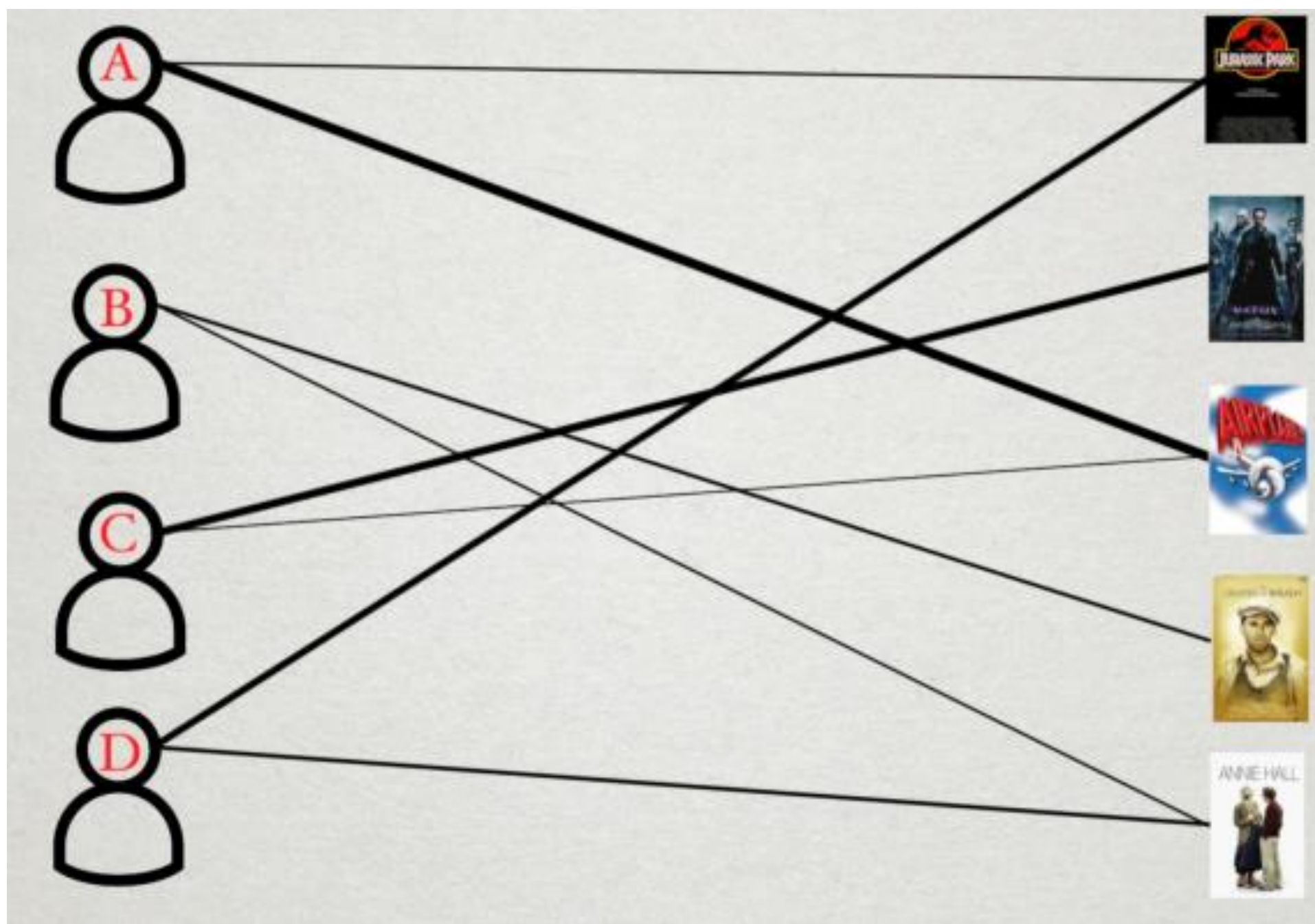


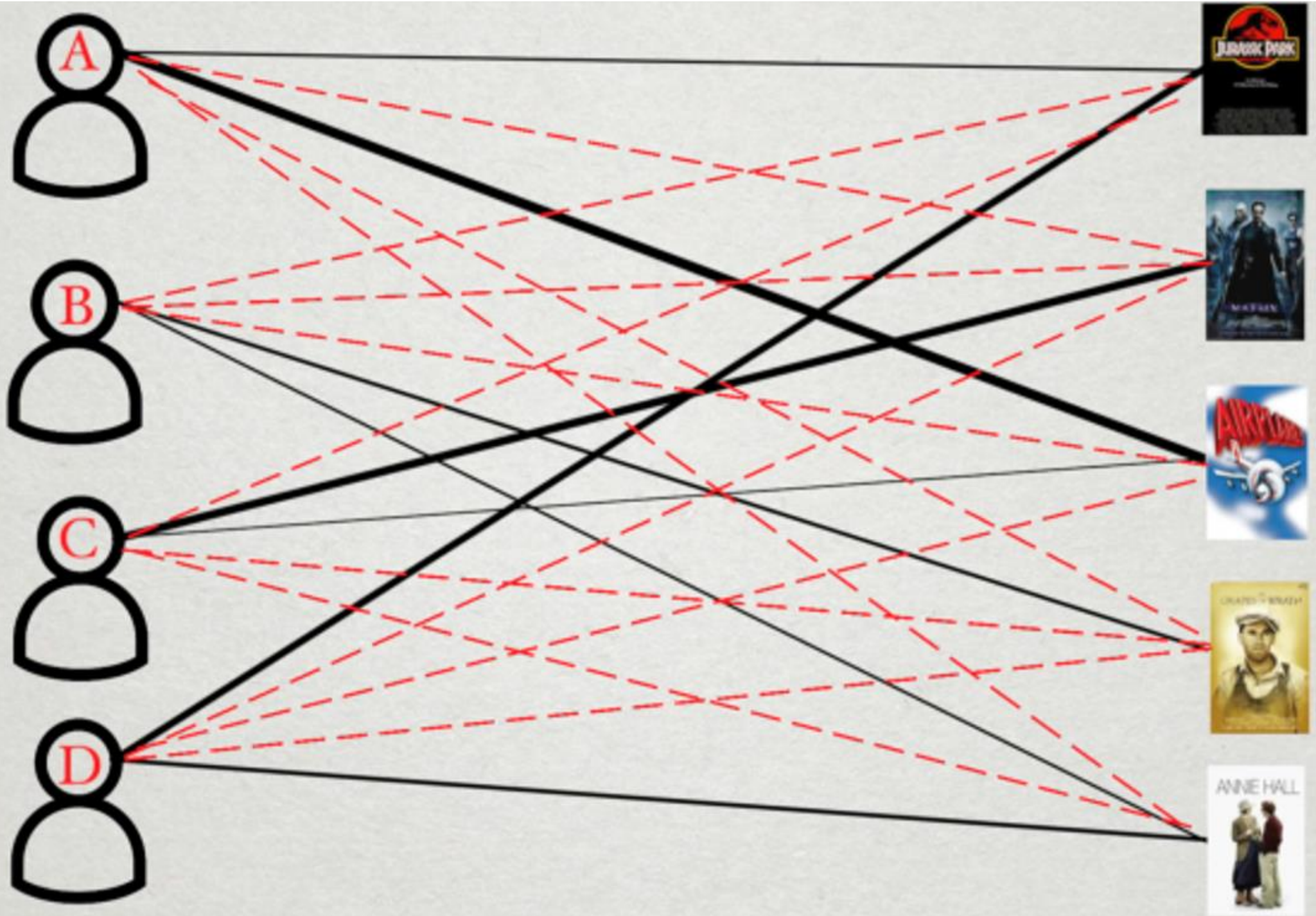
Content and Collaborative Filtering

Alfonso Meléndez

Netflix



Completar



