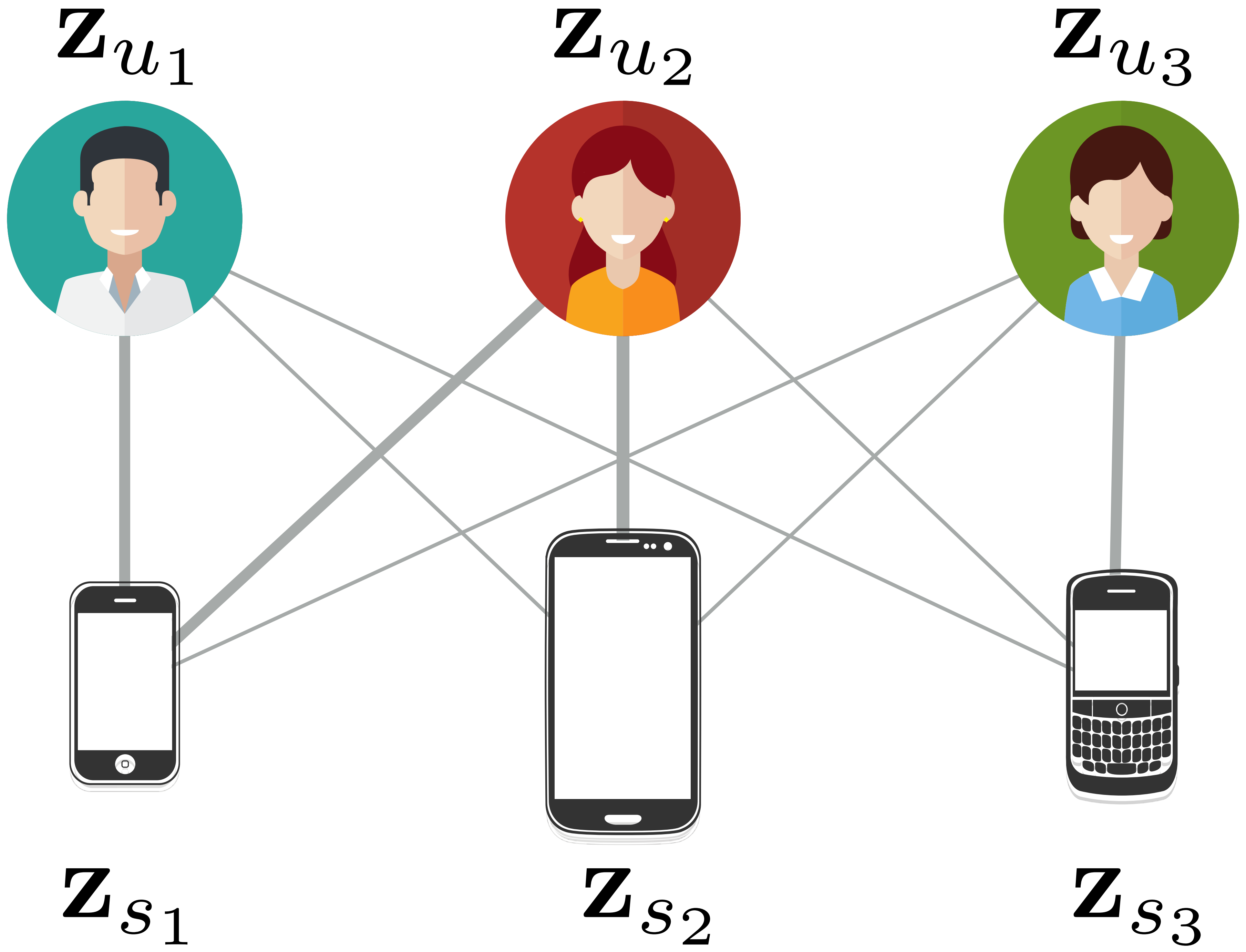
such that:

$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Re-Formalizing with a Latent Space

(Real) Latent Space



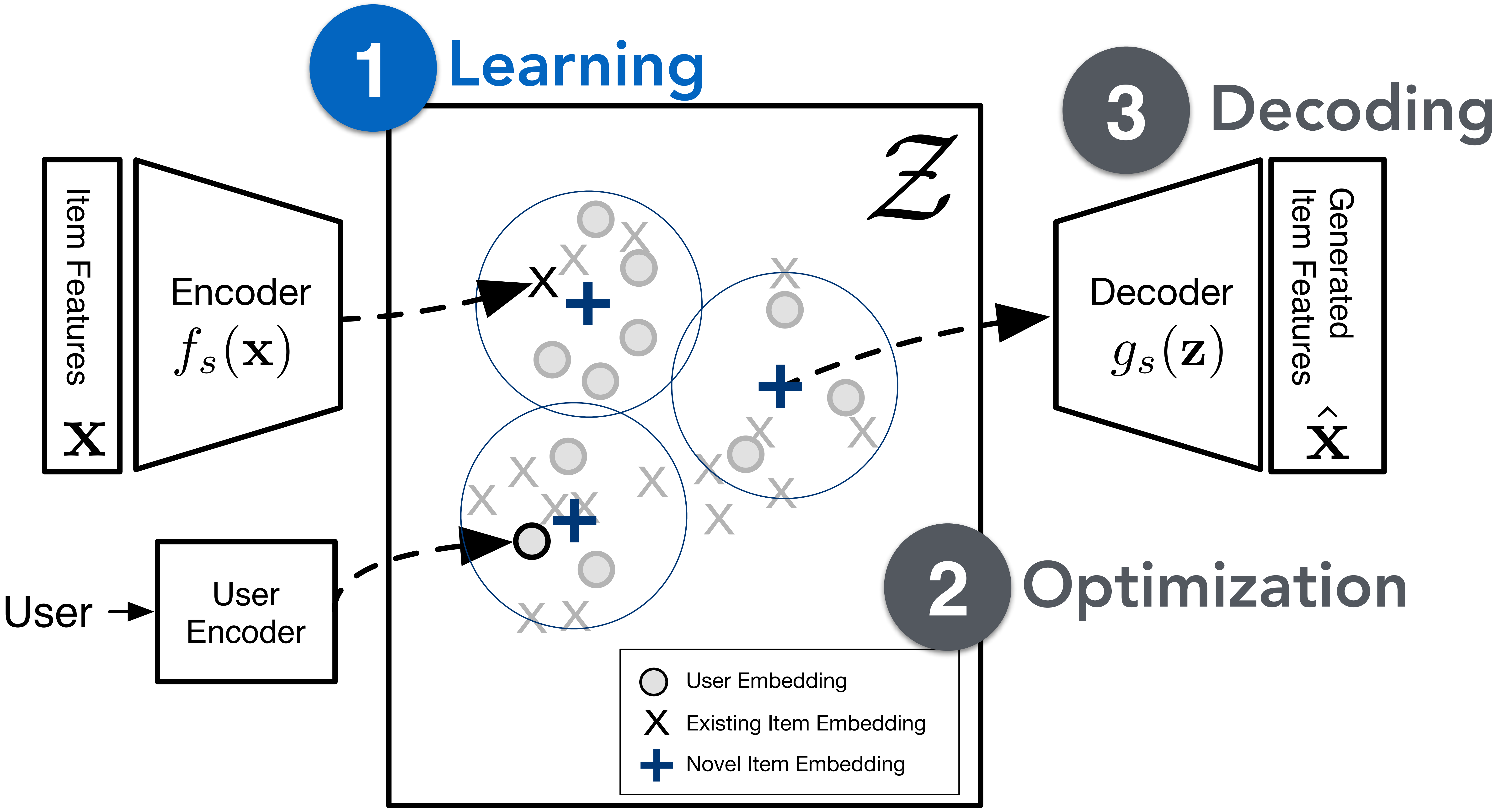


Provided we can learn Z!

$$\boxed{\operatorname{argmax}_{\mathcal{V}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} \hat{r}(\mathbf{z}_u, \mathbf{z}_v) c_{uv}}$$

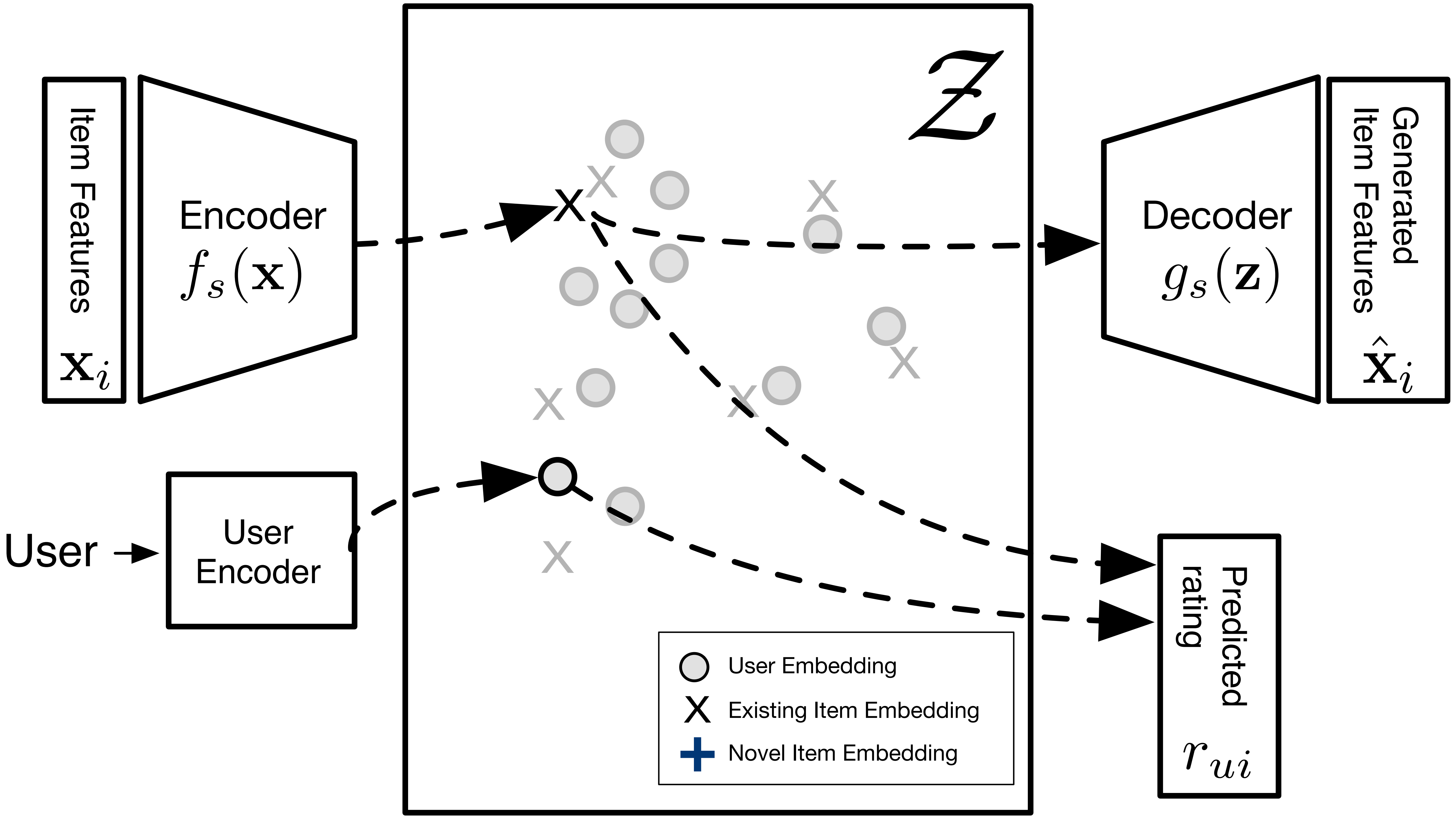
such that:

$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$



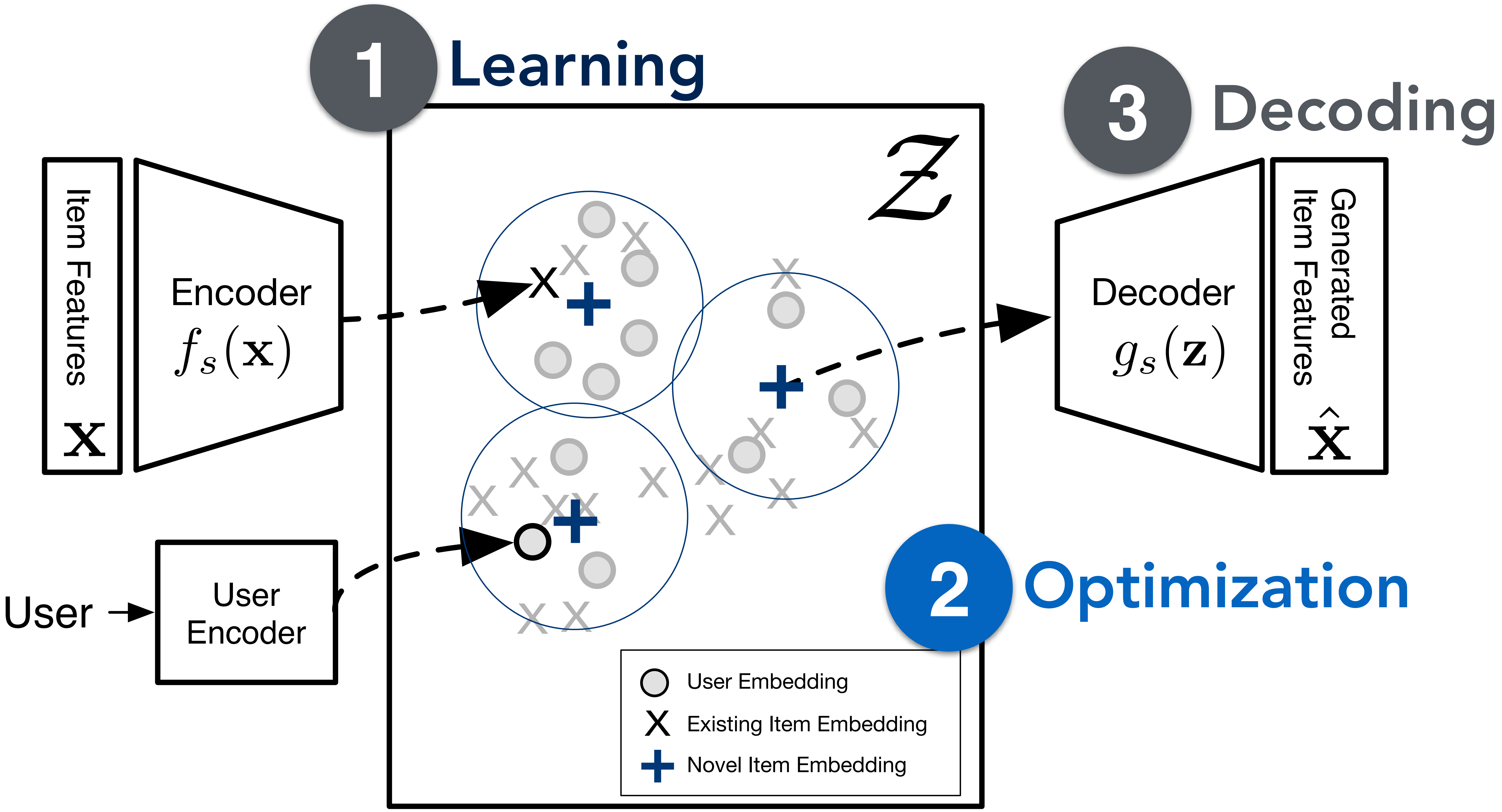
$$\begin{aligned}\mathcal{L} &= \text{Rating Loss} + \text{Reconstruction Loss} + \text{Regularizer} \\ &= \mathcal{L}_r(\theta) + \lambda_1 \mathcal{L}_g(\phi) + \lambda_2 \mathcal{R}(\theta, \phi)\end{aligned}$$

Collaborative VAE



$$\ell(\theta, \phi; d) = \underbrace{\mathbb{E}_{q(\mathbf{z}_i | f_\phi(\mathbf{x}_i))}[\log p(r_{ui}|\mathbf{z}_u, \mathbf{z}_i)]}_{\text{Rating prediction}} + \underbrace{\mathbb{E}_{q(\mathbf{z}_i | f_\phi(\mathbf{x}_i))}[\log p(\mathbf{x}_i | g_\theta(\mathbf{z}_i))]}_{\text{Item reconstruction}} - \underbrace{\mathbb{D}_{\mathrm{KL}}[q(\mathbf{z}_i | f_\phi(\mathbf{x}_i)) || p(\mathbf{z}_i)]}_{\text{Prior Penalty}}$$

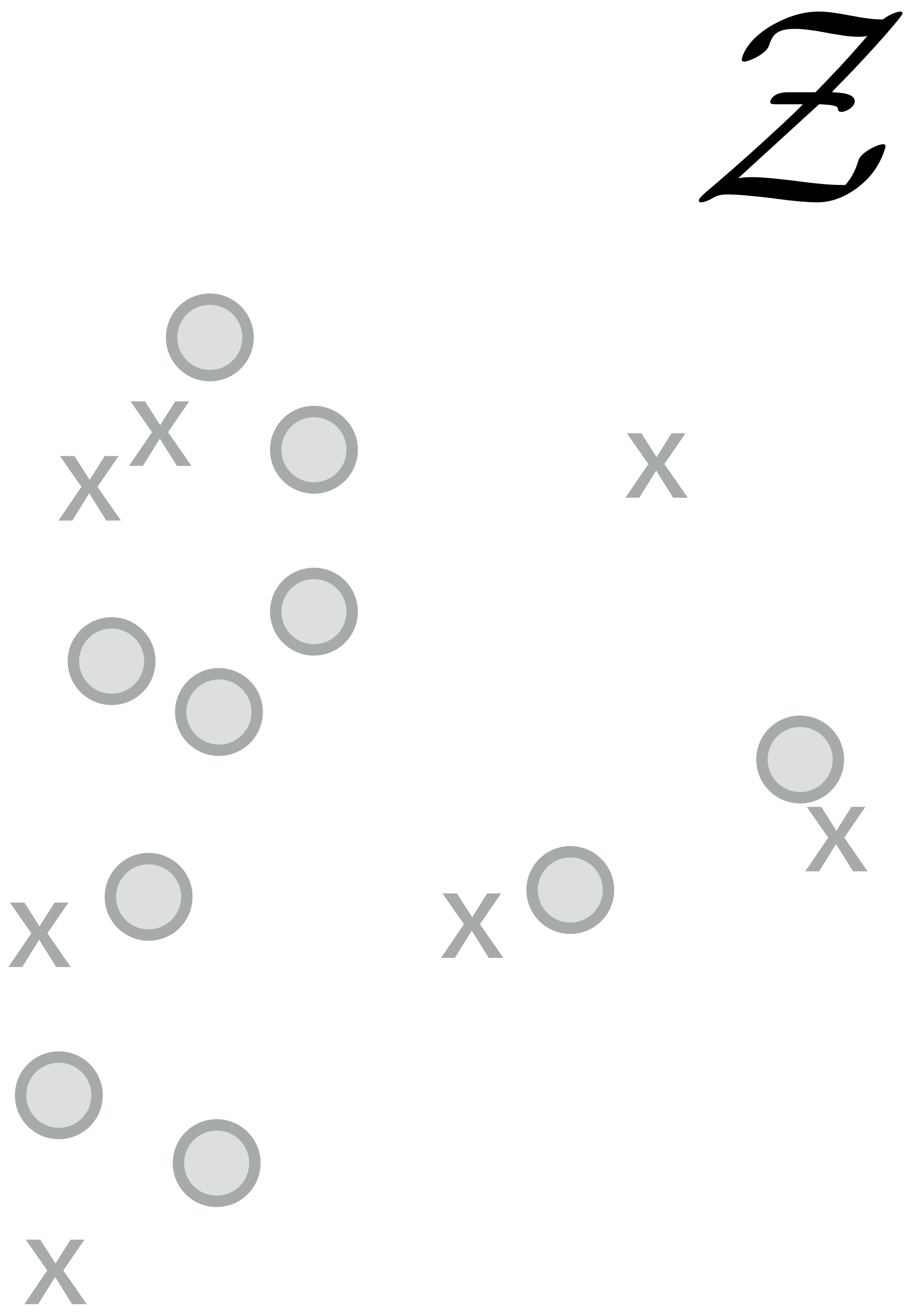
[Li, She, KDD 2017; Liang et al, WWW 2018; Vinh, Soh, 2018]