Forma
Matricial

					
A					
B					
C					
D					

Ratings
Numéricos



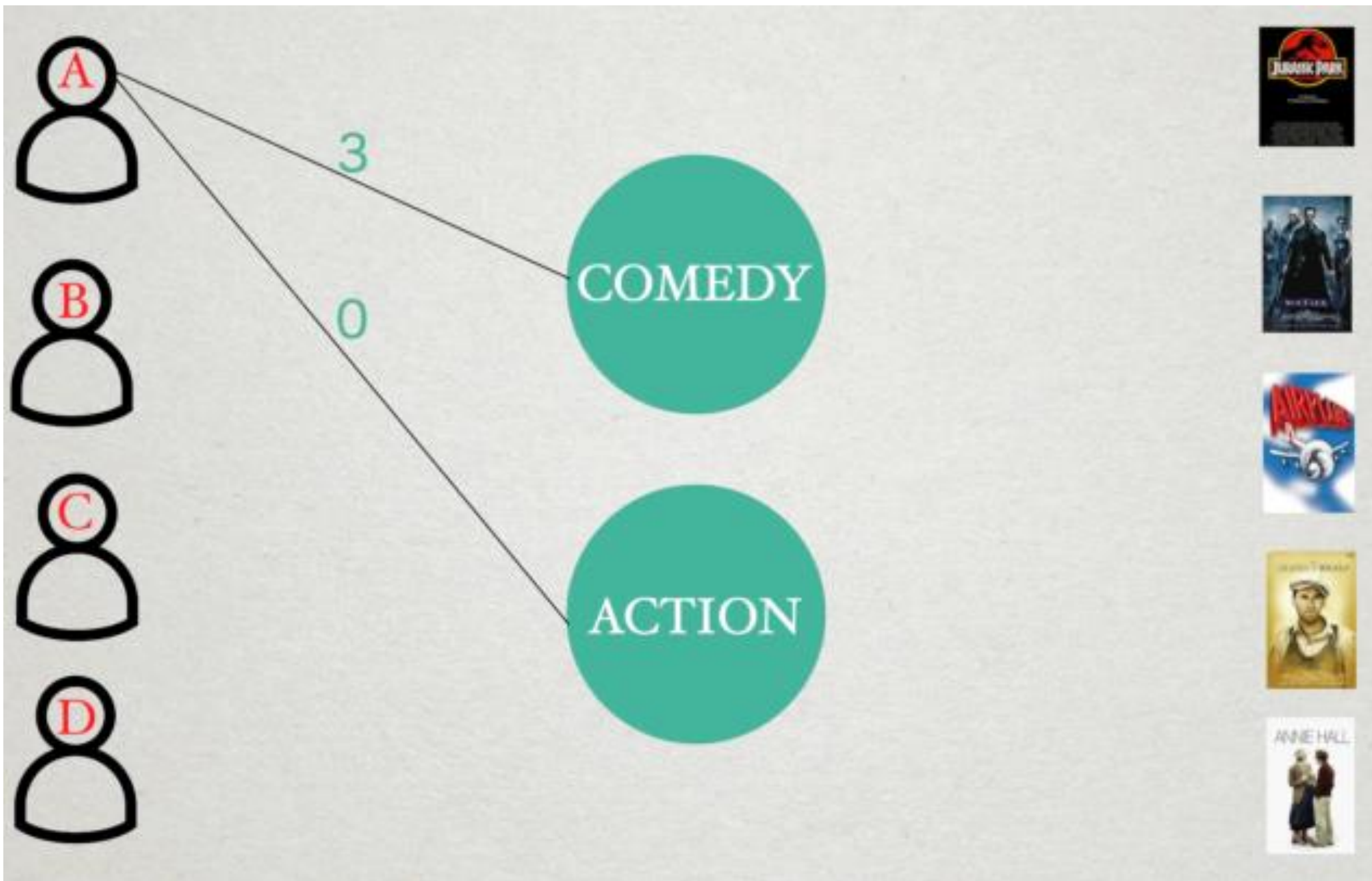
?	4	?	0	1
3	1	2	?	?
?	2	3	1	?
?	0	?	4	?

Content Filtering

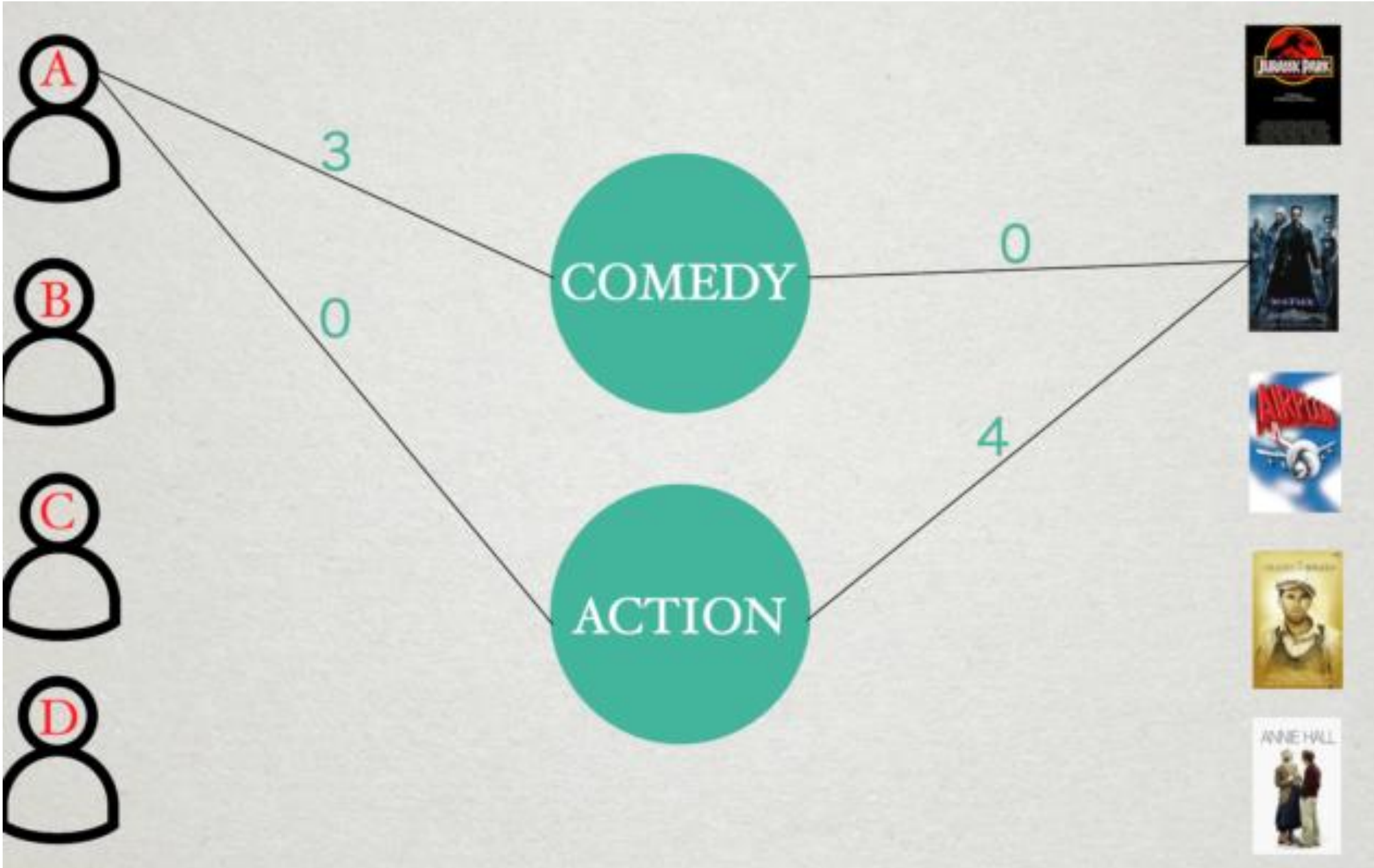
Categorías de Contenido



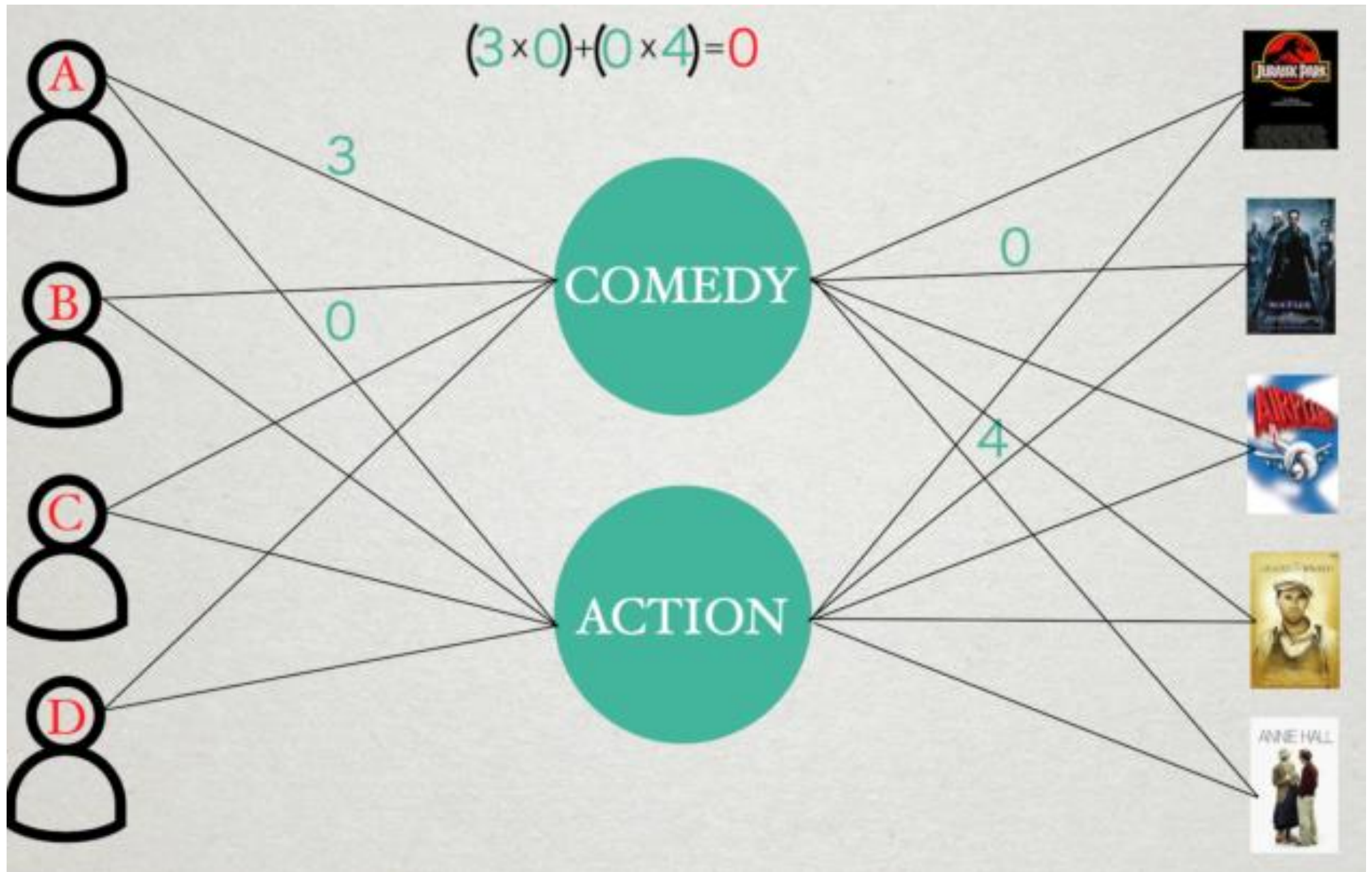
Preferencias
usuarios

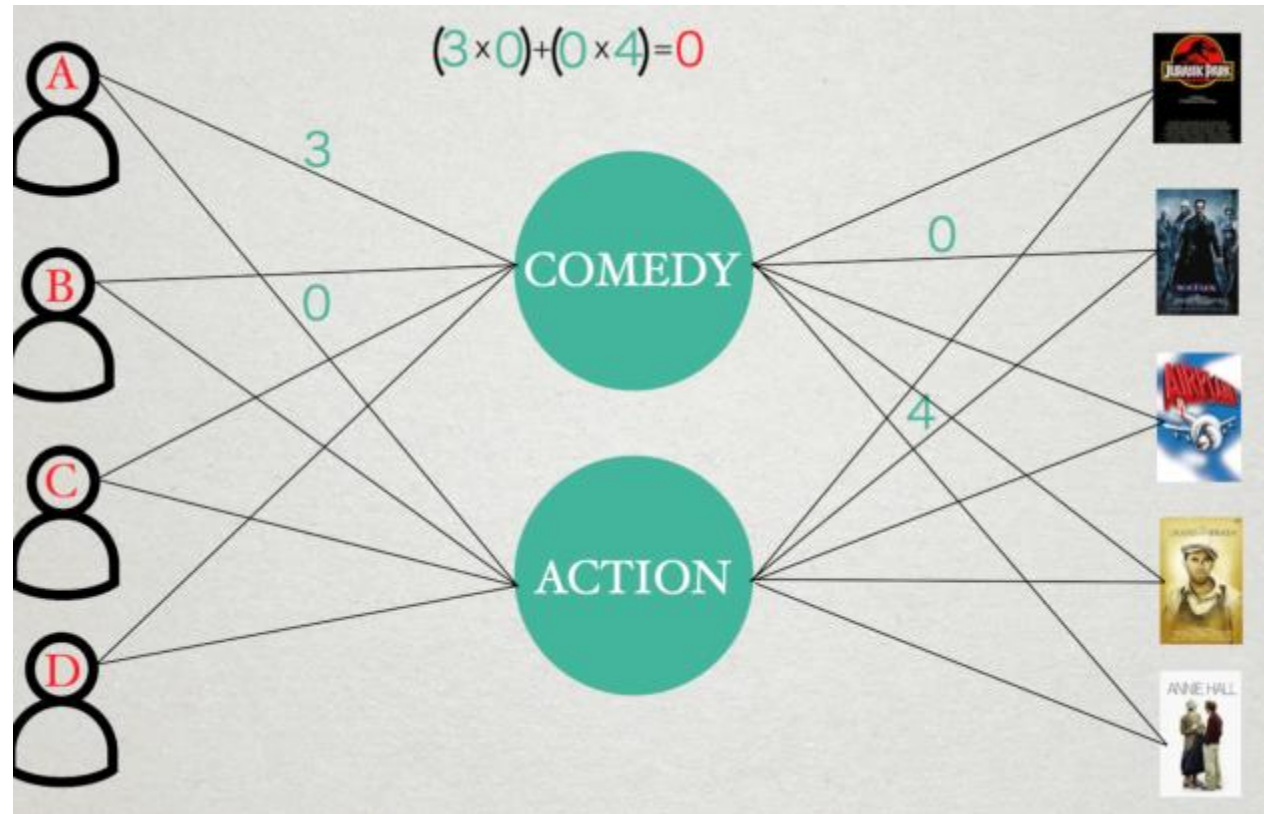


Tipos de película



Rating