1 Learning
3 Decoding
2 Optimization
Item Features
$\mathbf{x}$
Encoder
$f_s(\mathbf{x})$
$\mathcal{Z}$
Decoder
$g_s(\mathbf{z})$
Generated Item Features
$\hat{\mathbf{x}}$
User
User Encoder
User Embedding
Existing Item Embedding
Novel Item Embedding

Greedy Maximum Coverage

User

X Existing Item

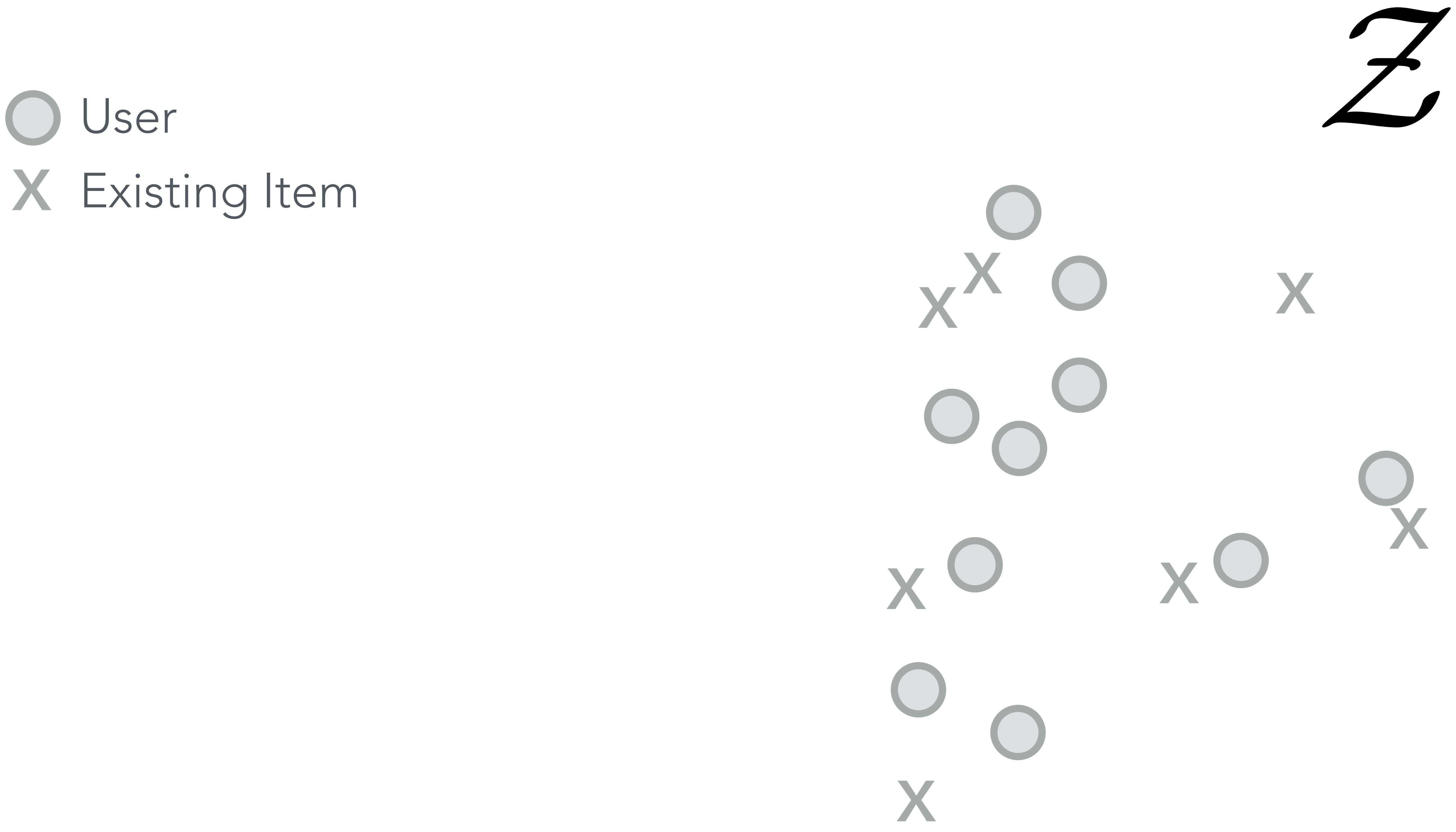


$$\text{argmax}_{\mathcal{V}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} \hat{r}(\mathbf{z}_u, \mathbf{z}_v) c_{uv}$$

such that:

$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Greedy Maximum Coverage



$$\boxed{\operatorname{argmax}_{\mathcal{V}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} \hat{r}(\mathbf{z}_u, \mathbf{z}_v) c_{uv}}$$

such that:

$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Greedy Maximum Coverage

$$\operatorname{argmax}_{\mathcal{V}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} \hat{r}(\mathbf{z}_u, \mathbf{z}_v) c_{uv}$$

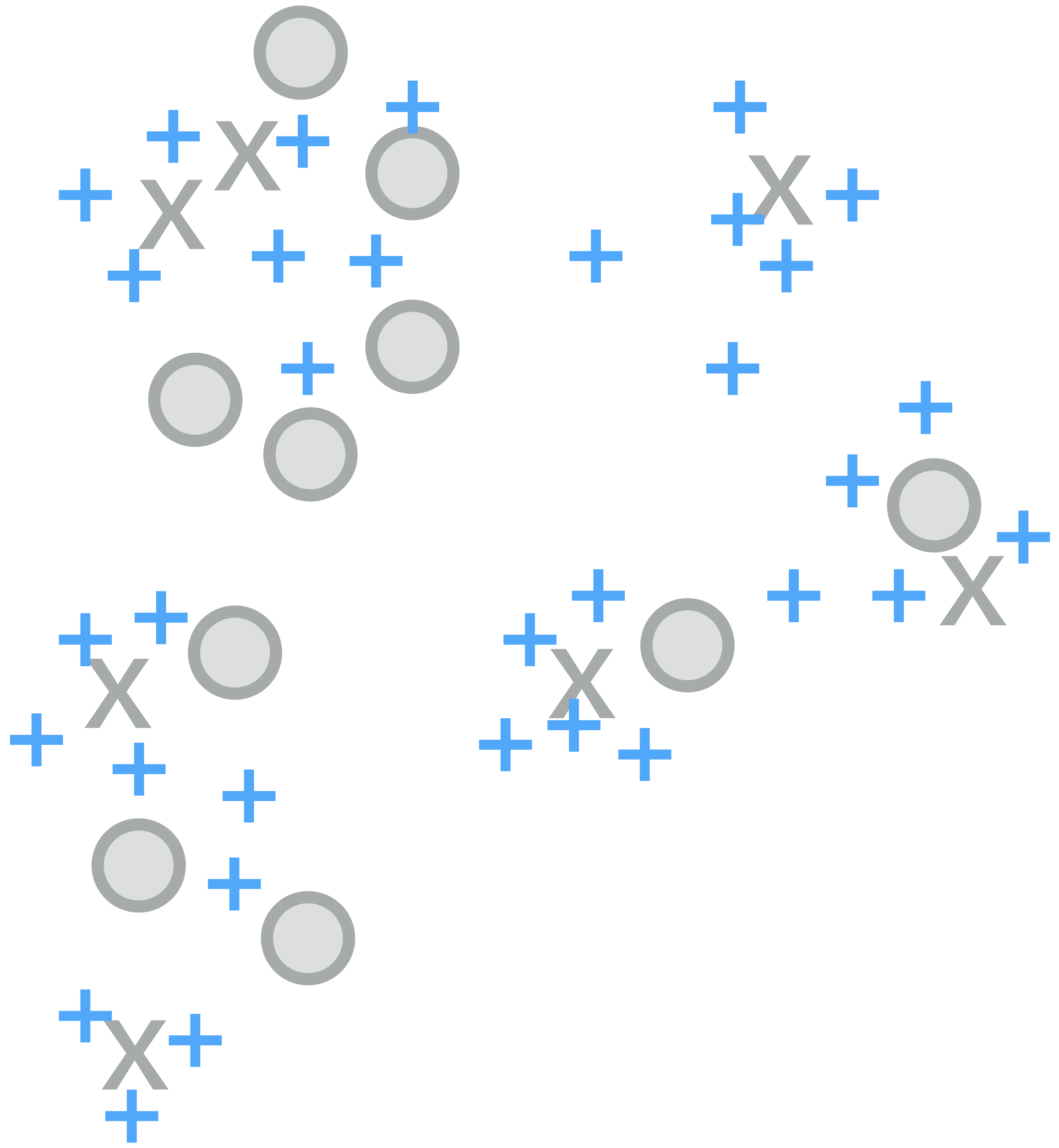
such that:

$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Greedy Maximum Coverage

- ◯ User
- X Existing Item
- + Candidate Novel Item

$\mathcal{Z}$



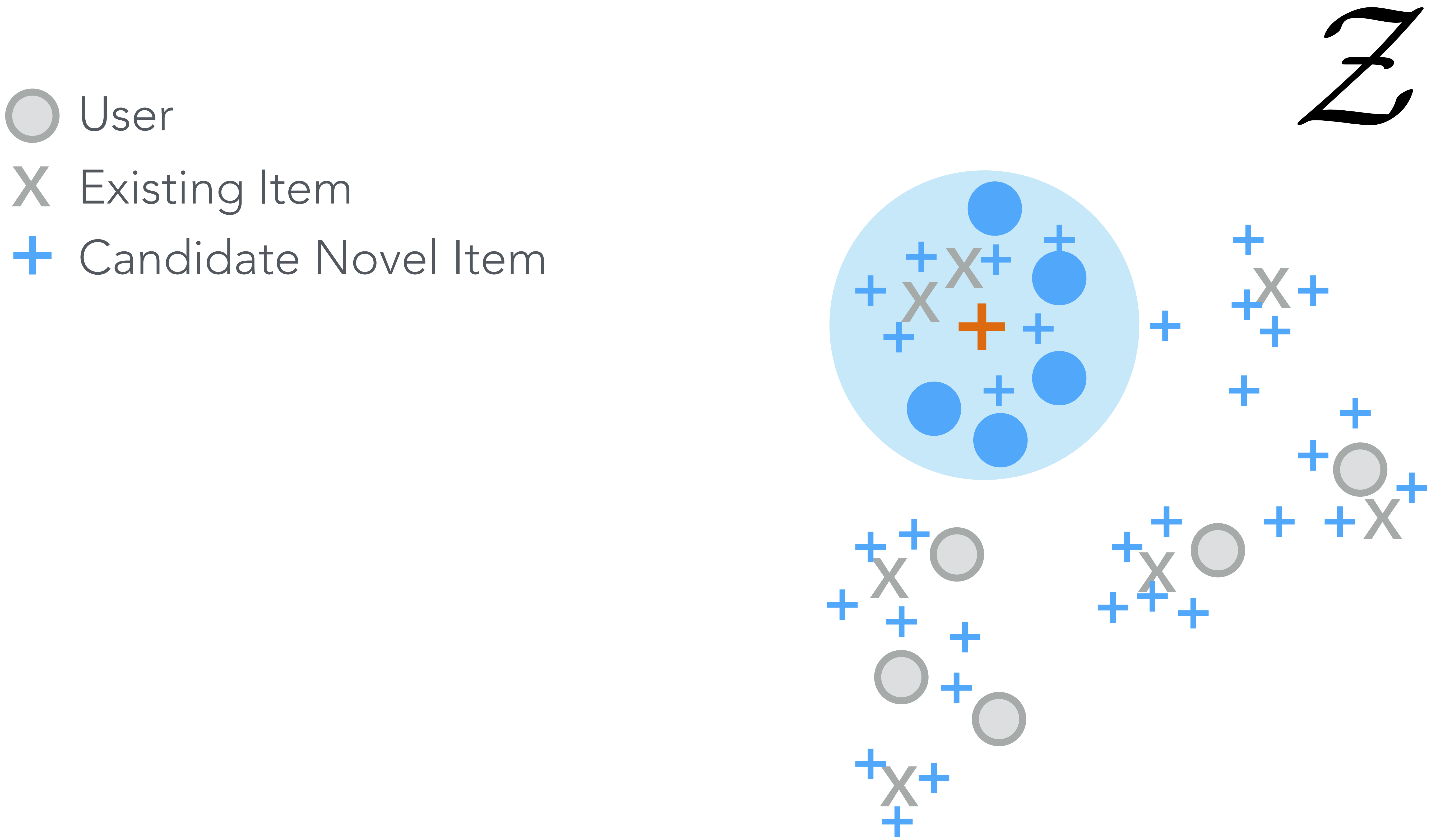
- Submodular [Nemhauser, Wolsey, Fisher, 1978]
- Achieves $1 - \frac{1}{e}$ approximation ratio
- best polynomial time algorithm unless P = NP

$$\boxed{\operatorname{argmax}_{\mathcal{V}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} \hat{r}(\mathbf{z}_u, \mathbf{z}_v) c_{uv}}$$

such that:

$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Greedy Maximum Coverage



- Submodular [Nemhauser, Wolsey, Fisher, 1978]
- Achieves $1 - \frac{1}{e}$ approximation ratio
- best polynomial time algorithm unless P = NP

$$\boxed{\operatorname{argmax}_{\mathcal{V}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} \hat{r}(\mathbf{z}_u, \mathbf{z}_v) c_{uv}}$$

such that:

$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Greedy Maximum Coverage

- Submodular [Nemhauser, Wolsey, Fisher, 1978]
- Achieves $1 - \frac{1}{e}$ approximation ratio
- best polynomial time algorithm unless P = NP

$$\boxed{\text{argmax}_{\mathcal{V}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} \hat{r}(\mathbf{z}_u, \mathbf{z}_v) c_{uv}}$$

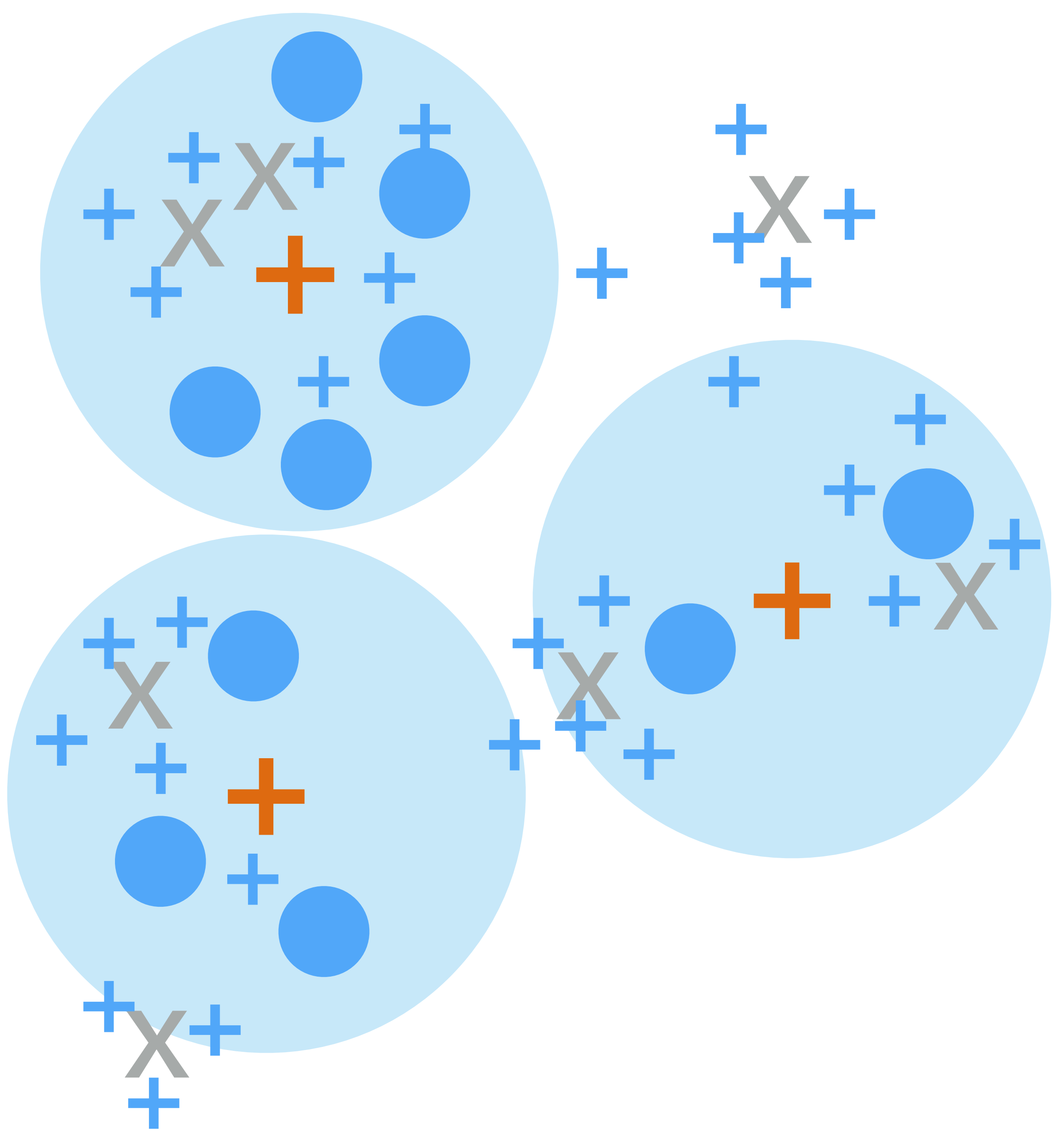
such that:

$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Greedy Maximum Coverage

- User
- X Existing Item
- + Candidate Novel Item

$\mathcal{Z}$

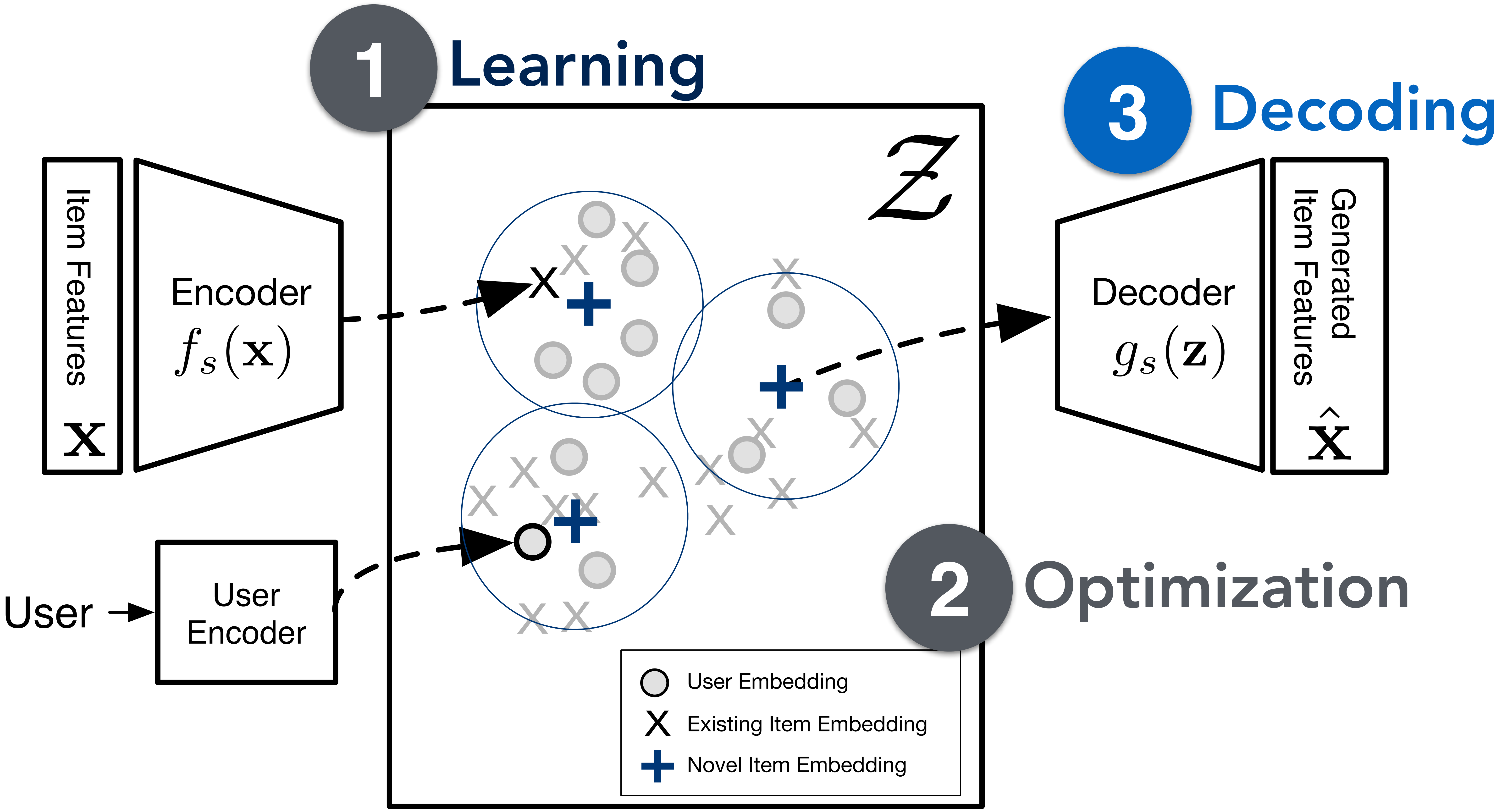


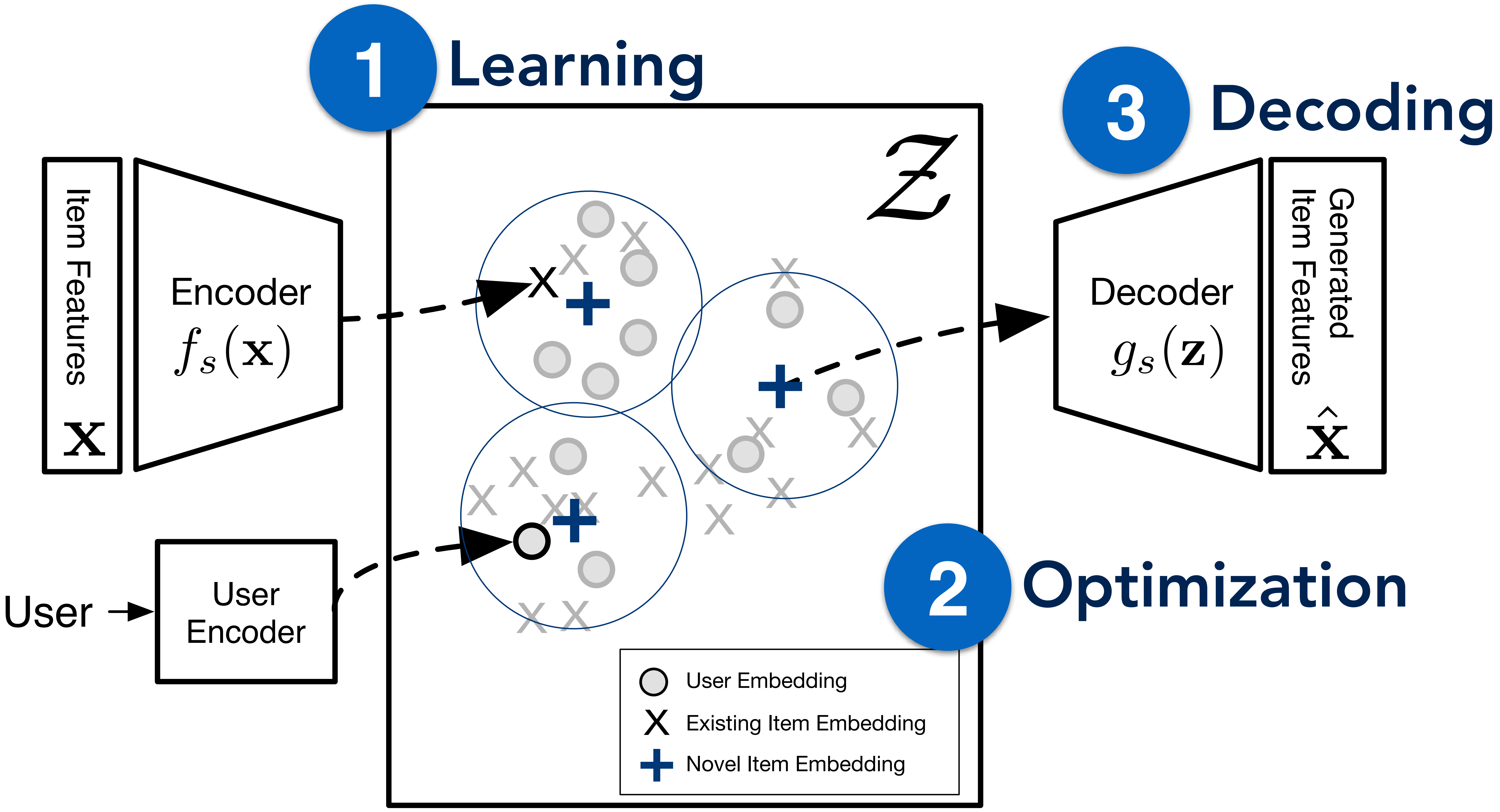
- Submodular [Nemhauser, Wolsey, Fisher, 1978]
- Achieves $1 - \frac{1}{e}$ approximation ratio
- best polynomial time algorithm unless P = NP

$$\boxed{\operatorname{argmax}_{\mathcal{V}} \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} \hat{r}(\mathbf{z}_u, \mathbf{z}_v) c_{uv}}$$

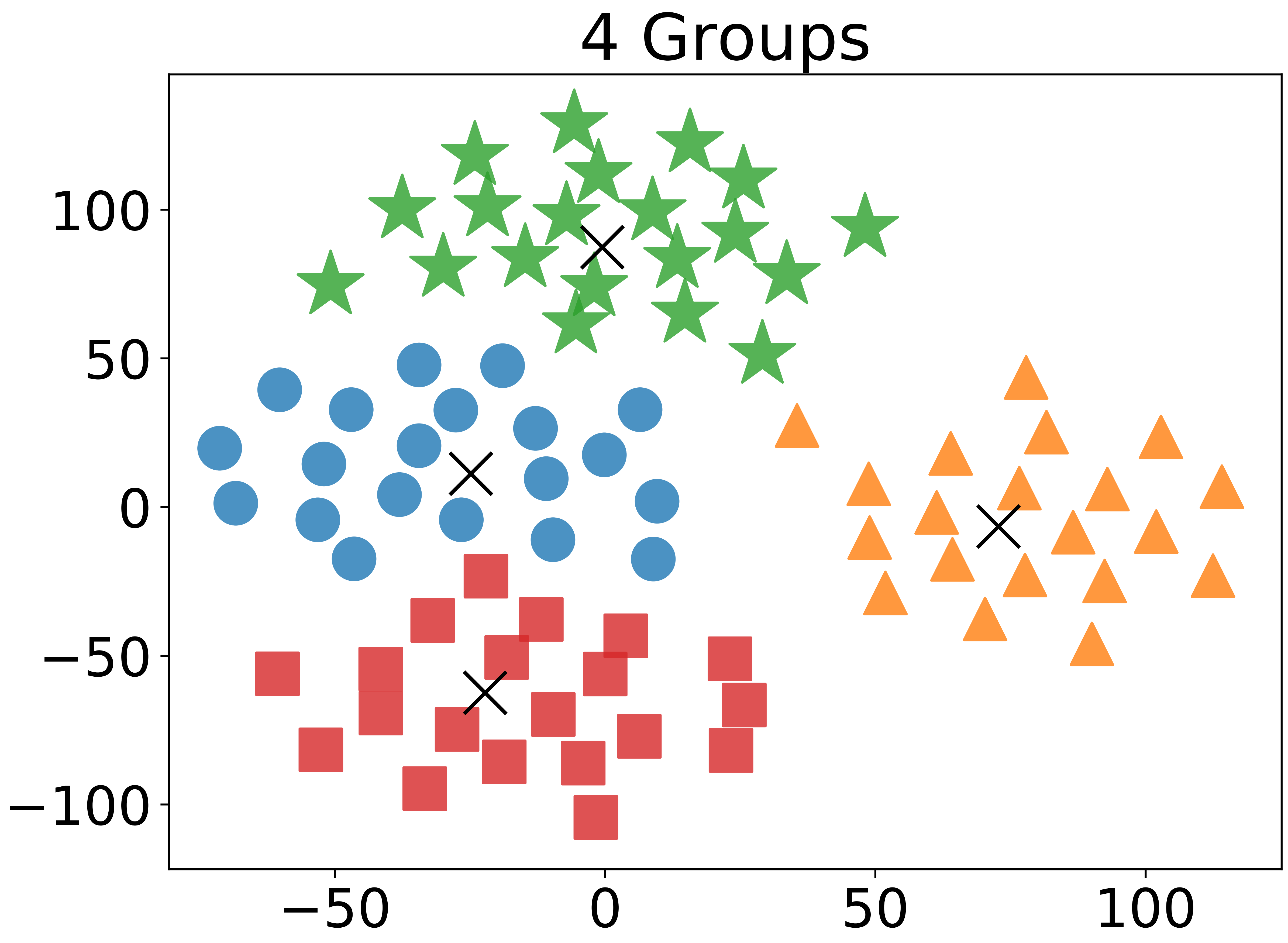
such that:

$$\sum_{v \in \mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$


1 Learning
3 Decoding
2 Optimization
Item Features $\mathbf{x}$
Encoder $f_s(\mathbf{x})$
$\mathcal{Z}$
Decoder $g_s(\mathbf{z})$
Generated Item Features $\hat{\mathbf{x}}$
User
User Encoder
User Embedding
Existing Item Embedding
Novel Item Embedding

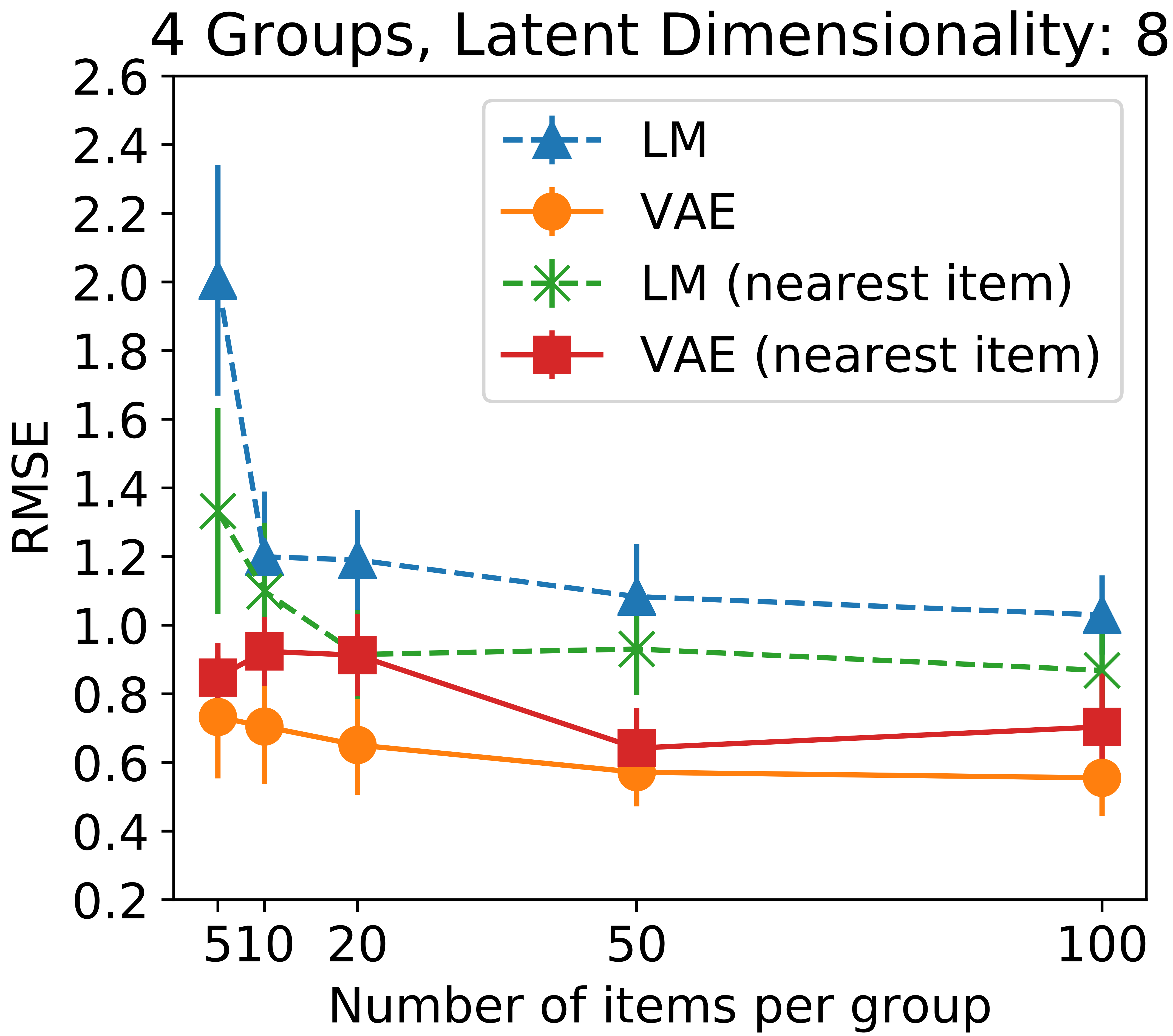

1 Learning
2 Optimization
3 Decoding
Item Features
$\mathbf{x}$
Encoder
$f_s(\mathbf{x})$
User
User Encoder
$\mathcal{Z}$
User Embedding
Existing Item Embedding
Novel Item Embedding
Decoder
$g_s(\mathbf{z})$
Generated Item Features
$\hat{\mathbf{x}}$

Quantitative Results: Synthetic Data

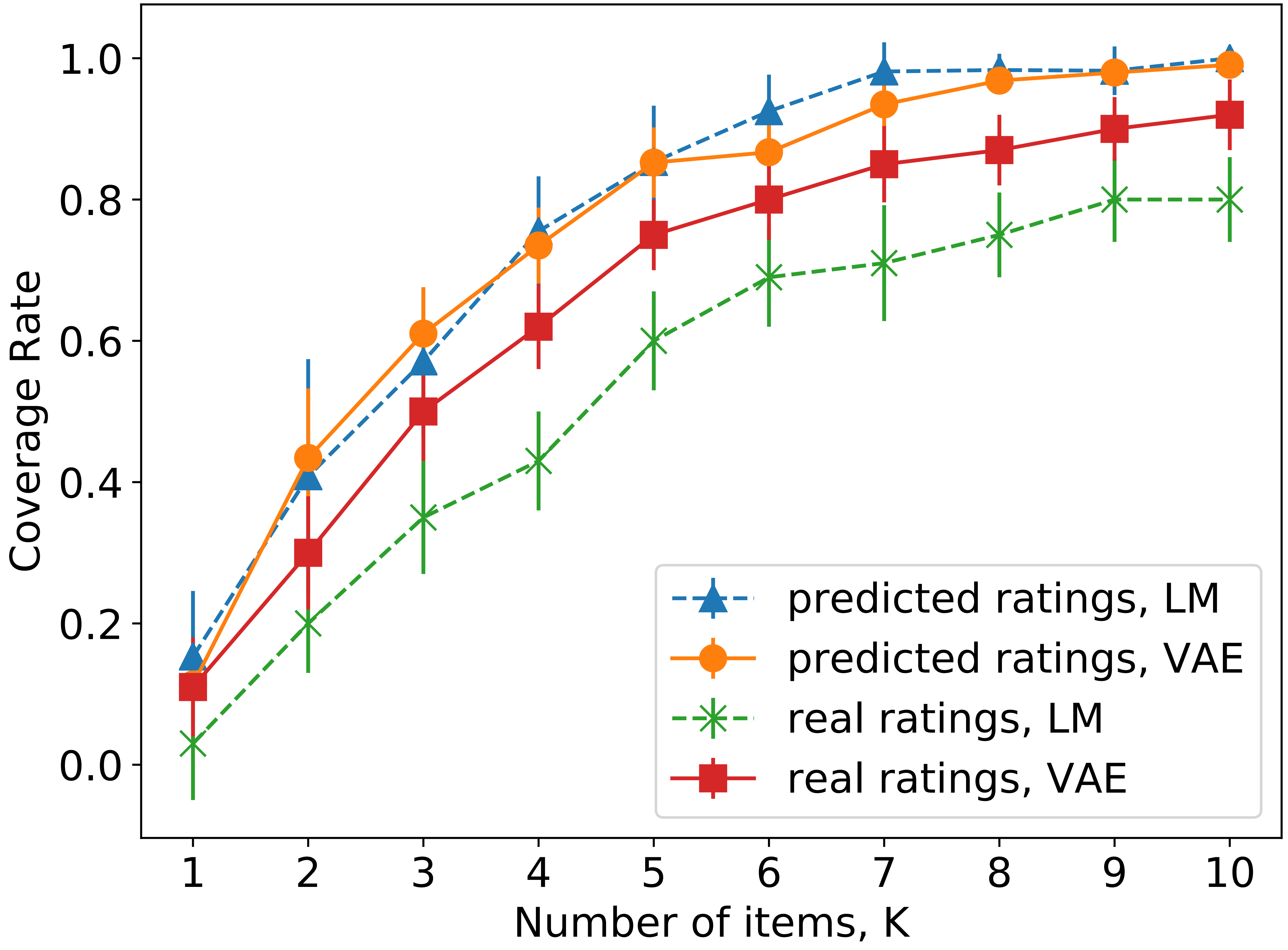


- 20 dimensional items
- Visualized in 2D (via t-SNE)