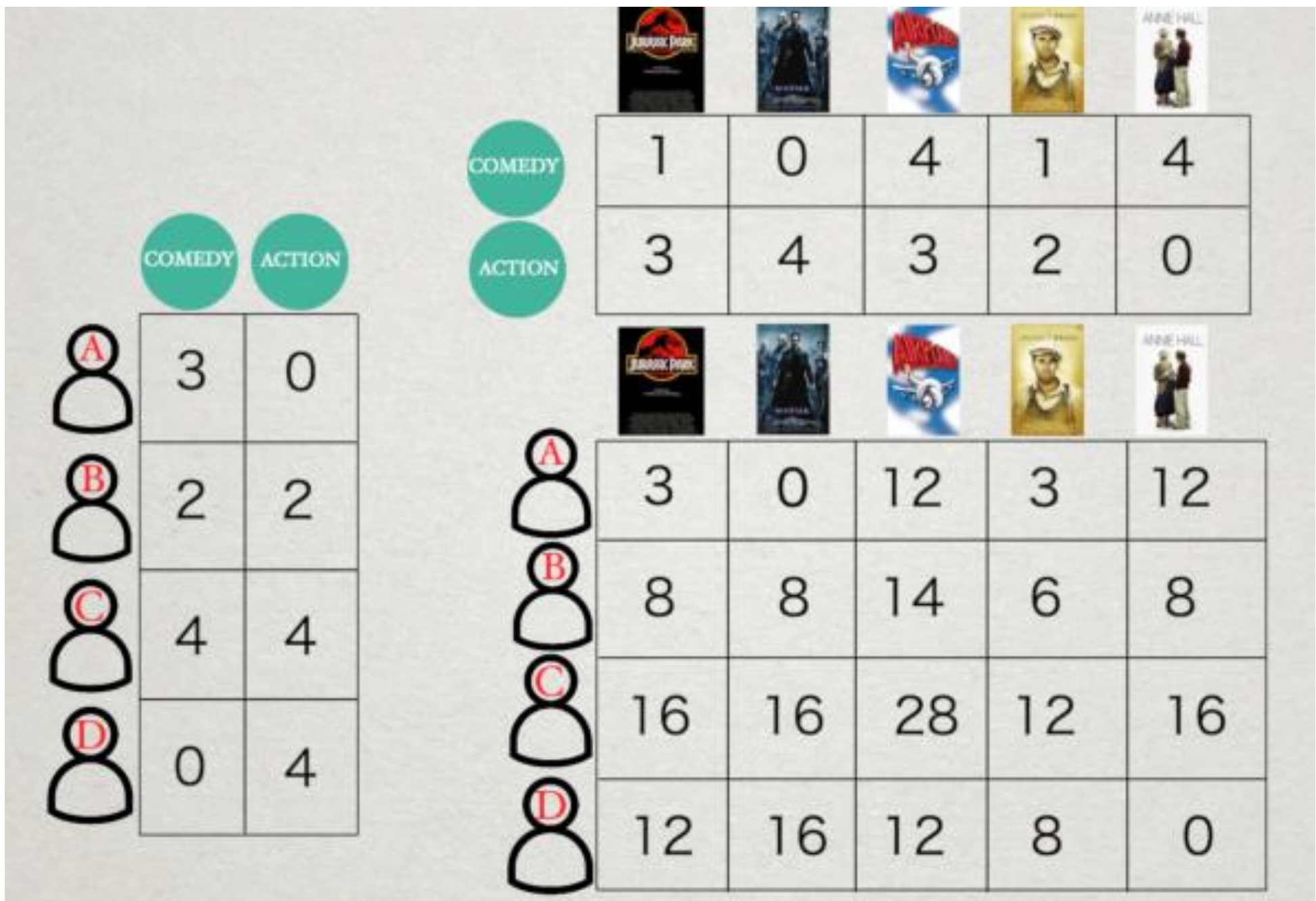
Matrices de Preferencias

	COMEDY	ACTION
A	3	0
B	2	2
C	4	4
D	0	4

COMEDY
ACTION

	JURASSIC PARK	THE DARK KNIGHT TRILOGY	THE SIMPSONS	THE MONUMENTS MEN	THE WEDDING SINGER
COMEDY	1	0	4	1	4
ACTION	3	4	3	2	0

Reconstrucción de la Matriz



Matriz Normalizada
(dividir por 8)

	COMEDY	ACTION
A	3	0
B	2	2
C	4	4
D	0	4

COMEDY
ACTION

	Jurassic Park	Star Wars	Indiana Jones	Amélie	When Harry Met Sally
COMEDY	1	0	4	1	4
ACTION	3	4	3	2	0

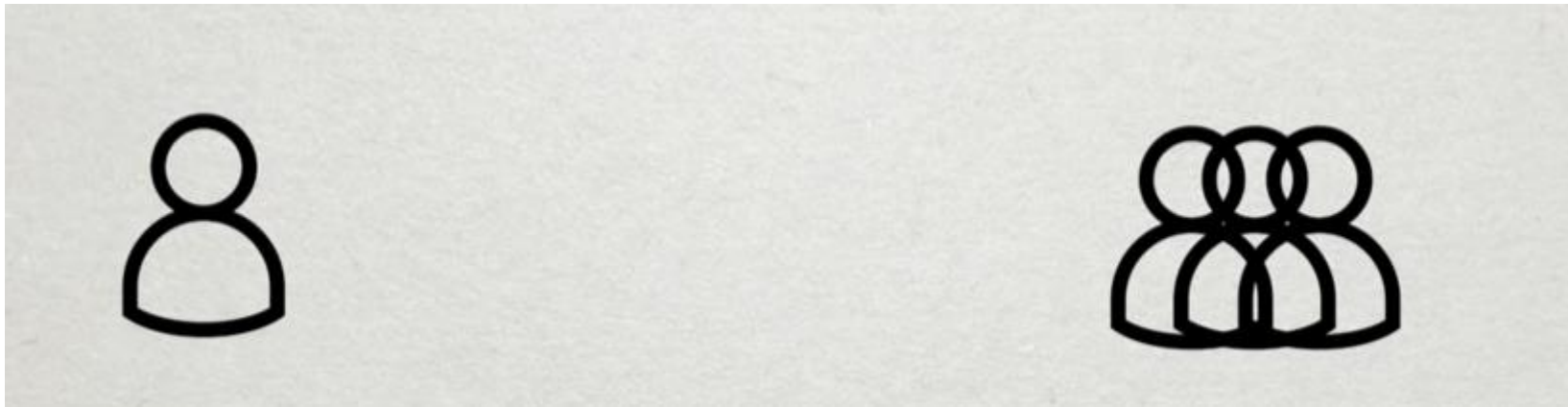
	Jurassic Park	Star Wars	Indiana Jones	Amélie	When Harry Met Sally
A	0	0	1.5	0	1.5
B	1	1	2	1	1
C	2	2	3.5	3	2
D	1.5	2	1.5	1	0

Observaciones

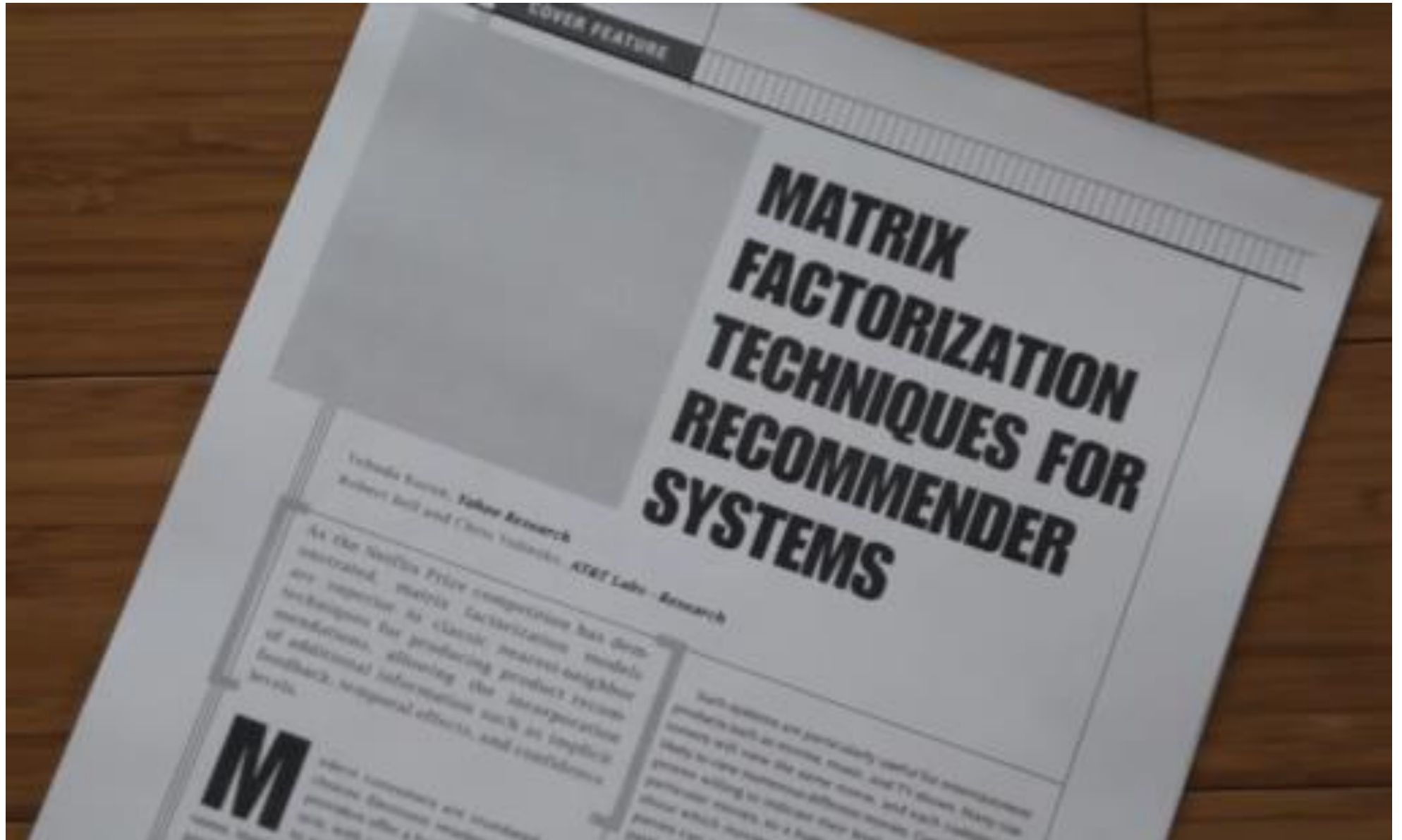
- Es simple
- No es muy exacto (Pocas Categorías)
- Mejora → Más Categorías
- Obtener la información es muy difícil:
 - Personas (Encuestas: comedia,acción)
 - Películas(¿de qué tipo es?)