Quantitative Results: Synthetic Data



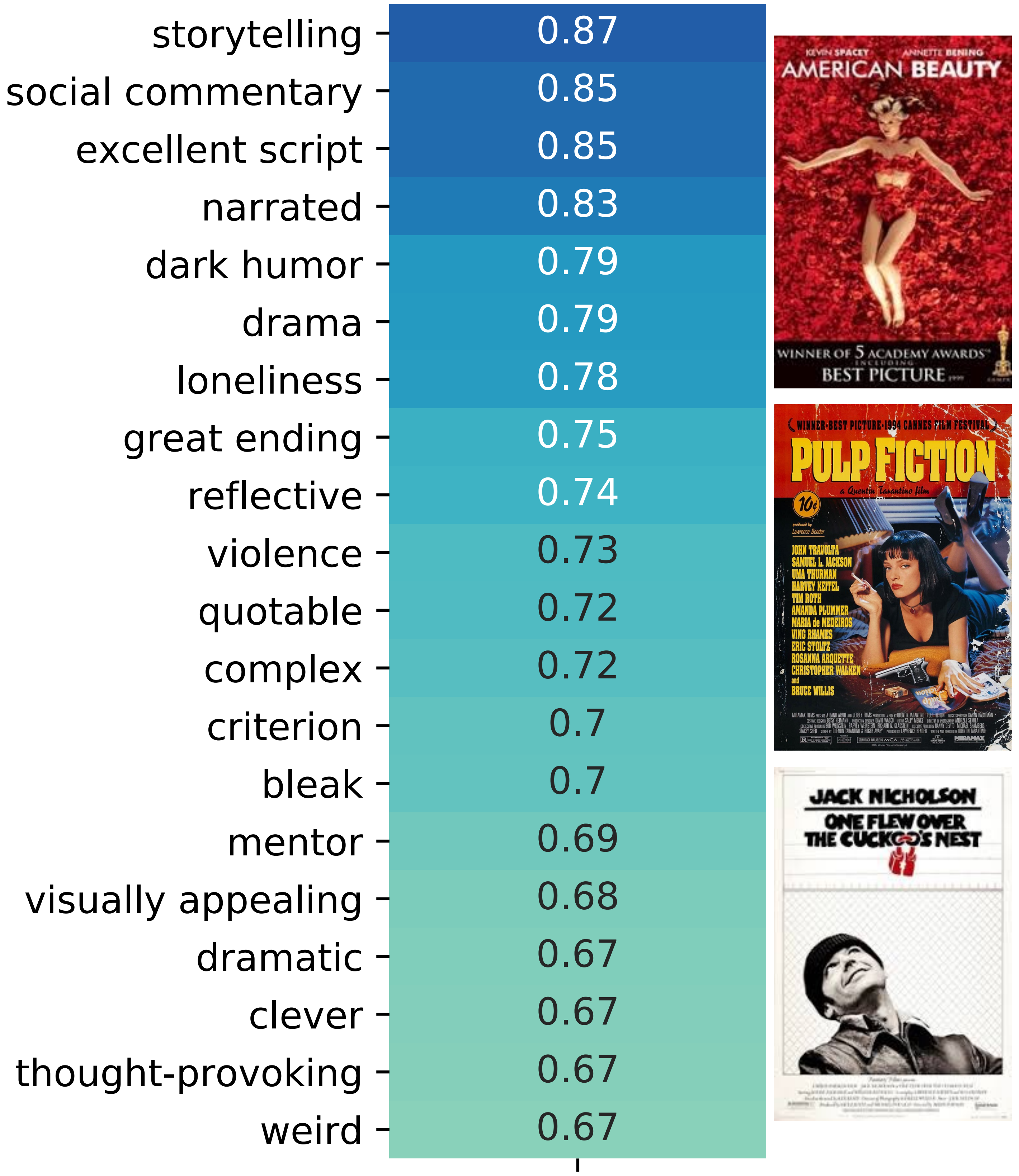
Quantitative Results: Synthetic Data



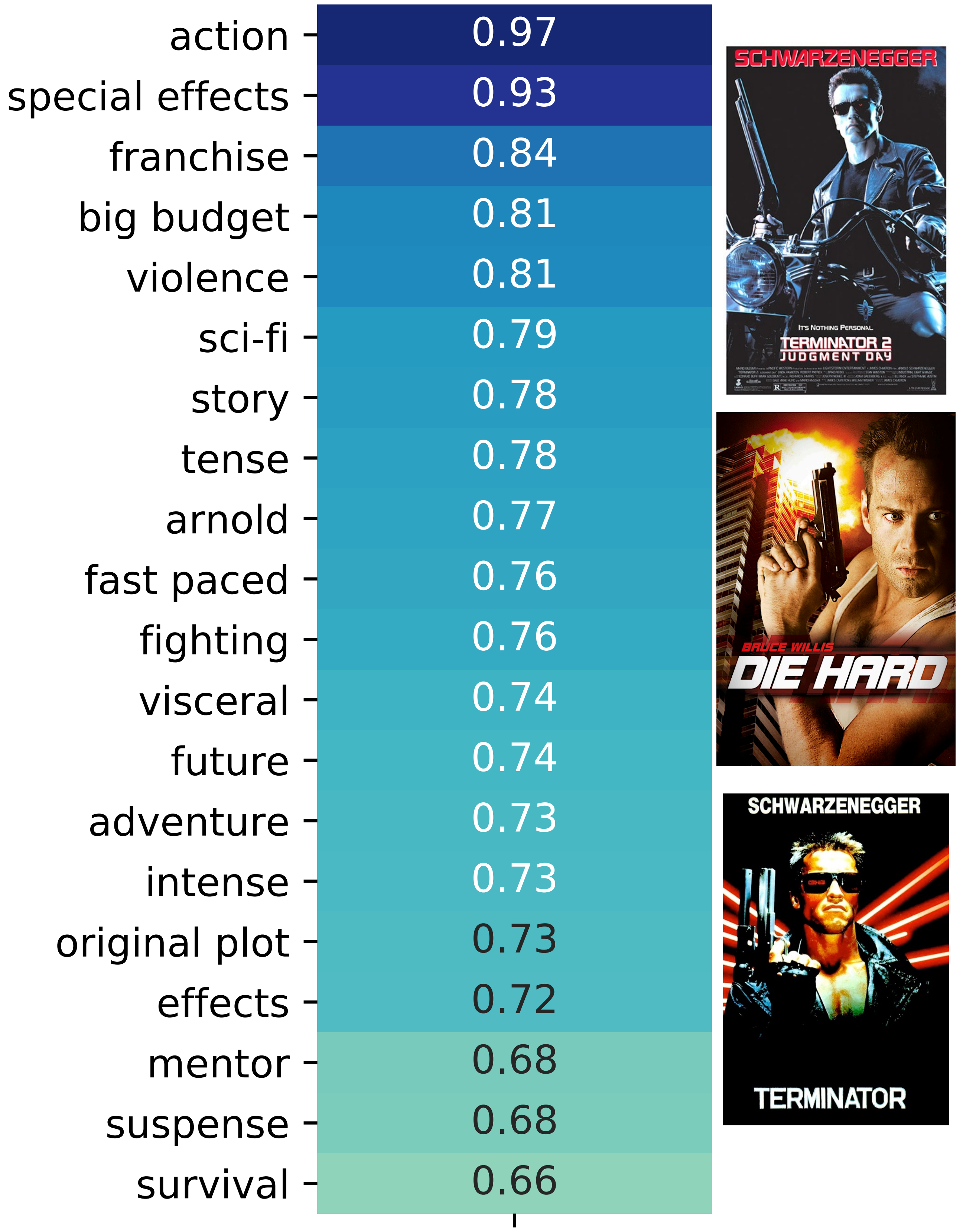
Generated Movie “Genomes”

Generated Movie 1
PredR(all): 0.87±0.07, PredR(> τ): 0.89±0.05

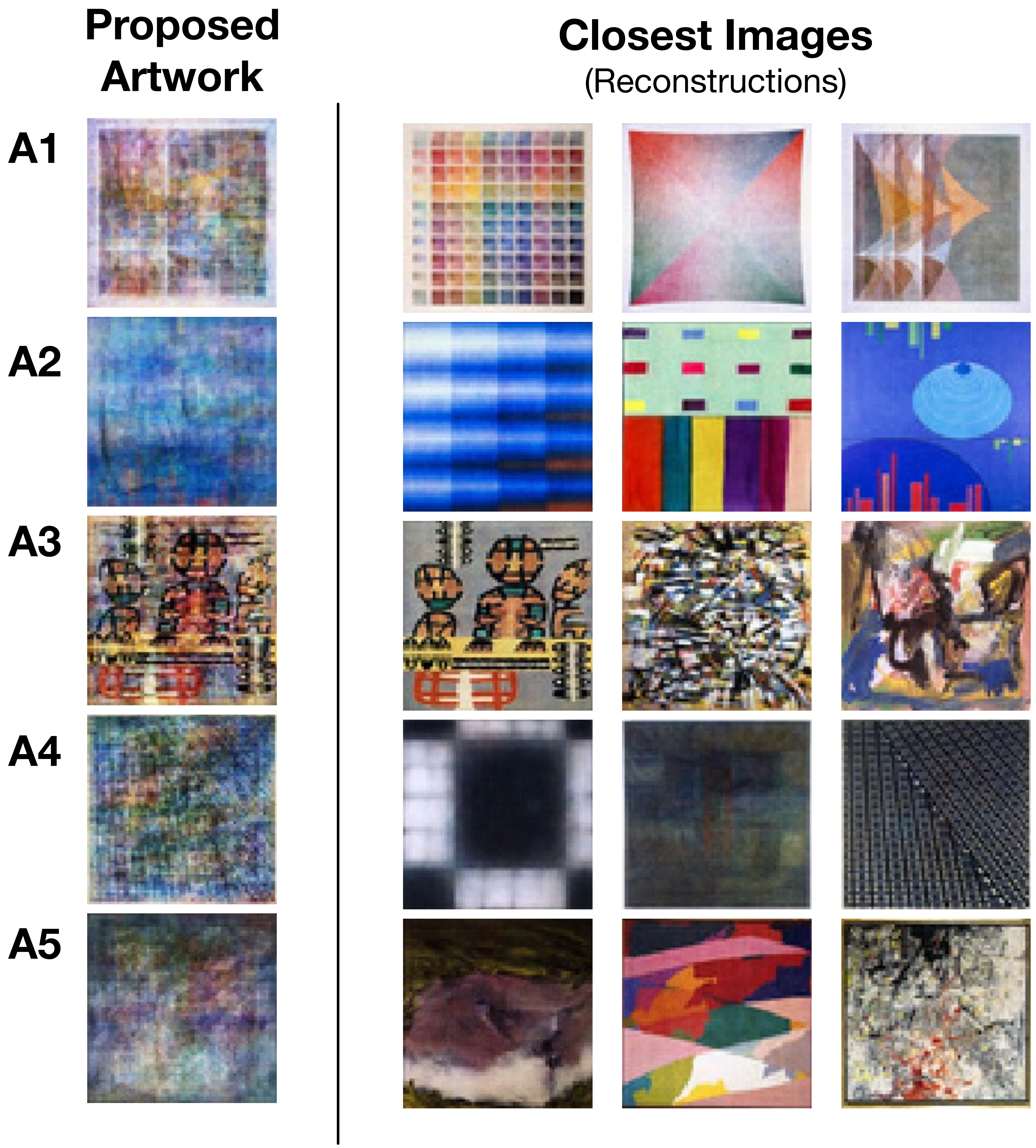
Generated Movie 2
PredR(all): 0.84±0.1, PredR(> τ): 0.89±0.05