Collaborative Filtering

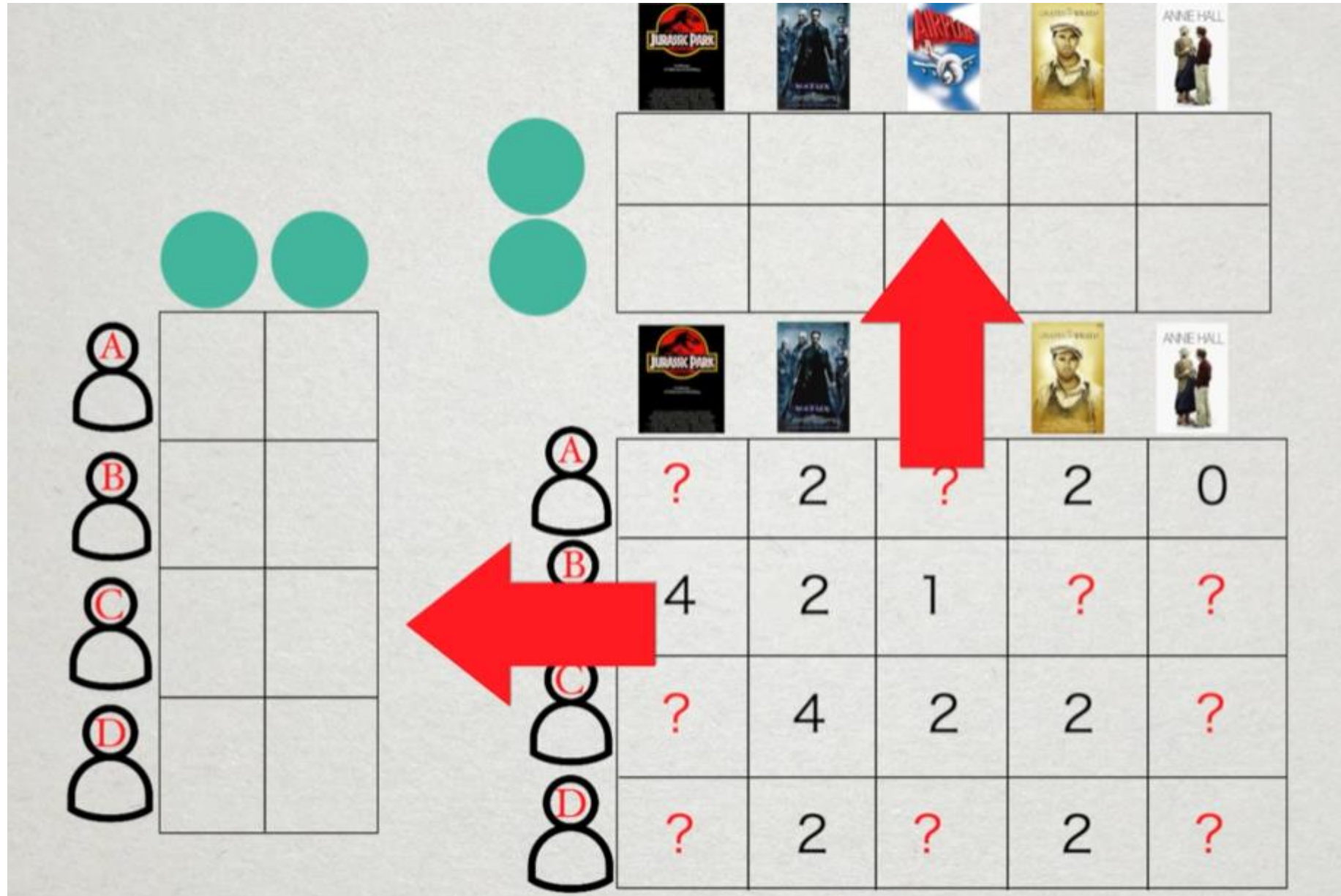
Idea: a Ud le gusta una película si a personas
“similares” les gusta