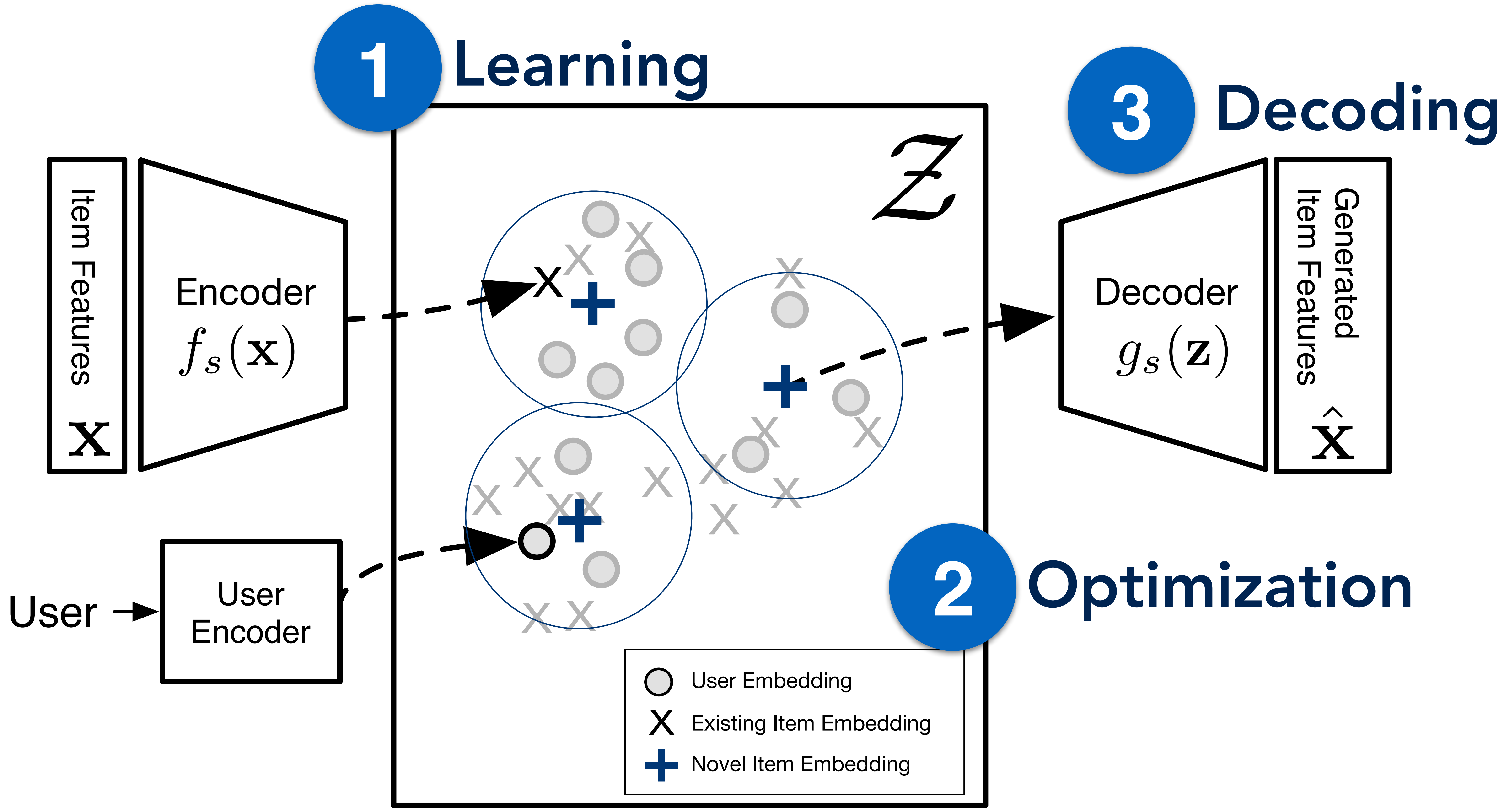


Generated Movie 3
PredR(all): 0.72±0.16, PredR(> τ): 0.88±0.05



Generated Abstract Art




1 Learning
Item Features
x
Encoder
$f_s(\mathbf{x})$
User
User Encoder
$\mathcal{Z}$
2 Optimization
3 Decoding
Decoder
$g_s(\mathbf{z})$
Generated Item Features
$\hat{\mathbf{x}}$
User Embedding
Existing Item Embedding
Novel Item Embedding

Future Work: What's Next?

Future Work: What's Next?

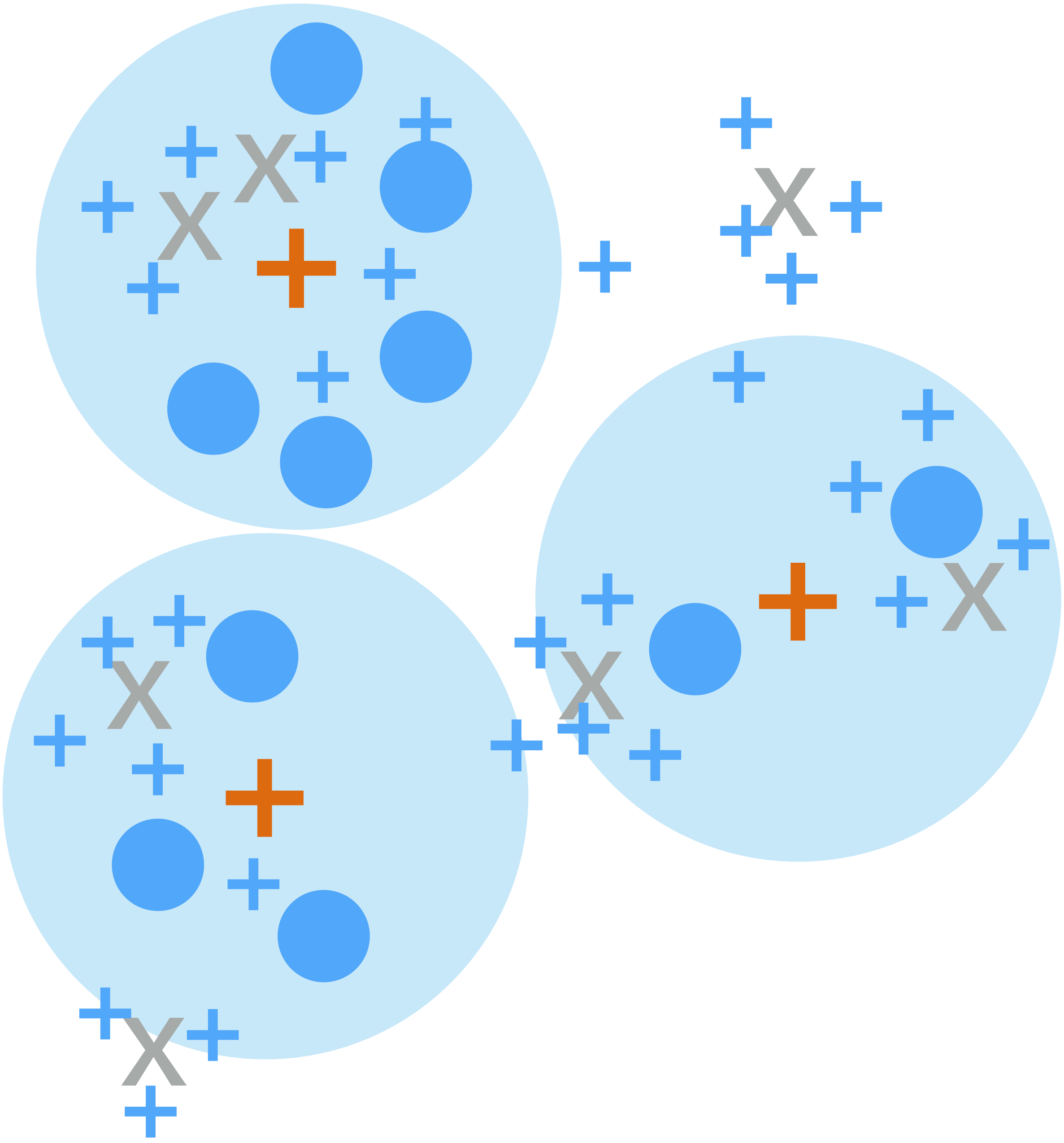
Alternative Problem Formalizations

Domain-specific Item Generation

Collaborative Work with Human-in-the-loop

Future Work: What's Next?

Alternative Problem Formalizations



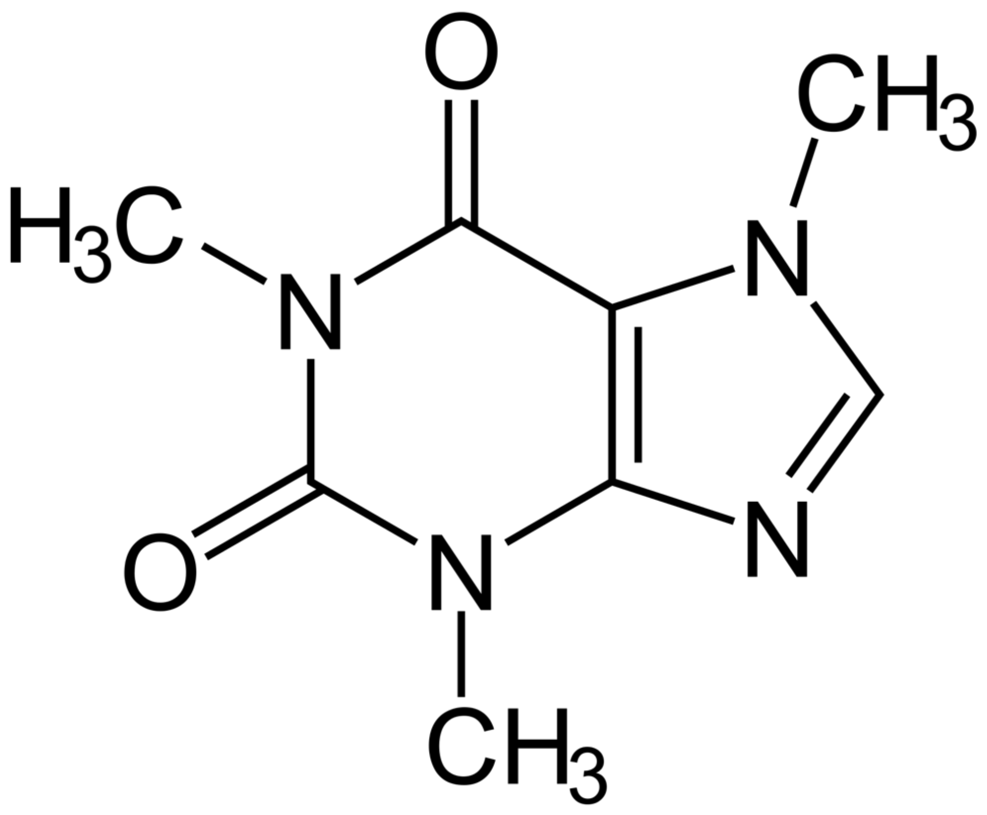
$$\boxed{\operatorname{argmax}_{\mathcal{V}} \sum_{u\in\mathcal{U}} \sum_{v\in\mathcal{V}} \hat{r}(\mathbf{z}_u, \mathbf{z}_v) c_{uv}}$$

such that:

$$\sum_{v\in\mathcal{V}} c_{uv} = 1 \text{ and } |\mathcal{V}| = k$$

Future Work: What's Next?

Domain-specific Item Generation



Future Work: What's Next?

[Ansari & Soh, arXiv 2018]

Collaborative Work with Human-in-the-loop

Thank you. Questions?

Alternative Problem Formalizations

Domain-specific Item Generation

Collaborative Work with Human-in-the-loop

Generation Meets Recommendation:

Proposing Novel Items for Groups of Users

Thanh Vinh Vo and Harold Soh
{vinhvt, harold}@comp.nus.edu.sg

NUS National University of Singapore

If you're in a hurry (TL;DR)

- We address the problem of ***recommendation for item makers*** instead of consumers: *how can we generate a set of **new** items that appeals to a large number of people?*
- We present a problem formalization based on ***learned latent real-vector spaces.***

Experiments & Results

- Validated model on a 20-dimensional synthetic dataset.
- **Key Question:** Can we recreate the missing items?
- **Key Results:** Yes, to a good