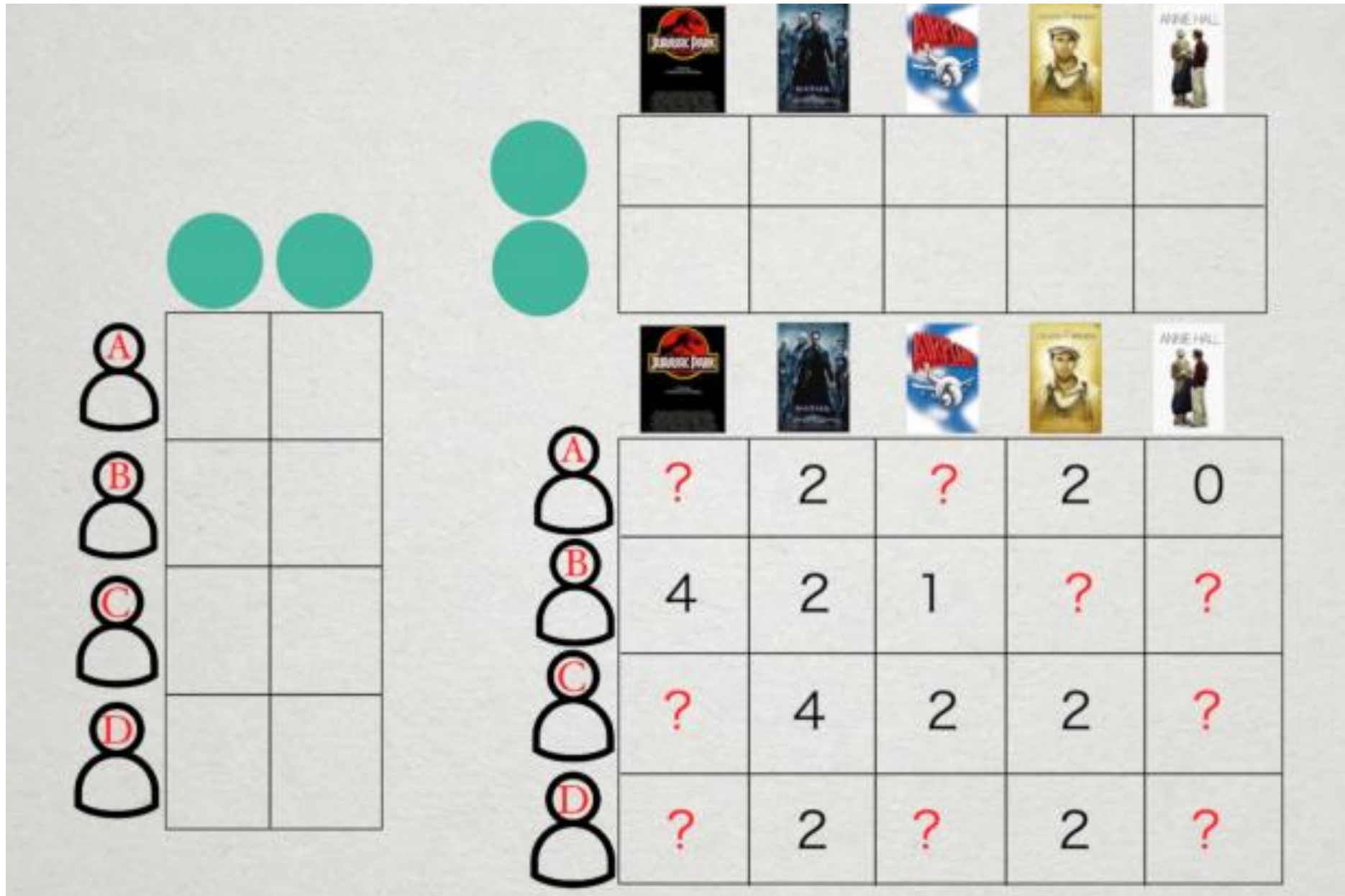
Artículo
Original



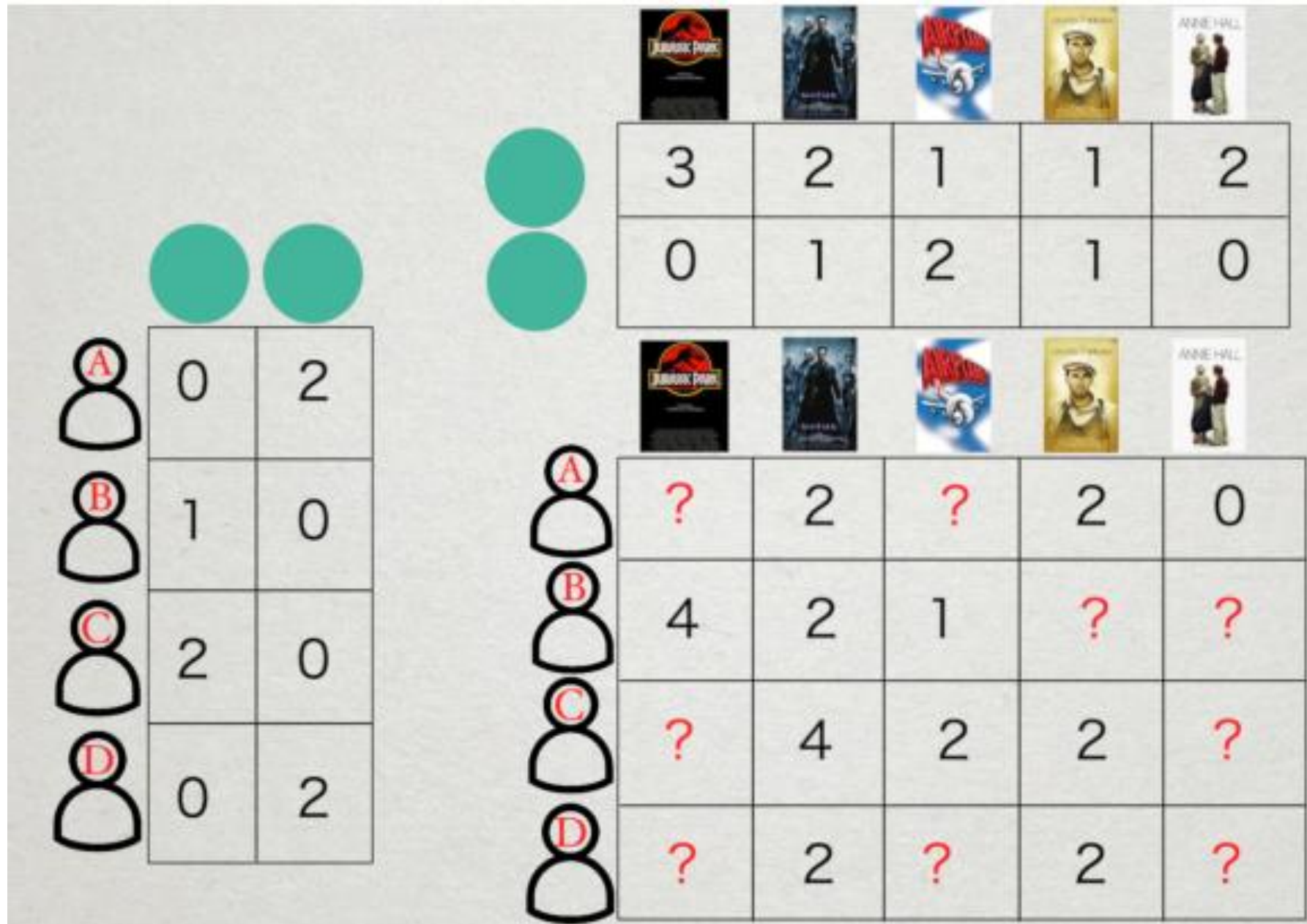
De la matriz a las categorías



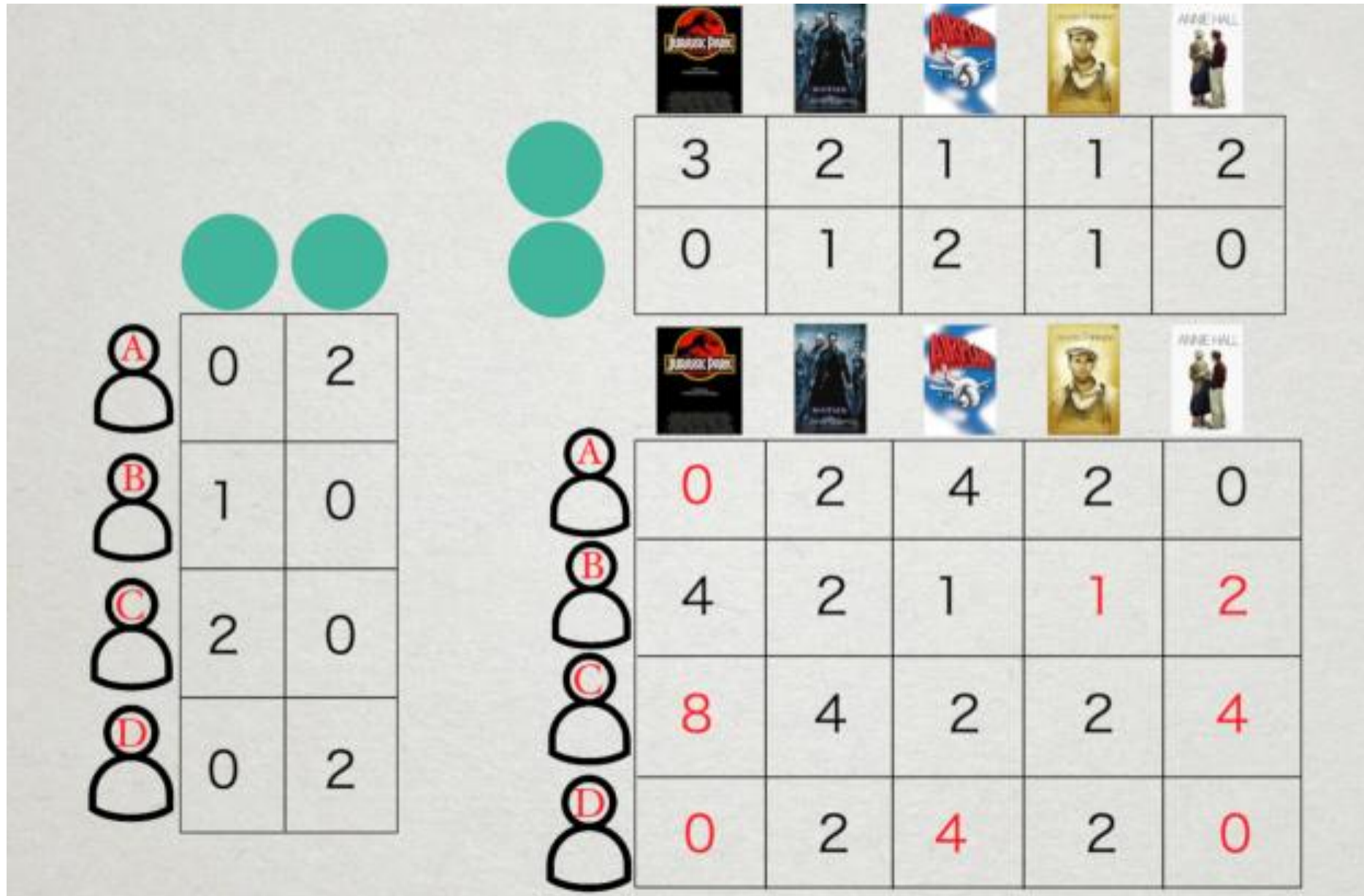
Idea: Optimizar factorización, minimizando el error (Machine Learning)



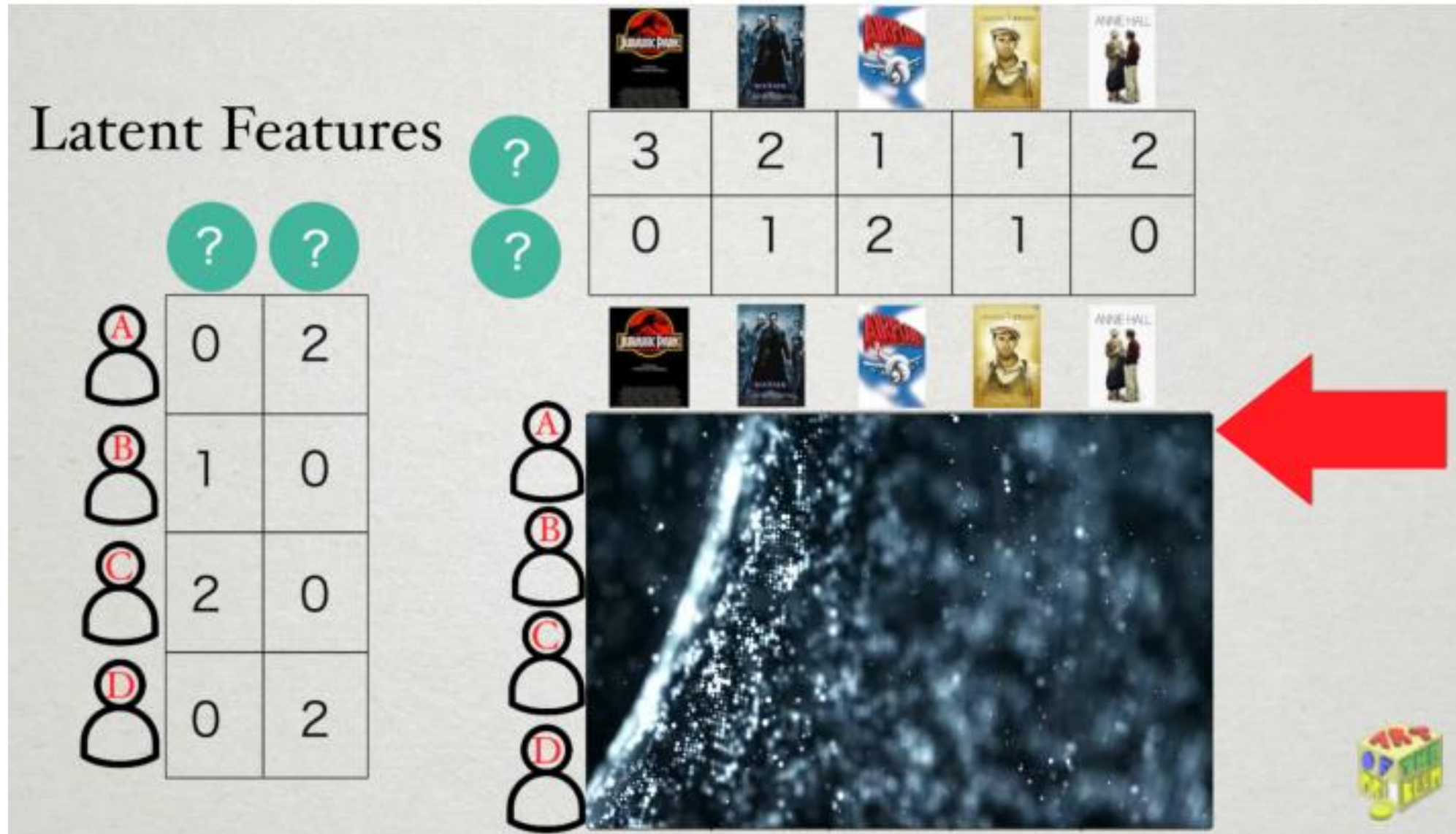
Factorización óptima



Ahora, se completa la matriz



En el fondo es encontrar patrones en los datos



Content Filtering



Collaborative Filtering



Es una forma de compresión de datos

